



Trinity College Dublin

Coláiste na Tríonóide, Baile Átha Cliath

The University of Dublin

PATHOGENIC PROTEIN AGGREGATION *IN VITRO AND IN VIVO*

Thesis submitted to the

University of Dublin

for the degree of

Doctor in Philosophy

Daniel Linus Fortunati

Supervisors:

Prof. Mani Ramaswami

School of Genetics and Microbiology, Trinity College Dublin

Prof. Kenneth Hun Mok

School of Biochemistry and Immunology, Trinity College Dublin

March 2022

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 my original research work. Wherever contributions of others were involved, every effort was made to indicate this clearly, with due reference to the literature, and acknowledgment of collaborative research and discussions.

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).

Daniel Fortunati

March 2022

SUMMARY

Neurodegenerative disorders (ND) such as Alzheimer's disease (AD), Parkinson's disease (PD) and amyotrophic lateral sclerosis (ALS) are one of the leading causes of death and disabilities worldwide. While their aetiologies and clinical manifestations can vary wildly, they are all characterized by the progressive loss of specific neuronal populations. The affected area of the brain or neuronal sub-type determines the primary clinical features of the disease. A common hallmark of most NDs is the presence of misfolded protein aggregates, which are thought to be drivers of disease progression; experimental evidence exists suggesting these misfolded aggregates are capable of spreading from cell to cell, propagating the disease phenotype in a prion-like manner.

An example where this may be occurring are Lewy body disorders (LBs), which are correlated to the aggregation of the protein α -synuclein (α SN); recent findings suggest that intervention on the early stages of aggregation may be necessary to alter disease progression. There is evidence that early-stage α SN oligomers rather than amyloid fibrils are responsible for cytotoxicity, but the detailed structural characterization of these oligomers has only occurred in recent years, and experimentally following their build-up in real-time has proven elusive.

In the first chapter of this work, the formation of pathophysiologically relevant α SN oligomers prepared at physiological conditions was monitored using a battery of physico-chemical techniques including size-exclusion high performance liquid chromatography (SE-HPLC), atomic force microscopy (AFM), and diffusion nuclear magnetic resonance spectroscopy (DOSY NMR). The biological activity of the oligomers produced was assayed by testing primary cortical neuron cytotoxicity and *in vivo* electrophysiological effects on rats. Using time-resolved techniques, the early stages of α SN oligomer growth were quantitatively monitored as individual components within a mixture. The early aggregation pathway of α -synuclein was then compared with that of the non-disease related protein hen egg white lysozyme (HEWL). In general, the ability to kinetically monitor these aggregates without the need for fluorescent markers and/or other modifications shows promise towards the study of proteins correlated with amyloidosis.

In the second chapter, the function of the protein Ataxin-2 (Atxn2) was investigated. The Ataxin-2 (ATXN2) gene is implicated in spinocerebellar ataxia type II (SCA2), amyotrophic lateral sclerosis (ALS) and PD. The encoded protein is a therapeutic target for ALS and other related conditions. Atxn2 (or Atx2 in insects) has been shown to

play a role in most stages of the messenger RNA (mRNA) life cycle, including; translational activation, translational repression, mRNA stability and in the assembly of mRNP-granules, a process mediated by intrinsically disordered regions (IDRs). Building on previous work from the laboratory demonstrating that the LSm (Smith antigen-like) domain of Atx2, which can help stimulate mRNA translation, antagonizes mRNP-granule assembly, a series of experiments was carried out on *Drosophila* and human Ataxin-2 proteins. Results of Targets of RNA-Binding Proteins Identified by Editing (TRIBE), co-localization and immunoprecipitation experiments indicated that a Poly-A-binding protein (PABP) interacting, PAM2 motif of Ataxin-2 may be a major determinant of the mRNA and protein content of Ataxin-2 mRNP granules. Co-immunoprecipitation experiments demonstrated that the Atxn2 LSm domain interacts with its binding partner LSM12 through a canonical LSm:LSm interaction. Transgenic experiments in *Drosophila* indicated that while the Atx2-LSm domain may protect against neurodegeneration, the structured PAM2- and unstructured IDR- interactions both support Atx2-induced cytotoxicity. Point mutations targeting both the Atxn2 PAM2:PABP interaction and the Atxn2 LSm:LSM12 interactions were designed and shown to successfully inhibit these associations, a promising finding from a drug-design perspective. Finally, preliminary experiments were carried out on the sequence-specific effect of the IDRs from Atxn2, Atx2 and ataxin-2-like (Atxn2L) on granule condensation, showing that these IDRs are not functionally equivalent and have different condensation propensities. Taken together the data obtained leads 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.

AKNOWLEDGEMENTS

I was originally planning on keeping this short, but there are a lot of people I really need to thank. I think it is appropriate, as without their help and support I would not have managed to reach the point of actually submitting this thesis. I hope this does not come off as overly sentimental.

First off, I would like to thank my supervisors: Ken, I took my first steps in a laboratory under your guidance, and you taught me the basics of proper scientific thinking; for that, I will always be grateful. Mani, I really cannot thank you enough for everything, I feel like I have grown an astonishing amount in your lab both technically and mentally. My approach to reading papers has changed completely and you have guided me towards more precise and critical scientific thinking. Although I still have a long way to go, I feel like I have made significant progress in that direction. I am sorry I was not always at my best during my time in your lab.

To the rest of the members of the Mok lab, JiYoon, Pancham, Anna and SoYoung: Thanks for everything you have taught me in the lab, and for always being kind and helpful; JiYoon, thank you and Ashton for being such nice and positive friends to me.

To the wonderful post docs from the Ramaswami lab, Joern, Marcia and Jens: Thanks for teaching and helping me with both practical and theoretical scientific problems, and being patient with me during the (many, many) times I made mistakes or messed up.

To my fellow PhDs (friends) from the Ramaswami lab: Camilla, Arnas, Guillaume, Georgia and Emily, thank you for readily accepting me as a member of your group, helping me with scientific issues and going on coffee breaks with me, and thanks for all the chats about scientific or less scientific issues. It has been wonderful to work with you all. Thanks also to Marlena and Zaneta for scientific and non-scientific support in the last couple of years.

To James, Livia and Sarah, thanks for staying by my side and keeping me company during the times when I was really struggling, I will be forever in your debt. To the rest of the reading room crowd in TBSI, thank you for all the pre-COVID collective lunches, banter, scientific advice and occasional board games.

To my friend and housemate Michelangelo thanks for always having my back and cheering me up whenever things got difficult. To Giulio, thank you for helping me with the thesis plan and for all the advice you gave me while I was writing, it was really helpful and saved me a lot of time. To the rest of the Italian group, I really cannot thank you enough: Adalberto, Giorgia, Bartolo, Stefania(s), Carlo, Chiara, Linda and Marco. Thank you for

taking me into your group, you have been like a family to me since the strange COVID-19 isolation times, and I could not have asked for a better one.

To Patrick, Phil, Dan Nelis, Dan Casey, Rory and the rest of the gang: Thanks for all the evenings spent playing board games, all the dinners and lunches and laughs we've had.

To my friends back in Italy (or wherever in the World you are now): Riccardo, Lorenzo, Matteo, Laura, Eugenio and everyone else. Even though we do not see each other much or talk that often, whenever I am back it always feels like no time has passed at all. Thank you for staying in my life.

To the people I have loved during these years, thank you for the time you spent by my side.

Finally, I would like to thank my family: Mom, Dad and Matthew, thank you for always believing in me and supporting me, regardless of how things were going.

If someone were to ask me what I like the most about myself, I would answer "the people I surround myself with". Thank you all for everything, really.

TABLE OF CONTENTS

DECLARATION	3
SUMMARY	5
AKNOWLEDGEMENTS	7
TABLE OF CONTENTS.....	10
LIST OF FIGURES.....	13
LIST OF TABLES.....	16
LIST OF ABBREVIATIONS	17
1. INTRODUCTION.....	21
1.1 Protein misfolding & aggregation in neurodegenerative diseases.....	21
1.2 Liquid-liquid phase separation.....	24
1.3 Stress granules and processing bodies	28
1.4 Neuronal proteostasis and clearance pathways.....	32
2. MATERIALS & METHODS.....	40
2.1 Key resources & Materials	40
2.2 Instrumentation	45
2.3 Protein aggregation protocols	47
2.3.1 Fibril formation (α SN)	47
2.3.2 Oligomer formation for cytotoxicity & animal experiments (α SN).....	47
2.3.3 Oligomer formation for SE-HPLC (α SN)	47
2.3.4 Hen egg white lysozyme aggregation	48
2.4 Size exclusion high performance liquid chromatography (SE-HPLC)	48
2.5 Gel electrophoresis, western blotting & immunoprecipitation.....	48
2.5.1 Gel electrophoresis of α -synuclein	48
2.5.2 Western blotting & Immunoprecipitation of minimized Atxn2 constructs	49
2.5.3 Western blotting of transgenic fly heads.....	50
2.6 Spectroscopic techniques	50
2.6.1 Thioflavin T fluorescence spectrophotometry.....	50
2.6.2 Circular dichroism spectropolarimetry	50
2.6.3 Diffusion-Ordered NMR Spectroscopy (DOSY)	51
2.7 Microscopy techniques.....	53
2.7.1 Atomic force microscopy	53
2.7.2 Transmission electron microscopy	53
2.7.3 Immunohistochemistry and imaging of cultured cells.....	54

2.8 Cell culture.....	54
2.8.1 Isolation of primary cortical neurons	54
2.8.2 Cell culture, transfection and stress induction of S2R+, HEK293T and U2OS cells	55
2.8.2 Cell Viability	55
2.9 <i>In vivo</i> electrophysiology	56
2.10 Statistical treatment of the data	56
3. TIME RESOLVED MONITORING OF α -SYNUCLEIN OLIGOMERS.....	58
3.1 Introduction.....	58
3.1.1 Parkinson's disease.....	58
3.1.2 Lewy bodies and α -synuclein aggregates	61
3.1.3 The role of pathology spreading & selective vulnerability in PD	65
3.1.4 Why study α SN oligomers?	67
3.2 Results	70
3.2.1 Characterization of α -synuclein monomer and aggregates	70
3.2.2 Time-resolved monitoring of α -synuclein aggregation	74
3.2.3 Time-resolved monitoring of the aggregation of a non-disease related protein.....	78
3.2.4 Cytotoxicity and electrophysiological effects of the α -synuclein aggregates	81
3.2.5 Size and morphology of aggregating protein samples using atomic force microscopy	83
3.2.6 DOSY-based size estimations of aggregating samples	87
3.3 Discussion	96
4. DIVERSE ROLES OF STRUCTURED & DISORDERED REGIONS OF ATAXIN-2 IN mRNP GRANULE FORMATION	100
4.1 Introduction.....	100
4.1.1 Spinocerebellar ataxias.....	100
4.1.2 Ataxin-2 and neurodegeneration	102
4.1.3 Atxn2 domains and their functions	106
4.1.4 Atxn2, mRNP granules and disease	109
4.2 Results	111
4.2.1 The structured PAM2 domain of Atx2 is necessary for the correct protein and mRNA content in Atx2 granules	111
4.2.2 PAM2 interactions with PABP are sufficient for Atx2 to associate with stress granules	117
4.2.3 Minimized Atxn2/Atx2 constructs are directed to stress granules <i>via</i> the PAM2:PABP interaction	122
4.2.4 Atxn2 interacts with the protein LSM12 <i>via</i> a canonical LSm-LSm interaction	128
4.2.5 The Atx2/Atxn2/Atxn2L IDRs are not functionally equivalent	131

4.2.6 The IDR and PAM2 domains promote and the LSm domain inhibits cytotoxicity in Drosophila neurodegeneration models.....	135
4.3 Discussion.....	140
4.3.1 Molecular mechanisms of Atxn2 function.....	140
4.3.2 Atxn2 as a therapeutic target	141
4.3.2 Interactions between structured domains may be the main determinant of mRNP composition.	143
5. CONCLUSIONS & OUTLOOK	145
5.1 Conclusions	145
5.2 Outlook	147
6. SUPPLEMENTARY INFORMATION	149
7. BIBLIOGRAPHY	156
8. PUBLICATIONS & PRESENTATIONS	183
9. APPENDIX.....	184

LIST OF FIGURES

Figure 1.1: Structure of amyloid fibril from A β (1 - 42) peptide.....	22
Figure 1.2: The seeding-nucleation model of protein aggregation and aggregate spread.	23
Figure 1.3: Thermodynamic principles of phase separation.....	25
Figure 1.4: Summary of the diverse protein-protein interactions driving condensation of RNP granules.	26
Figure 1.5: Partial list of proteins found in stress granules, processing bodies, or both.	29
Figure 1.6: Change in conformation of the chaperone HSP90 upon substrate/co-factor binding.	33
Figure 1.7: Structure of the 20S (uncapped) proteasome core catalytic particle. ..	34
Figure 3.1: Clinical progression of Parkinson's disease.....	59
Figure 3.2: Schematic diagram showing different cellular processes implicated with Parkinson's pathogenesis.	60
Figure 3.3: Domains and structural features of α SN.	62
Figure 3.4: Lewy body stainings in neurons of the nigrostriatal pathway from PD patients.	64
Figure 3.5: Summary of putative PD vulnerability determinants of dopaminergic neurons from the <i>substantia nigra pars compacta</i>	66
Figure 3.6 Fibrillation of α -synuclein.	70
Figure 3.7: SE-HPLC standard curve for protein size determination.....	71
Figure 3.8: Chromatogram of native α SN.	72
Figure 3.9: Far-UV circular dichroism spectropolarimetry spectra of monomeric and oligomeric α SN.....	73
Figure 3.10: SDS PAGE analysis of α SN sample.	74
Figure 3.11: Chromatogram of oligomeric aggregates of α SN.	75
Figure 3.12: Time resolved monitoring of early α SN aggregation.	77
Figure 3.13: Early aggregation of HEWL.	78
Figure 3.14: Aggregating HEWL peak deconvolution over time.	79
Figure 3.15: Quantification of HEWL peak components at different time points....	80
Figure 3.16: Cytotoxic effect of α SN oligomers on primary cortical neurons.....	81
Figure 3.17: Effects of α SN oligomers on baseline synaptic transmission in rats.	82
Figure 3.18: AFM imaging of α SN following the full oligomerization protocol.	83
Figure 3.19: Size distribution of α SN aggregates imaged in liquid.....	84
Figure 3.20: TEM imaging of α SN aggregates.....	85

Figure 3.21: AFM imaging of HEWL aggregation at different time points.....	86
Figure 3.22: Aggregating HEWL sample composition at different time points.	87
Figure 3.23: Diffusion projection of NMR spectrum from oligomerized α SN.....	88
Figure 3.24: Oligomerized α SN DOSY NMR measurements.	89
Figure 3.25: Single component fitting of diffusion projections from native/aggregating HEWL.....	91
Figure 3.26: Differences between protein R_H values obtained during the DOSY experiments.....	92
Figure 3.27: Three-component poly-gaussian fit of aggregating HEWL DOSY results.	93
Figure 3.28: Results of DOSY NMR experiment on a mixture of small molecules. .	94
Figure 4.1: Molecular determinants of pathology observed through different types of spinocerebellar ataxias.	101
Figure 4.2: Domain structure schematic of Ataxin-2, human and fly.	104
Figure 4.3: Atxn2 can have different functions depending on the mRNA transcript/protein complex it is associated with.....	105
Figure 4.4: Structure of LSm domain and mechanism of LSm:LSm binding.	107
Figure 4.5: Ataxin-2 interacts with PABPC1 <i>via</i> the PAM2:MLLE interaction.	108
Figure 4.6: Summary of the TRIBE analysis pipeline and ADARcd constructs used.	112
Figure 4.7: The PAM2 domain plays a directing role in Atx2 RNA target selection.	114
Figure 4.8: The PAM2 domain affects protein composition of Atx2-GFP overexpression induced granules in S2 cells.	116
Figure 4.9: Schematic of minimized ataxin-2 constructs.	117
Figure 4.10: The Ataxin-2 PAM2 domain is conserved across human and fly, as is the PABP MLLE domain.	119
Figure 4.11: SNAP co-immunoprecipitation experiment carried out with wild-type and domain deleted minimized constructs.	120
Figure 4.12: Single point mutations in key residues of the PAM2 domain are sufficient to replicate the effect of a full PAM2 deletion.....	121
Figure 4.13: Mini Atx2 is recruited to SGs upon oxidative stress in S2 cells.	123
Figure 4.14: Mini Atxn2 is recruited to SGs upon oxidative stress in U2OS cells.	124
Figure 4.15: Deletion of the PAM2 domain, but not of the LSm and LSm-AD domains prevent mini Atx2 SG recruitment in S2 cells.	125
Figure 4.16: Deletion of the PAM2 domain, but not of the LSm and LSm-AD domains prevent mini Atx2 SG recruitment in U2OS cells.	126

Figure 4.17: Point mutants in the PAM2 domain prevent fly mini Atx2 from being recruited to oxidative stress induced SGs in S2 cells.....	127
Figure 4.18: Point mutants in the PAM2 domain prevent human mini Atxn2 from being recruited to oxidative stress induced SGs in U2OS cells.	128
Figure 4.19: LSM protein alignment and mutant design.....	129
Figure 4.20: Point mutations in the Ataxin-2 LSm domain are sufficient to prevent the mini Ataxin-2:LSM12 interaction.....	130
Figure 4.21: Lsm14a is not pulled down by mini Atxn2, regardless of the presence of the LSm domain.....	131
Figure 4.22: Human Atxn2/Atxn2L:Fly Atx2 structured domain swaps.....	132
Figure 4.23: Construct with human Atxn2 structured domains and fly Atx2 IDRS (HSDI) expressed in U2OS cells, stained for PABPC1 and wild type Ataxin-2.	133
Figure 4.24: Construct with human Atxn2 structured domains and fly Atx2 IDRS (HSDI) expressed in U2OS cells, stained for G3BP and LSM12.....	133
Figure 4.25: Construct with <i>Drosophila</i> Atxn2 structured domains and human Atxn2 IDRS (DSHI-Atxn2) expressed in S2 cells, stained for SG markers Caprin and dFMRP.....	134
Figure 4.26: Construct with <i>Drosophila</i> Atxn2L structured domains and human Atxn2 IDRS (DSHI-Atxn2L) expressed in S2 cells, stained for SG markers Caprin and dFMRP.....	135
Figure 4.27: The IDR and the PAM2 domains promote while the LSm domain inhibits neurodegeneration in <i>Drosophila</i>	138
Figure 4.28: Fly head western blots to compare protein expression levels/ensure temperature-dependent expression.....	139
Figure 6.1: SE-HPLC chromatogram of α SN freshly dissolved in 8 M Urea.	149
Figure 6.2: HEWL aggregation progression monitored using thioflavin T fluorescence spectrophotometry.....	150

LIST OF TABLES

Table 2.1: Key resources & materials	40
Table 3.1: Hydrodynamic radii of proteins used for the standard curve	72
Table 3.2: Validation of HEWL deconvolution approach using a mixture of small molecules.	95
Table 6.1: Hydrodynamic radi and elution volumes of standard curve proteins as obtain upon injection using 8 M urea as a mobile phase	149
Table 6.2: Mapping percentages of TRIBE replicates	150
Table 6.3: Targets obtained from TRIBE experiments	151

LIST OF ABBREVIATIONS

aa	Amino acid
Ab	Antibody
AD	Alzheimer's disease
ADAR	Adenosine deaminase acting on RNA
ADARcd	ADAR catalytic domain
ADP	Adenosine di-phosphate
AFM	Atomic force microscopy
Ago	Argonaute
Ala	Alanine
ALS	Amyotrophic Lateral Sclerosis
APP	Amyloid precursor peptide
Asn	Asparagine
ASO	Antisense oligonucleotide
ATF6	Activating transcription factor 6
ATG	Autophagy related gene
ATP	Adenosine triphosphate
Atx2	Ataxin-2 protein (<i>Drosophila</i>)
Atxn2	Ataxin-2 protein (Human)
ATXN2	Ataxin-2 gene
Aβ	Amyloid- β
BSA	Bovine serum albumin
CD	Circular dichroism
CMA	Chaperone-mediated autophagy
CP	Catalytic core particle
CSF	Cerebrospinal fluid
Cyt C	Cytochrome C
DA	Dopamine
DCP1	mRNA-decapping enzyme 1
DDX6	ATP-dependent RNA helicase DDX6
dFMRp	<i>Drosophila</i> Fragile X mental retardation protein
DLS	Dynamic light scattering
DMEM	Dulbecco modified Eagle media
DNA	Deoxyribonucleic acid
DOSY	Diffusion-ordered-spectroscopy
DTT	dithiothreitol
EDC3	Enhancer of decapping protein 3
EDTA	Ethylene-di-amine-tetra-acetic acid
eIF3	Eukaryotic translation initiation factor 3
eIF5	Eukaryotic translation initiation factor 5
eIF4F	Eukaryotic translation initiation factor 4F
eIF4G	Eukaryotic translation initiation factor 4G
eIF2α	Eukaryotic translation initiation factor 2 α
EM	Electron microscopy
ER	Endoplasmic reticulum
FCS	Fluorescence correlation spectroscopy
FD	Frontotemporal dementia

FID	Free induction decay
FMRp	Fragile X mental retardation protein
FRAP	Fluorescence recovery after photo-bleaching
FRET	Förster resonance energy transfer
FTD	Frontotemporal dementia
FUS	Fused-in-Sarcoma
G3BP	Ras-GTPase-activating protein (SH3 domain)-binding protein
GCN2	General control non-de-repressible 2
GFP	Green fluorescent protein
Gln	Glutamine
Gly	Glycine
GW182	Glycine-tryptophan protein of 182 kDa
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HEWL	Hen egg white lysozyme
HRI	Heme-regulated inhibitor
HRP	Horseradish peroxidase
HSA	Human serum albumin
HSP	Heat shock protein
IDP	Intrinsically disordered protein
IDR	Intrinsically disordered region
IgG	Immunoglobulin G
IP	Immunoprecipitation
IRE1α	Inositol-requiring protein 1 α
ISF	Interstitial fluid
ISR	Integrated stress response
ISRIB	Integrated stress response inhibitor
kDa	Kilo-Dalton
KH	hRNP K homology
LAS	Lysosomal-autophagy system
LB	Lewy body
LBD	Dementia with Lewy bodies
LC	Low complexity
LDS	Lithium dodecyl sulfate
Leu	Leucine
LLPS	Liquid-liquid phase separation
LSm	Smith antigen-like
LTP	Long term potentiation
mAb	Monoclonal antibody
MALS	Multiangle light scattering
Me31b	ATP-dependent RNA helicase me31b
Mini Atx2	Minimized ataxin-2 (<i>Drosophila</i>)
Mini Atxn2	Minimized ataxin-2 (human)
miRNA	Micro RNA
MRE	Mean residue ellipticity
mRNA	Messenger RNA
mRNP	Messenger ribonucleoprotein
MS	Mass spectrometry
mTOR	Mammalian target of rapamycin
NAADP	Nicotinic acid adenine dinucleotide phosphate

NAC	Non-A β -component
ND	Neurodegenerative disease
NMR	Nuclear magnetic resonance
NTA	Nitrilotriacetate
pAb	Polyclonal antibody
PABP	Poly-A binding protein
PABPC1	Poly-A binding protein cytoplasmic 1
PAGE	Poly-acrylamide gel electrophoresis
PAIP1	Poly-A binding protein interacting protein 1
PAIP2	Poly-A binding protein interacting protein 2
PB	Processing body
PBS	Phosphate buffered saline
PBST	Phosphate buffered saline + Tween
PCR	Polymerase chain reaction
PD	Parkinson's disease
<i>per</i>	<i>Period</i> gene
PERK	PKR-like ER kinase
Phe	Phenylalanine
PIC	Pre-initiation complex
PMSF	Phenylmethylsulfonyl fluoride
Poly-A	Poly-adenosine monophosphate
Poly-Q	Poly-glutamine
PrD	Prion disease
PRK	Protein kinase R
PrP	Prion protein
PrP^{Sc}	Misfolded prion protein
QOL	Quality of life
RBP	RNA-binding protein
RBP7	Retinoid binding protein 7
R_{eq}	Equivalent radius
R_G	Radius of gyration
R_H	Hydrodynamic radius
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RP	Regulatory particle
RRM	RNA recognition motif
SAXS	Small angle X-ray scattering
SCA	Spinocerebellar ataxia
SDS	Sodium dodecyl sulphate
SE-HPLC	Size exclusion high performance liquid chromatography
SG	Stress granule
SLiM	Short linear motif
Sm	Smith antigen
SNCA	α -synuclein gene
SNpc	<i>Substantia nigra pars compacta</i>
SOD1	Superoxide dismutase 1
ssNMR	Solid-state NMR
TDP-43	TAR DNA-binding protein 43
TEM	Transmission electron microscopy
ThG	Thyroglobulin

ThT	Thioflavin T
TIA-1	T-cell intracellular antigen 1
TRIBE	Targets of RNA-binding proteins identified by editing
tRNA	Transfer RNA
Trp	Tryptophan
TYF	Twenty four
XRN1	5' – 3' exonuclease 1
Ub	Ubiquitin
UPR	Unfolded protein response
UPS	Ubiquitin-proteasomal system
UTR	5' untranslated region
WB	Western blot
WT	Wild type
αSN	α -synuclein
β-LG	β -Lactoglobulin
γG	γ -Globulin

1. INTRODUCTION

1.1 Protein misfolding & aggregation in neurodegenerative diseases

Neurodegenerative diseases (NDs) are, by definition, characterized by the progressive loss of specific sub-populations of neurons in the nervous system. The areas/cell types initially affected within the nervous system cause enormous variation in the clinical features of these diseases. Despite their clinical heterogeneity, many common features exist; progressive deterioration with age, damage to the synaptic transmission network, neuronal cell death and consequent loss of brain mass in selected areas. A further common event in all NDs is the misfolding, aggregation and toxic-gain-of-function of different proteins, which is now thought to be a key driver of neurodegeneration (Chiti & Dobson, 2017; Ross & Poirier, 2004; Soto, 2003).

Different proteins misfold and aggregate in different diseases: some prevalent examples are amyloid- β peptide ($A\beta$) in Alzheimer's disease (AD), tau in AD and frontotemporal dementia (FD), α -synuclein (α SN) in Parkinson's disease (PD) and dementia with Lewy bodies (LBD) and TAR DNA-binding protein 43 (TDP-43) in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) (Chiti & Dobson, 2017). Surprisingly, although these proteins are not particularly similar in terms of sequence, size, function or subcellular localization, the aggregation process and the aggregates produced share many common features across most NDs. In most ND diseases, the aggregation end-product are highly structured fibrillar species termed amyloid fibrils (**Figure 1.1**).

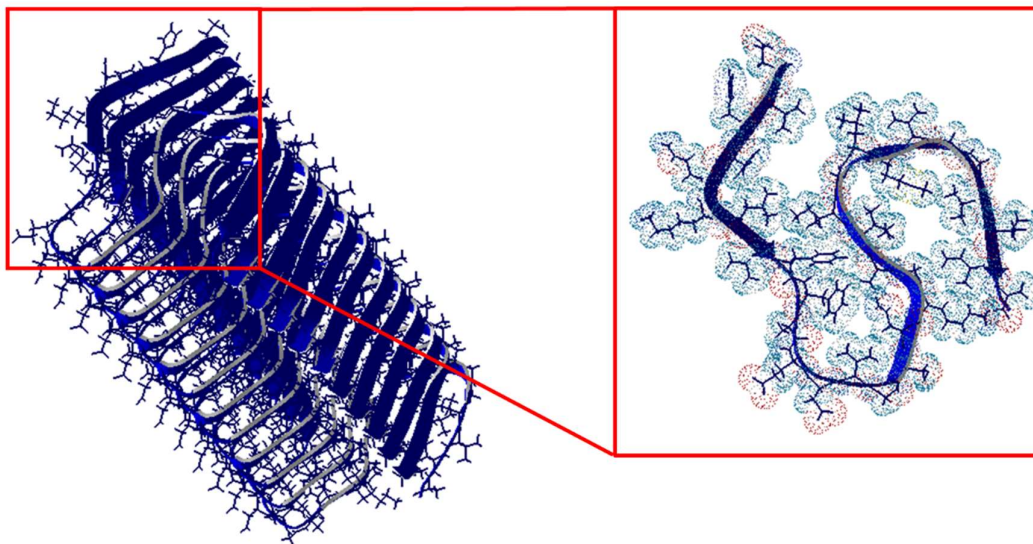


Figure 1.1: Structure of amyloid fibril from A β (1 - 42) peptide. Amyloid fibril structure of A β (1 – 42) peptide (PDB ID: 2mxu), determined *via* solid state NMR (ssNMR). Amyloid aggregates of A β are one of the hallmarks of Alzheimer's disease (AD), and the fibrils of the A β (1 – 42) peptide have been shown to be especially cytotoxic. Regardless of differences in sequence, most aggregating proteins form similar β -sheet rich fibrillar structures reactive to Congo Red and Thioflavin T (ThT) (Xiao *et al*, 2015).

Amyloid fibrils can be of varying length but are typically 100-200 Å in diameter; structurally they are arranged in a characteristic cross- β fold, which is based on main chain interactions and reactive to the Congo Red and Thioflavin T (ThT) dyes. While the general structure is similar across aggregating proteins, specific details of shape, function and cytotoxicity are dependent on the protein involved (Chiti & Dobson, 2017; Selkoe, 2003).

Protein folding is a stochastic and complex process, which can be prone to errors. The probability that a protein will assume a non-native fold depends on a variety of different factors, including genetic mutations, physical or chemical stress, temperature, and exogenous compounds not commonly found in the protein's intended sub-cellular environment (Dobson, 2004). These unfolded structures are in some cases metastable intermediates (that is: local minima of the protein folding potential energy surface); specialized pathways exist to deal with these, readily recognizing these aggregates and directing them for degradation by the proteasomal machinery of the cell (Chen *et al*, 2008; Hollien, 2013; Rousseau & Bertolotti, 2016). Often, proteins involved in NDs contain intrinsically disordered regions (IDRs) that do not have a well-defined structure in solution. The absence of a well-defined native state may serve a variety of different functions,

ranging from being a means of non-specifically driving liquid-liquid phase separation (LLPS) by associating with other IDR-containing proteins with similar physical characteristics (Alberti & Dormann, 2019; Lin *et al*, 2017), to acting as low-affinity binding partners for different cellular targets (Wright & Dyson, 2015). The lack of a natively well-defined structure may cause increased vulnerability to aggregation-inducing mutations or outside insults. Regardless, both natively well-folded and natively intrinsically disordered proteins (IDPs) given the right conditions can form aggregates.

Physically, the aggregation observed in NDs is well-described by the seeding-nucleation model (**Figure 1.2**), in which the rate determining step is the (slow and thermodynamically unfavourable) nucleation phase, which is then followed by a rapid elongation phase (Jarrett & Lansbury, 1993; Meisl *et al*, 2016). Once the initial stable seeds are formed, they drive polymerization by rapidly adding on monomeric protein units to the ends of the nascent fibril (Soto & Pritzkow, 2018).

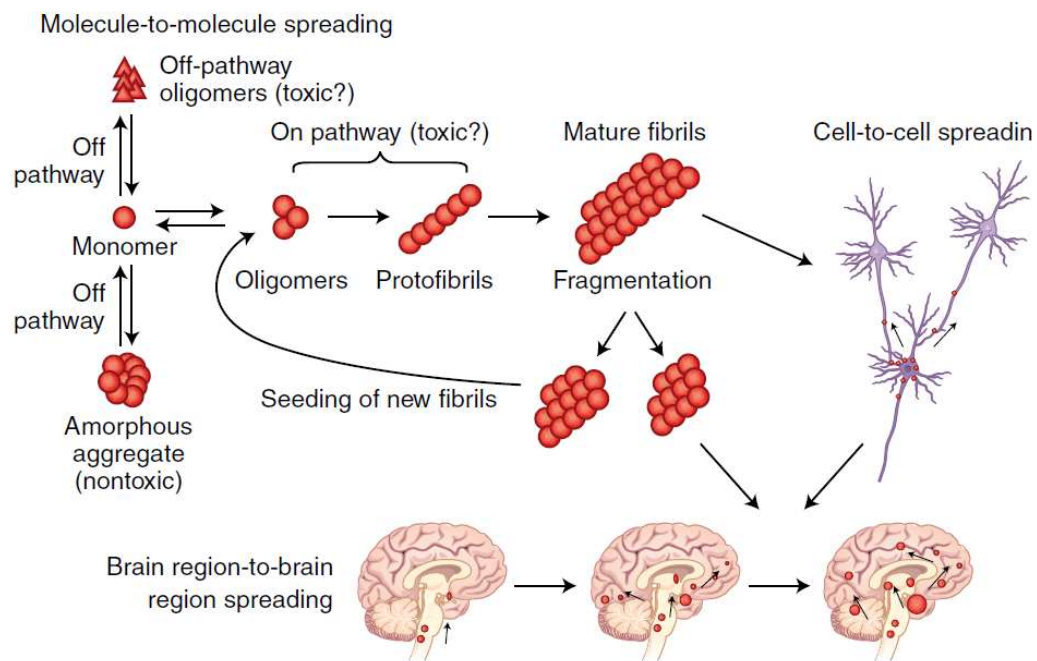


Figure 1.2: The seeding-nucleation model of protein aggregation and aggregate spread.

In the presence of certain triggers or specific mutations, the formation of aggregates becomes kinetically and/or thermodynamically favourable compared to the native fold of the protein. These can drive disease in a variety of ways (starting from loss-of-function: as in loss of the natively folded protein and its functions), but mainly cytotoxicity is thought to be induced through a toxic-gain of function mechanism, possibly driven by both the insoluble fibrillar aggregates and earlier soluble oligomers. There is also substantial evidence that misfolded/soluble oligomer species

have the ability to spread from one cell to its neighbours, and eventually across brain regions, in a manner similar to infectious prions. (*Adapted from: Soto & Pritzkow, 2018.*)

This model states that all misfolded protein aggregates have the ability to seed further misfolding and aggregation. This mechanism could imply the possibility of cell-to-cell spreading of the disease mediated directly by these misfolded seeds. It has been shown that this is the case with prion diseases (PrDs), in which the ND is initiated by exposure to an external misfolded protein agent, such as prion protein (PrP^{sc}), that acts as a seed, initiating aggregation of the host's native PrP (Prusiner, 1998). While exogenous infection has been demonstrated conclusively only with PrDs, it has been shown that amyloid aggregation (and the related disease phenotype) can be initiated in mice by injecting either homogenized tissue from affected individuals or even *in vitro* produced amyloid aggregates (Eisele *et al*, 2010; Ganowiak *et al*, 1994; Luk *et al*, 2012a; Luk *et al*, 2012b; Meyer-Luehmann *et al*, 2006; Volpicelli-Daley *et al*, 2011)

1.2 Liquid-liquid phase separation

A process that is becoming increasingly relevant in the field of protein aggregation, and indeed as a general principle of intracellular organization, is phase separation.

A phase of matter is physically defined as a region of space in which all physical properties of the material are uniform (solid, liquid and gaseous are examples of phases). A phase transition occurs when conditions (like pressure or temperature) are changed in the selected region of space causing a change of state (for example ice melting upon heating, switching from solid to liquid phase). However, phase separation occurs upon mixing two (or more) components that are in the same state of matter, and these autonomously segregate into distinct, homogeneous regions. An example of this is oil droplets forming in an aqueous solution.

Compartmentalization is a key feature of all cellular life. Separation of different regions of the cytoplasm through membranes allows for the formation of distinct chemical environments where different processes and reactions can occur. These membrane-defined enclosures are called organelles. Yet the presence of organelles is not sufficient to explain many cellular processes and functions, such as the localized translation of certain mRNAs, ribosome localization for export from the nucleus or some cytoskeletal dynamics (Banani *et al*, 2017). Significant evidence exists that many macromolecules can form dynamic liquid compartments that lack membranes but essentially act as organelles (Hyman *et al*, 2014; Li *et al*, 2012a; Milovanovic & De Camilli, 2017; Shin & Brangwynne, 2017). These liquid compartments form through a network of weak interactions that pay

the entropic cost of reducing the overall disorder of the system through a collective gain in enthalpy (**Figure 1.3**). Biochemical reactions must still be capable of occurring inside these phase separated clusters, so it is important that these interactions are weak and the molecules in the liquid phase are still capable of freely diffusing (Hyman *et al.*, 2014).

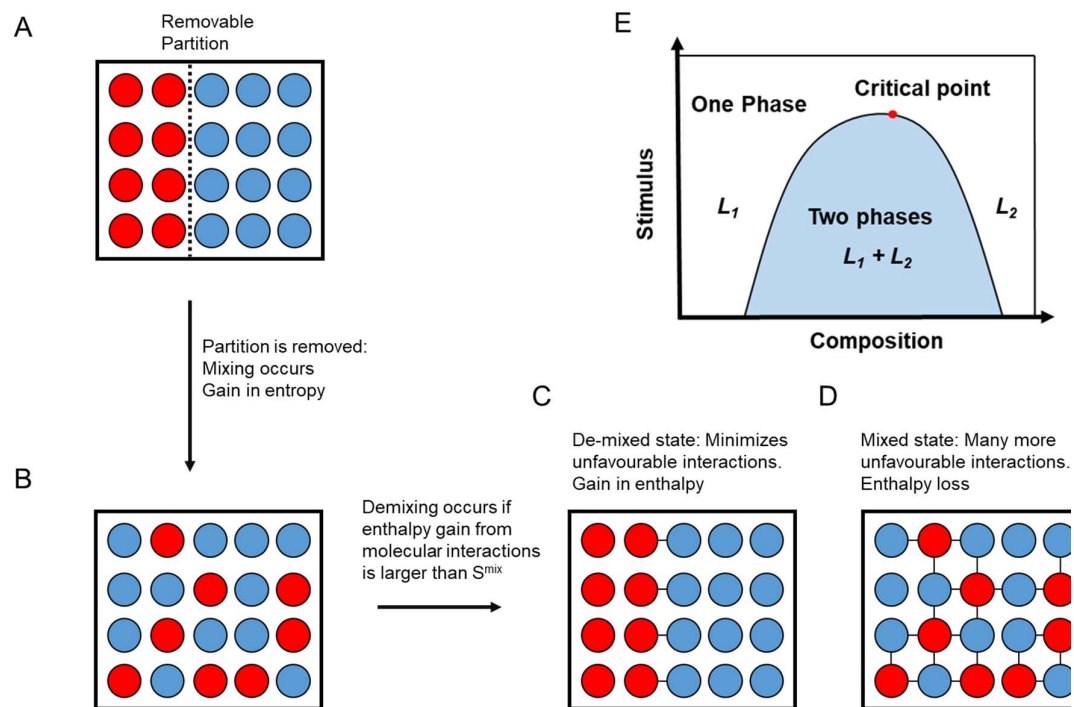


Figure 1.3: Thermodynamic principles of phase separation. (A – D) Mixing of two previously segregated liquids causes an increase in entropy of the solution. Separation in distinct phases occurs when the enthalpy gain associated with (i) minimizing unfavourable interactions between species and (ii) maximising favourable interactions between similar species is larger than the entropy of mixing (S^{mix}). This occurs for example if a very hydrophobic liquid is mixed with an aqueous liquid. (E) Liquid-liquid phase separation diagram. Stimuli include, amongst others: temperature, pH, salt concentration, crowding agents. Depending on these factors and the relative concentration of the two liquids (L_1 and L_2) mixture can either present itself as a single, well-mixed solution or as two distinct phases

In biological systems, the necessary enthalpy for condensates to form is provided by polar and ionic interactions between biomolecules. These mostly can be protein-protein interactions or protein-RNA interactions (**Figure 1.4**). Specificity of the components in these condensates is often given by highly specific interactions between well-folded protein domains and conserved motifs found in their binding partners (Li *et al.*, 2012a). Another common characteristic is the presence of intrinsically disordered regions (IDRs), amino acid stretches without a well-defined secondary/tertiary structure at physiological

conditions (Malinovska *et al*, 2013; Martin & Holehouse, 2020). Often short linear motifs (SLiMs) involved in specific protein-protein interactions can also be found embedded in these IDRs (Boeynaems *et al*, 2021; Davey *et al*, 2012). Protein-RNA interactions are also a key driving factor of the formation of many types of biological condensates. Unlike protein-protein interactions driving condensation, protein-RNA interactions that lead to phase separation are characterized by highly structured protein domains such as: RNA recognition domains (RRMs), Zinc-finger domains, Smith antigen-like domains (LSm) or hRNP K homology domains (KH domain) recognizing specific RNA bases or base sequences. In addition to these specific interactions, general non-specific association between basic amino acid (aa) stretches (usually found in IDRs) and the negatively charged RNA sugar-phosphate backbone also occur (Elbaum-Garfinkle *et al*, 2015; Lin *et al*, 2015; Van Treeck & Parker, 2018).

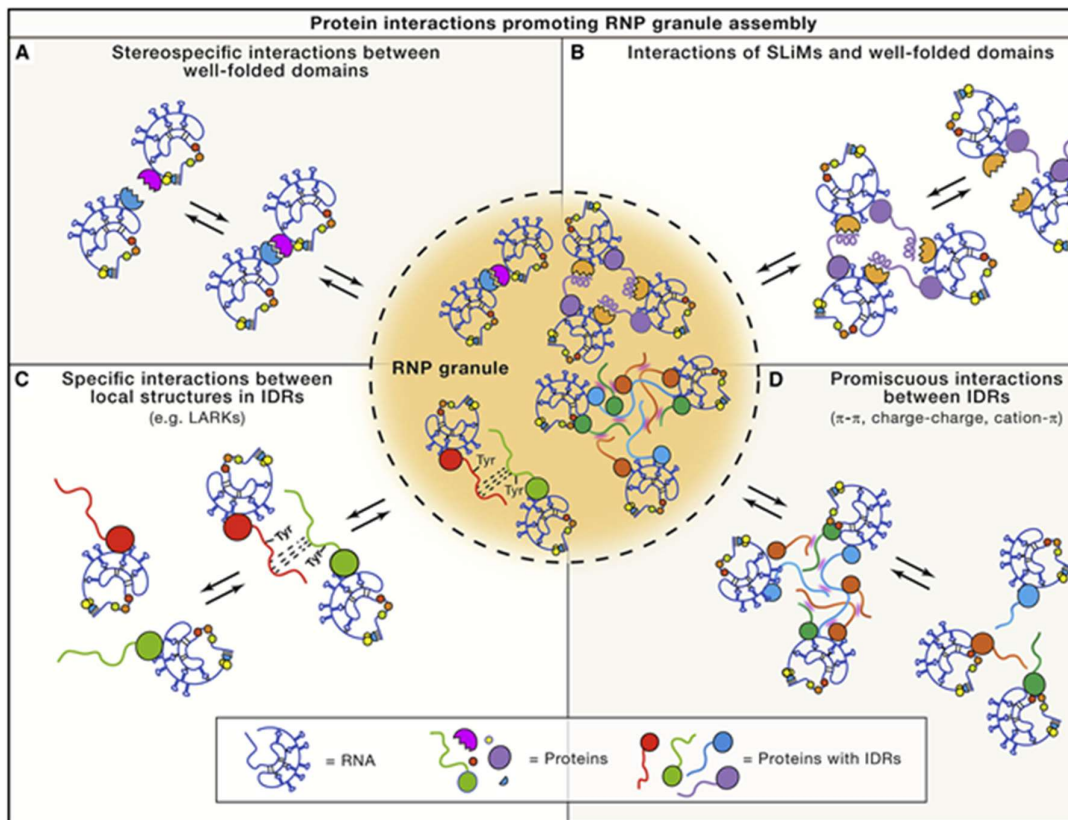


Figure 1.4: Summary of the diverse protein-protein interactions driving condensation of RNP granules. A variety of different interactions participate in the formation of protein-RNA condensates in cells, determining the composition, function and physical properties of the granules themselves. (A) Canonical interactions between well folded protein domains, (B) Interactions of well-folded domains with their peptide ligands, short sequences often embedded in intrinsically disordered stretches of the protein (C) Specific IDR-IDR interactions, such as basic

residue stretches binding negatively-charged stretches and (D) non-specific IDR-IDR interactions. (*Image adapted from: Van Treeck & Parker, 2018*).

The lack of a well-defined structure gives IDPs unique roles and characteristics within the proteome. From a thermodynamic point of view, they often have a tendency for non-specific self-association. This is because the lack of a definite fold means a much larger number of accessible conformations with very little energy required to transition from one state to another, so nearby IDR containing proteins are subjected to a much larger overall enthalpy gain from like-to-like surfaces interacting non-specifically with each other (much like oil droplets in water). This isn't commonly the case for well-folded proteins, as solvent-accessible surface area of these is much more restricted and is hydrogen-bonded to the solvent, meaning a much higher activation energy is typically required to cause two well-folded proteins to lose their native structure and interact non-specifically with each-other (Chiti & Dobson, 2017). Despite this, due to their conformational flexibility and their ability to form condensates that mediate precise biochemical processes, it is gradually becoming clear that a large chunk of the proteome is made up by IDR-containing proteins (Wright & Dyson, 2015).

In the context of neurodegenerative proteinopathies, it is becoming increasingly clear that in the majority of cases IDR-containing proteins are involved at some point of disease progression. This is hardly surprising, as their constitutional binding promiscuity and tendency to self-associate makes them particularly susceptible to mutations in their sequence or that of its interactors as well as cellular stressors such as heat or oxidative damage. Depending on the disease in question, dysfunction may occur in a variety of manners. Most commonly, IDR related diseases are associated with the formation of non-native protein aggregates, which may drive the disease either by loss of the biological functions associated to the proteins involved or through a toxic-gain of function of either the early oligomeric aggregates or the later-stage highly phosphorylated insoluble aggregates (typically a hallmark of the diseases where they are observed).

Membrane-less organelles built through phase separation are involved in a surprising amount of biological processes, such as: stress granules (SGs), processing bodies (PBs), the nucleolus, germ granules, neuronal transport granules and synaptic vesicles (Boisvert *et al*, 2007; Kedersha & Anderson, 2007; Kiebler & Bassell, 2006; Park *et al*, 2021; Voronina *et al*, 2011).

1.3 Stress granules and processing bodies

From the moment they are synthesized to when they are degraded, mRNA transcripts always exist as messenger ribonucleoproteins (mRNPs), being associated with a host of RNA-binding proteins (RBPs), whose identities vary across individual mRNAs (Anderson & Kedersha, 2009). Individual mRNPs are also remodelled, as associated RBPs change with the translational or metabolic state of each mRNA. Thus, different RBPs are associated with and control different activities of individual mRNAs including translation, localization, modification and degradation, which in turn influence expression levels of proteins, potentially essential to cellular survival, growth, differentiation or death. There are three key aspects to such regulation of mRNAs:

- (i) mRNA localization – To where in the cell will the mRNA transcript be transported before it is translated?
- (ii) mRNA translational control – When and where will the protein coded by the mRNA transcript be expressed, stored or degraded
- (iii) mRNA stabilization/ degradation – For how long is a transcript available before it is degraded?

mRNPs often assemble into structures known as mRNP granules. Many different kinds of mRNP granules have been discovered, each with specific functions and protein/RNA composition; two of them are stress granules (SGs) and processing bodies (PBs). These are membrane-less organelles that condense in the cytoplasm of most cell types, composed by a variety of proteins and translationally inactive mRNAs (Kedersha *et al*, 2005). Many protein factors are shared in both SGs and PBs, but some are unique to one or the other, and their structures are functionally and morphologically distinct (Ivanov *et al*, 2019) (**Figure 1.5**). Both play a role in translational regulation and post-transcriptional control and are in a highly dynamic relationship with actively translating polyribosomes as well as with each-other, as highlighted by the observation that a proportion of PBs appears to be docked to SGs (Cougot *et al*, 2004; Kedersha *et al*, 2000; Kedersha *et al.*, 2005; Moon *et al*, 2019). As with most mRNP granules, SGs and PBs most likely condense through a mixture of protein-protein interactions (Mittag & Parker, 2018), protein-RNA interactions (Vandelli *et al*, 2022) and RNA-RNA interactions (Van Treeck & Parker, 2018). These can involve different proteins that remodel mRNA structures, changing how they interact with other RBPs. Of the proteins that coalesce in SGs and PBs, some are strong drivers of LLPS (known as “scaffolds”) which play a physical role in the condensation process, an extreme example of this being Ras-GTPase-activating protein (SH3 domain)-binding protein 1 (G3BP1) for SGs (Yang *et al*, 2020), which when knocked out prevents

SG formation in response to stress. Others instead, while co-localizing in SG or PB *foci*, do not play a physical role in granule formation, and are known as “clients” (Banani *et al.*, 2017). Both SGs and PBs play a role in human disease. SGs in particular have been shown to be important in NDs such as ALS and FTD (Li *et al.*, 2013), as well as playing a role in the viral immune response (McCormick & Khapersky, 2017; Reineke *et al.*, 2015) and chemotherapeutic resistance in cancer cells (El-Naggar & Sorensen, 2018).

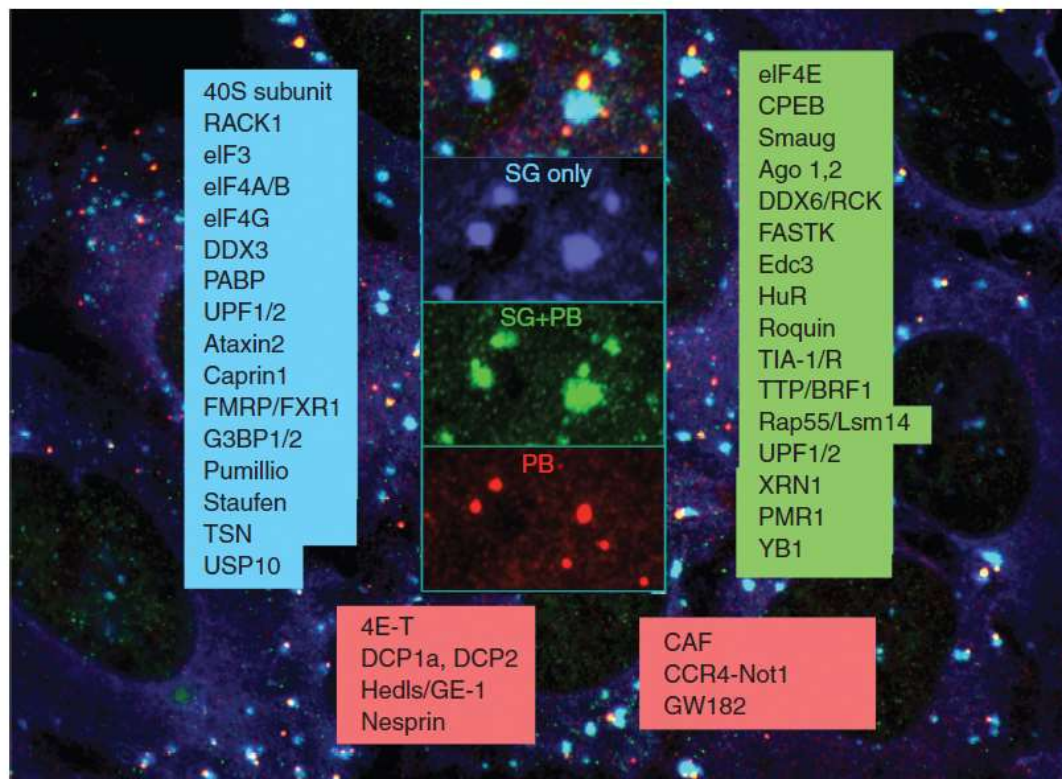


Figure 1.5: Partial list of proteins found in stress granules, processing bodies, or both. Image obtained from U2OS cells after arsenite stress, stained for eukaryotic initiation factor 3b (eIF3b, shown in blue), mRNA-decapping enzyme 1a (DCP1a, shown in red), and eukaryotic initiation factor 3b (eIF4E, shown in green), respectively markers for SGs, PBs or both. Shown in blue is a partial list of proteins found exclusively in SGs, in red a partial list of proteins found exclusively in PBs and in green a list of proteins that can be found in both. (*Image adapted from: Ivanov et al, 2019*)

While the role PBs play in disease is less clear, many protein components of PBs directly regulate and process mRNA transcripts involved in the inflammatory response (Anderson, 2010).

Stress Granules form and can be visualized in cells in the presence of stress. Initially they were discovered as heat-shock protein (HSP) containing *foci* observed after subjecting plant cells to heat shock (Nover *et al*, 1983), but it later became clear that different types of stress could induce similar granules in various cell types (Protter & Parker, 2016). SGs contain stalled pre-initiation complexes (PICs) together with small (40S) ribosomal subunits (Kimball *et al*, 2003). They also contain T-cell intracellular antigen (TIA-1), poly-A binding protein (PABP) and G3BP. Most of the mRNAs found in SGs are poly-adenylated (Anderson & Kedersha, 2009; Ivanov *et al.*, 2019). Over-expression of “scaffold” components of SGs can drive SG formation even in the absence of stress. Disassembly of SGs can be induced through cycloheximide treatment, while SG assembly can be promoted through puromycin treatment. Cycloheximide is an inhibitor of translation elongation, causing stalled PICs to complete the translation initiation process, trapping them in polysomes and causing a reduction of free, stalled mRNPs that can condense into SGs (Kedersha *et al.*, 2000). Puromycin instead causes early termination of translation, increasing the pool of mRNPs available for SG formation. Interestingly, although stalled PICs are necessary for SG formation, the opposite is not true and SGs are not required for translational arrest (Kedersha *et al*, 2016; Ohn *et al*, 2008; Singh *et al*, 2022). Fluorescence recovery after photo-bleaching (FRAP) studies have shown that many SG components continuously exchange between condensates and existing diffusely in the cytosol. Some, like G3BP and TIA-1 show quick rates of exchange (10 - 30 s) while others, like PABP take longer (30 – 60 s). Some exhibit even slower rates (minutes), like fragile X mental retardation protein (FMRp) (Bley *et al*, 2015; Kedersha *et al.*, 2005). The presence of protein diffusion in and out of SGs highlights both the highly dynamic and the tightly regulated nature of these condensates. A central regulator of SG formation, and indeed a key event in the integrated stress response (ISR), is phosphorylation of eukaryotic translation initiation factor 2 α (eIF2 α) (Kedersha *et al*, 1999). SGs formed in this manner contain most of the usual SG markers but lack eukaryotic translation initiation factors 2 and 5 (eIF2 and eIF5) – as non-phosphorylated eIF2 α is necessary for their assembly (Kedersha *et al*, 2002; Kimball *et al.*, 2003). Different eIF2 α kinases have been identified, each responding to a different type of cellular stressor:

- General control non-de-repressible 2 (GCN2), responsive to charged transfer RNA (tRNA) levels (Wek *et al*, 1995).
- Heme-regulated inhibitor (HRI), responsive to oxidative stress and heme availability (McEwen *et al*, 2005).
- Protein kinase R (PRK), responsive to double-stranded RNA (dsRNA) levels (Srivastava *et al*, 1998).

- PRK-like ER kinase (PERK), responsive to levels of unfolded protein (Harding *et al*, 2000).

Knockdown of each of these kinases will inhibit the cell's ISR (and SG formation) upon exposure to the specific stressor they respond to. Similarly, pharmacological intervention using integrated stress response inhibitor (ISRIB), an inhibitor of PERK, disassembles SGs but not PBs (Sidrauski *et al*, 2015), and has shown positive effects on neurodegeneration models (Wong *et al*, 2018).

Processing bodies are dynamic protein-RNA membrane-less organelles. Although they share many protein components with SGs, they also exhibit significant differences in terms of properties and function (Decker & Parker, 2012; Stoecklin & Kedersha, 2013). Unlike SGs, they are constitutively present in the cytoplasm of most cells and do not contain stalled PIC proteins, or in general most proteins involved in translation initiation. Instead, they are enriched in factors involved in mRNA decay and silencing, such as mRNA-decapping enzymes 1 and 2 (DCP1, DCP2), 5' – 3' exonuclease 1 (XRN1), enhancer of decapping protein 3 (EDC3), Argonaute (Ago) and GW182 proteins (Youn *et al*, 2019). This cohort of protein components suggests that PBs are cytoplasmic locations where active mRNA decay occurs (Cougot *et al.*, 2004). However, their function is actually more complex and nuanced than this, as mRNAs in PBs have been shown to exchange with the cytosol and can re-enter the pool of translating mRNAs (Brenques *et al*, 2005). Furthermore, micro-RNA (miRNA) silencing and other general mRNA decay pathways can occur even if PB condensation is disrupted, demonstrating that PB formation is not essential for these processes (Eulalio *et al*, 2007). These discoveries lead to a model in which mRNAs present in PBs are repressed but not immediately degraded, making them central hubs of translational control, from which mRNA transcripts can be rapidly mobilized either towards decay pathways or shuttled back towards translation (Standart & Weil, 2018). This is consistent with the fact that many of the mRNA transcripts found in PBs code for regulatory proteins, as well as proteins involved in the inflammatory response, which require rapid changes in expression levels upon specific cellular stimuli. Keeping a pool of these continuously stored would allow quicker response times by avoiding transcription, processing and nuclear export steps (Hubstenberger *et al*, 2017; Ivanov *et al.*, 2019).

A plethora of genetic and biochemical studies have implicated RBPs with NDs, most notably ALS, FTDs and AD (Bassell & Warren, 2008; Elden *et al*, 2010; Kim *et al*, 2013; Lefebvre *et al*, 1995; Neumann *et al*, 2006; Vance *et al*, 2009; Wegmann *et al*, 2018). From the study of mRNPs, and in particular SGs, plausible mechanistic explanations for

their known involvement in NDs are emerging. The finding that two proteins central to ALS pathology, TAR DNA-binding protein 43 (TDP-43) and Fused-in-Sarcoma (FUS) drive SG formation, together with the fact that other SG RBP-linked mutations influence disease progression gave the fundamental insight that SGs and other mRNPs could act as the initial site where protein aggregation occurs. In many ways, SGs are ideal hub for this, as the threshold between LLPS and a liquid to solid transition is small (Shin & Brangwynne, 2017). A possible sequence of events leading to protein aggregation seeded from SGs could be:

1. In normal conditions, the ISR is transient, resolved quickly after the stressor has ceased to act upon the cells exposed to it. In certain cell-types, or in response to specific environmental factors, this may occur more frequently, not allowing enough time for SGs to dissolve, causing the pathological transition of the granules.
2. SGs, by definition, contain aggregation-prone proteins at higher concentrations than in the surrounding cytoplasm. Given enough time, or in the presence of mutations favouring the aggregation process, these can transition from being fluid, dynamic organelles to solidified, β -sheet rich amyloid structures. Cells contain various safeguard mechanisms to contain and degrade unfolded and aggregating proteins, but over time these may become less effective.
3. With a timeline dependent on the disease and proteins involved, clearance pathways become overloaded by the continued presence of these solidified granules, failing to clear these efficiently. These aggregates become cytotoxic and in some cases may spread to the surrounding cells, seeding aggregation and causing progression of the ND.

1.4 Neuronal proteostasis and clearance pathways

Protein homeostasis (proteostasis) is defined as the regulation of the proteome's integrity. This includes all steps of the life cycle of a protein: synthesis, folding, correct quaternary structure achievement, turnover and degradation (Powers *et al*, 2009). Proteostasis is especially important in neuronal cells, as they generally do not undergo mitosis. Neuronal cells also typically have high levels of metabolic activity, and are thus more likely to undergo oxidative damage, which can cause the build-up of toxic misfolded and aggregated protein (Fu *et al*, 2018). A general method of waste clearance in the brain is believed to be a "glymphatic system," which is a paravascular system dedicated to the exchange of cerebrospinal fluid (CSF) and interstitial fluid (ISF). The Glymphatic system cycles these fluids through neuronal cells, eliminating toxic waste products (Iliff *et al*, 2013;

Jessen *et al*, 2015). Specialized, cell-autonomous methods of dealing with misfolded protein in neurons include chaperones. Their function is to help proteins achieve their correct fold and are connected to interlaced pathways that degrade terminally misfolded or aggregated proteins, namely the ubiquitin-proteasomal system (UPS), the lysosome-autophagy system (LAS) and the unfolded protein response (UPR).

Chaperones comprise various families of large, multidomain proteins. Their functions include acting like folding aids for proteins as they are being translated, refold proteins that have misfolded due to stress and as a barrier to protein aggregation. Their expression is upregulated when cells are subjected to stress (Saibil, 2013). Unlike enzymes, that have evolved highly specific chemical mechanisms at their active site to catalyse the reactions they are involved in, chaperones function in a more generic manner, acting as heavy-duty machines on a wide range of protein targets. Due to their upregulation under conditions of stress, many chaperone families are denoted as heat-shock-proteins (HSPs). The mechanism through which they induce correct folding differ across families, but is often an ATP-dependent process during which the target is inserted in a chemical environment that drives achievement of the native structure (Ali *et al*, 2006; Dollins *et al*, 2007; Shiau *et al*, 2006) (**Figure 1.6**).

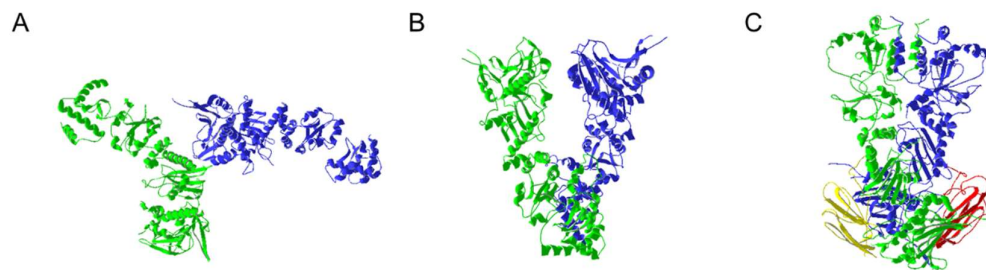


Figure 1.6: Change in conformation of the chaperone HSP90 upon substrate/co-factor binding. The HSP90 chaperone forms a dimer that changes in conformation to drive folding of its protein targets. (A) Fully open conformation of HSP90, no substrate or co-factors (PDB ID: 2IOQ). (B) HSP90 in its open but ADP bound form (PDB ID: 2O1V). (C) Closed, ATP and co-factor bound form of HSP90 (PDB ID: 2CG9). The HSP90 target protein binds to the central cleft, which closes in an ATP- dependent manner, forcing the protein in a chemical environment driving the correct native fold.

Chaperones are also connected with the cell's protein degradation pathways, and in some cases can direct misfolded or aggregated protein to the LAS and UPS, and even initiate apoptosis (Takayama *et al*, 2003).

The ubiquitin proteasomal system is a highly dynamic network of processes that mediate the degradation of misfolded and aggregated protein. It is an ATP dependent process during which ubiquitin (Ub), a small, highly conserved protein, is ligated to key lysine residues on the protein targeted for degradation, which is then shuttled to the proteasomal machinery (Hershko & Ciechanover, 1998). Usually, more than a single Ub monomer is appended, and typically chains (Ub has many lysine residues especially for this function) of Ub form. The way a protein is polyubiquitinated will ultimately determine its fate (Grabbe *et al*, 2011). A variety of enzymes, working in a concerted manner, mediate and give specificity to this process, and are commonly grouped as ubiquitin-activating (E1), ubiquitin-conjugating (E2) and ubiquitin-ligases (E3). Over 35 E2 and more than 600 E3 enzymes have been identified to date, the E3 determining the final target to which Ub will be attached (Hershko & Ciechanover, 1998; Rao *et al*, 2015).

The proteasome itself consists in a very large, multidomain cylindrical structure, composed by the 20S catalytic core particle (CP) and the 19S regulatory particles (RPs). The CP forms the barrel where the proteolytic reactions take place (Toste Rêgo & da Fonseca, 2019) (**Figure 1.7**). While the two RPs lock the polyubiquitinated protein inside the complex. Two central hetero-heptameric β -subunit rings sandwiched in-between two α -subunit hetero-heptameric rings make up the CP. The central β -subunit stack contains a diverse set of protease sites, which cleave peptide bonds after any acidic residue ($\beta 1$), any basic residue ($\beta 2$) or any hydrophobic residue ($\beta 5$). The 26S capped conformation is the resting state of the proteasome (Jung & Grune, 2012).

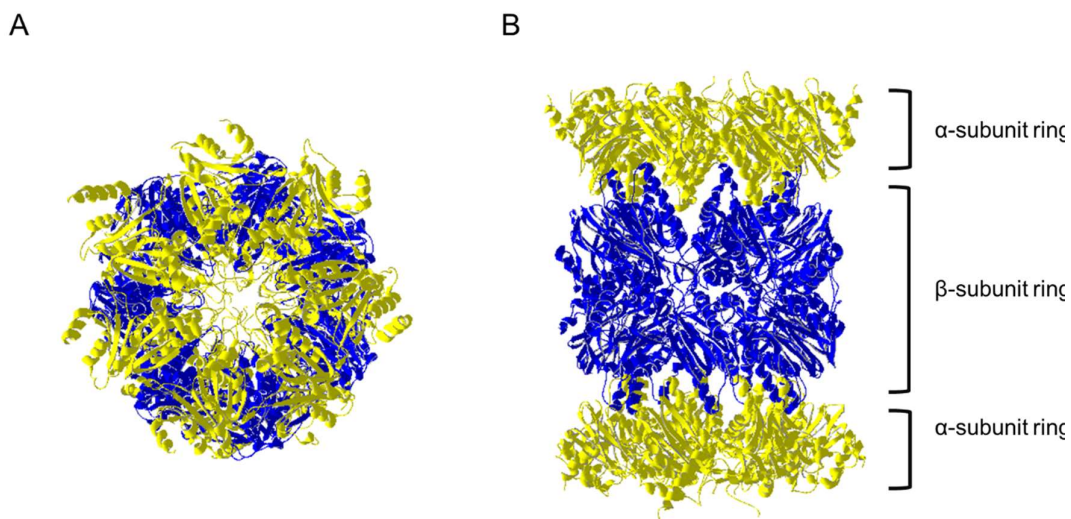


Figure 1.7: Structure of the 20S (uncapped) proteasome core catalytic particle. Crystal structure of the proteasomal 20S CP (PDB ID: 6RGQ). The polyubiquitinated protein sent for degradation enters the central pore. The full functional proteasome is then capped by two 19S

RPs, forming the full 26S proteasome (not shown). The 20S CP is made up by two hetero-heptameric β -subunit rings (blue) set between two hetero-heptameric α -subunit rings (yellow). (A) Top view of the 20S CP. (B) Side view (Toste Rêgo & da Fonseca, 2019).

The lysosome-autophagy system is a complicated network through which the cell sequesters and digests unwanted macromolecules and organelles back to their component parts, which can then be recycled and used to build new proteins, carbohydrates, lipids or nucleic acids (autophagy literally means “self-eating”) (Ichimiya *et al*, 2020). This occurs by shuttling waste products, *i.e.* misfolded proteins and damaged organelles, to lysosomes, specialized, membrane-bound, single-lipid bilayer organelles with an acidic (pH $\sim$ 4.5) internal milieu that are filled with proteolytic enzymes (Bouhamdani *et al*, 2021). In most cell-types this is a characteristic response to starvation (He *et al*, 2018), but in long-living cells such as neurons where apoptosis isn’t a sustainable response to cellular damage or stressors, the LAS takes on a more central role in maintaining cellular homeostasis (Finkbeiner, 2020). Autophagy is traditionally subdivided in three broad categories (although many more subtle pathway distinctions can be made):

- (i) Macroautophagy mainly takes care of damaged organelles such as mitochondria and peroxisomes. This occurs through the activation of a variety of autophagy related genes (ATGs) coordinated by the central regulator protein mammalian target of rapamycin (mTOR) in response to nutrient availability (Kim *et al*, 2002)
- (ii) Chaperone-mediated autophagy (CMA) is where proteins or peptides containing a specific recognition sequence bind to certain chaperones (like HSP90) which shuttle them to the lysosome through receptor mediated endocytosis (Dice, 1990; Kaushik & Cuervo, 2012).
- (iii) Microautophagy is a poorly understood mechanism through which targeted proteins enter the lysosome through invaginations of its lipid bilayer (Li *et al*, 2012b).

The unfolded protein response is triggered when protein folding at the endoplasmic reticulum (ER) is inhibited or somehow perturbed (Hetz *et al*, 2020). The ER is the main centre of secreted protein translation in cells, with $\sim$ 30% of the whole proteome being translated in this location (Schwarz & Blower, 2016). Essentially, the UPR is a network of ER signal transduction events that when triggered, upregulate in both a local and global manner some of the abovementioned protein-misfolding response pathways. Three main sensors involved in triggering the UPR have been discovered so far, inositol-requiring protein 1 α (IRE1 α), PERK and activating transcription factor 6 (ATF6) (Preissler & Ron, 2019). The activation of these sensors causes two temporally resolved waves of events

to occur: (i) Protein synthesis is stalled *via* PERK phosphorylation of eIF2 α , also causing SG condensation, and certain mRNA transcripts coding for ER proteins are degraded (Harding *et al.*, 2000). Macroautophagy is upregulated, probably to clear damaged sections of the ER and clumps of aggregating protein. These mechanisms prevent translocation to and expression of proteins in the ER, giving it time to achieve homeostasis again (Hetz, 2012). (ii) On a longer timescale, transcription of a variety of protective genes is heavily upregulated. These include a variety of chaperones, to aid protein folding at the ER as well as a variety of genes involved in autophagy, to increase the turnover speed of proteins and macromolecules, limiting their amounts in the ER lumen and surrounding cytoplasm (Hetz *et al.*, 2020; Hollien, 2013). This is an adaptive response to help protect from further insults. If regardless of these safeguards, toxic levels of unfolded protein persist at the ER, sustained UPR signalling will trigger apoptosis (Benbrook & Long, 2012).

Cellular proteostatic pathways are intricate, highly inter-connected and controlled by highly elaborate signalling pathways (Benbrook & Long, 2012; Pakos-Zebrucka *et al.*, 2016). As proteins are the “machines” responsible for the vast majority of essential biochemical processes that occur in cells, the presence of layers over layers of “safeguard” mechanisms to protect from potential errors and mistakes is absolutely necessary for the survival and growth of cells and organisms in general. This is especially true for complicated, long-lived cells with high-energy requirements, such as many types of neurons in animals with a complex nervous system. The devastating consequences of NDs are an excellent example of the fate of organisms in which a protein malfunction slips through the cracks of the proteostatic system.

1.5 Rationale (Statement of purpose)

The prevalence of neurodegenerative diseases is continuously rising in the global population, proportionally with the increase in life expectancy. These diseases impact the quality of life of both individuals affected and those surrounding them. Indeed, neurodegenerative diseases lead to premature death in most patients. Most neurodegenerative diseases are currently incurable, with the vast majority of the available treatments almost exclusively aimed at ameliorating symptoms/slowing down disease progression. Despite the enormous global burden of these diseases, many aspects of their molecular mechanisms of pathogenicity remain poorly understood, although it is generally accepted that for many of these diseases the aggregation of specific proteins is a driver of disease progression (Chiti & Dobson, 2017). In light of this, a better understating the mechanisms of protein aggregation is necessary for the elucidation of potential drug

targets and designing novel therapeutics. In the past two decades or so, there has been an increase in interest towards soluble oligomeric aggregates, rather than the insoluble, hyper-phosphorylated fibrils found in inclusion bodies as drivers of disease (Fusco *et al*, 2017; Thibaut *et al*, 2018). There are multiple reasons for this: (1) these may play a prion-like role in disease spreading, (2) there is direct evidence of their toxicity and (3) insoluble aggregates tend to be found in surviving cells, meaning there is a chance that insoluble aggregates may be due to cellular rescue pathways. Many proteins involved in neurodegenerative diseases contain intrinsically disordered stretches, which in turn have been shown to play an important role in LLPS (Mateju *et al*, 2017). Both the proteins investigated in this study (α SN and Atxn2) have been shown to participate in LLPS both *in vivo* and *in vitro*. Understanding the roles of their different domains in these processes is important as proteins in a liquid-liquid phase separated state are more likely to undergo a liquid to solid transition, meaning this is perhaps the initial physical state in which amyloid aggregates are formed.

1.6 Hypotheses

Chapter 3:

- Soluble oligomers of α SN are cytotoxic/pathogenic species involved in PD.
- These oligomers will exhibit physiological effects on long-term potentiation (LTP) *in vivo* in rats.
- SE-HPLC can be established as a powerful and accessible method for monitoring early aggregation steps of pathogenic proteins, provided that the appropriate aggregation protocol is selected.
- Diffusion-ordered spectroscopy NMR can be a useful corroborating technique to determine sizes of heterogeneous protein aggregates.

Chapter 4:

- The PAM2 domain of the protein Atxn2 is central in selecting its mRNA and protein targets.
- The PAM2 domain of the protein Atxn2 is an important regulator of the protein and mRNA composition of Atxn2 containing mRNP granules.
- The PAM2 domain of the protein Atxn2 is necessary and sufficient for Atxn2:PABPC1 binding.
- The Atxn2:PABPC1 interaction is necessary and sufficient for the correct localization of Atxn2 to stress granules.

- The LSm domain of the protein Atxn2 is necessary and sufficient for Atxn2:LSM12 binding.
- The binding of the LSm domain of Atxn2 binds the protein LSM12 through a canonical LSm:LSm interaction. This LSm:LSm interaction is specific for the Atxn2:LSM12 interaction.
- The Atx2/Atxn2/Atxn2L IDRs are not functionally equivalent, having instead different LLPS inducing competencies.
- The cIDR and PAM2 domains of Atxn2 promote neurodegeneration, while the LSm domain inhibits it.

1.7 Aims

Chapter 3:

- Investigate the progression of early α SN aggregation over time when using the chosen protocol (Danzon *et al*, 2007): Pinpoint at what time-point oligomeric species become observable.
- Attempt to isolate the species observed with SE-HPLC.
- Establish SE-HPLC as a useful, reproducible and accessible method for probing the progression of the early steps of protein aggregation. Cross-verify the sizes of species observed using other techniques.
- Assess the pathophysiological relevance of the oligomers produced in cells and *in vivo* in rats.

Chapter 4:

- Investigate the importance of the Atxn2 PAM2 domain in selecting its mRNA targets, as well as the protein components of mRNP granules.
- Using radically truncated constructs, understand if the interactions of Atxn2 with PABP and LSM12 depend solely on the Atxn2 PAM2 and LSm domains respectively.
- Characterize the binding of these proteins structurally, and attempt to devise ways to specifically inhibit these structured interactions (*i.e.* through point-mutations of minimized Atxn2 constructs).
- Compare LLPS competency information contained in the sequence of the IDRs of Atxn2, Atxn2L and Atx2.

- Investigate if the Atxn2 mRNP granule composition (protein and mRNA) is a central determinant of Atxn2-related cytotoxicity.

2. MATERIALS & METHODS

2.1 Key resources & Materials

Table 2.1: Key resources & materials

<i>Material</i>	<i>Source</i>
Acetonitrile (CH ₃ CN) HPLC grade	Fisher Scientific
Albumin from human serum ≥ 99%	Sigma-Aldrich
Alpha lactalbumin from bovine milk	Sigma-Aldrich
Bis(sulfosuccinimidyl)suberate (BS ³)	Thermo-Fisher Scientific
β-Lactoglobulin from bovine milk ~ 90%	Sigma-Aldrich
Chelex® 100 Resin	Bio-Rad
Cytochrome C from equine heart ≥ 90%	Fluka analytical
Cu-Zn Superoxide dismutase from bovine erythrocytes	Sigma-Aldrich
Digitonin	Sigma-Aldrich
Dithiothreitol (DTT) molecular biology grade	Thermo-Fisher Scientific
Dulbecco modified eagle media (DMEM)	Gibco
Ethanol (C ₂ H ₅ OH) HPLC grade	Sigma-Aldrich
γ-Globulins from bovine blood ≥ 99%	Sigma-Aldrich
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
Guanidine hydrochloride solution ~ 8.0 M (NH ₂ C(=NH)NH ₂ • HCl)	Fluka Biochemika
HEPES	Sigma-Aldrich
Hexylene glycol	Sigma-Aldrich
Hydrochloric acid (HCl)	Reidel-de Haën
Human recombinant α-Synuclein	Osaka University*

Human recombinant insulin	Sigma-Aldrich
Igepal	Sigma-Aldrich
Insulin Polyclonal Antibody (Rabbit IgG)	Invitrogen
IRDye® 800CW Goat anti-Rabbit IgG (H + L)	Li-Cor
L-Tryptophan	Sigma-Aldrich
Lysozyme from chicken egg white ≥ 90%	Sigma-Aldrich
Methanol (CH ₃ OH) HPLC grade	Fisher Scientific
NuPAGE® Antioxidant	Thermo-Fisher Scientific
NuPAGE® LDS Sample Buffer (4X)	Thermo-Fisher Scientific
NuPAGE® SDS Running Buffer (20X)	Thermo-Fisher Scientific
NuPAGE® Transfer Buffer (20X)	Thermo-Fisher Scientific
Odyssey Blocking Buffer	Li-Cor
PageRuler Prestained Protein Ladder	Thermo Scientific
Potassium Phosphate (K ₂ HPO ₄)	Fluka Biochemika
Protease inhibitor cocktail	Roche
PVDF blotting membrane	Sigma-Aldrich
RPMI 1640	Sigma-Aldrich
Sodium chloride (NaCl)	Sigma-Aldrich
Sodium hydroxide (NaOH)	Sigma-Aldrich
Sodium phosphate monobasic dihydrate (NaH ₂ PO ₄ • 2H ₂ O)	Reidel-de Haën
Sodium phosphate dibasic heptahydrate (Na ₂ HPO ₄ • 7H ₂ O)	Sigma-Aldrich
Sodium sulfate (Na ₂ SO ₄) 99.99%	Sigma-Aldrich
Thioflavin T (C ₁₇ H ₁₉ ClN ₂ S) (ThT)	Aquos

Thyroglobulin from bovine thyroid ≥ 90%	Sigma-Aldrich
Trypsin	Thermo-Fisher Scientific
Tween® -20	Sigma-Aldrich
Ubiquitin from bovine erythrocytes ≥ 98%	Sigma-Aldrich
Urea (CO(NH ₂) ₂) 98+%	Sigma-Aldrich
UAS-Atx2-WT-ADARcd	Singh et al. 2021
UAS-Atx2-ΔLSm-ADARcd	Singh et al. 2021
UAS-Atx2-ΔPAM2-ADARcd	Petrauskas & Fortunati et al. (in press)
UAS-Atx2-ΔcIDR-ADARcd	Singh et al. 2021
UAS-Atx2-only LSm/Lsm-AD -ADARcd	Singh et al. 2021
Mef2-Gal4; Tub-Gal80 ^{ts}	Bloomington Drosophila Stock center
Elav-Gal4; Tub-Gal80 ^{ts}	Bloomington Drosophila Stock center
S2R+ cells	DGRC
HEK293T cells	Gift: Adrian Bracken lab
U2OS cells	Gift: Martina Schroeder lab
pUASst-Atx2-SNAP_fly (Plasmid)	Genscript
pUASst-Atxn2-SNAP_hum (Plasmid)	Genscript
pUASst-mini-Atx2-SNAP_fly (Plasmid)	Genscript
pUASst-mini-Atxn2-SNAP_hum (Plasmid)	Genscript
pUASst-ΔLSm-mini-Atx2-SNAP_fly (Plasmid)	Genscript
pUASst-ΔLSm-mini-Atxn2-SNAP_hum (Plasmid)	Genscript
pUASst-ΔLSmAD-mini-Atx2-SNAP_fly (Plasmid)	Genscript
pUASst-ΔLSmAD-mini-Atxn2-SNAP_hum (Plasmid)	Genscript

pUAS Δ PAM2-mini-Atx2-SNAP_fly (Plasmid)	Genscript
pUAS Δ PAM2-mini-Atxn2-SNAP_hum (Plasmid)	Genscript
pUAS Δ L859A-mini-Atx2-SNAP_fly (Plasmid)	Genscript
pUAS Δ L914A-mini-Atxn2-SNAP_hum (Plasmid)	Genscript
pUAS Δ F866A-mini-Atx2-SNAP_fly (Plasmid)	Genscript
pUAS Δ F921A-mini-Atxn2-SNAP_hum (Plasmid)	Genscript
pUAS Δ L859A-F866A-mini-Atx2-SNAP_fly (Plasmid)	Genscript
pUAS Δ L914A-F921A-Q928A-mini-Atxn2-SNAP_hum (Plasmid)	Genscript
pUAS Δ I328A-L329A-F330A-mini-Atxn2-SNAP_hum (Plasmid)	Genscript
pUAS Δ I328A-L329A-F330A-mini-Atxn2-SNAP_fly (Plasmid)	Genscript
pUAS Δ (328:339)-mini-Atxn2-SNAP_hum (Plasmid)	Genscript
pUAS Δ (328:339)-mini-Atxn2-SNAP_fly (Plasmid)	Genscript
pUAS Δ Atx2-Structured-Atxn2-IDRs-SNAP_fly (Plasmid)	Genscript
pUAS Δ Atx2-Structured-Atxn2L-IDRs-SNAP_fly (Plasmid)	Genscript
pcDNA3.1-Atxn2-Structured-Atx2-IDRs-SNAP_human (Plasmid)	Genscript
pAcman-Atx2-GFP (Fosmid)	(Sudhakaran <i>et al</i> , 2014)
pAcman- Δ L5m-Atx2-GFP (Fosmid)	(Sudhakaran <i>et al.</i> , 2014)
pAcman- Δ PAM2-Atx2-GFP (Fosmid)	(Sudhakaran <i>et al.</i> , 2014)
pActin-Gal4	DGRC

pCMV-Gal4	Addgene
Anti-Atx2 (chicken polyclonal)	(Bakthavachalu <i>et al</i> , 2018)
Anti-Caprin (rabbit polyclonal)	(Papoulas <i>et al</i> , 2010)
Anti-dFMR (mouse monoclonal)	DSHB
Anti-GFP (chicken polyclonal)	Abcam
Anti-V5 (rabbit polyclonal)	Santa Cruz Biotechnology
PABP (rabbit polyclonal serum)	(Lee <i>et al</i> , 2017)
Me31B (rabbit polyclonal serum)	(Lee <i>et al.</i> , 2017)
Rox8 (rat polyclonal)	(Buddika <i>et al</i> , 2020)
SNAP (rabbit polyclonal)	NEB
PABPC1 (rabbit polyclonal)	Abcam
Lsm12 (rabbit anti human polyclonal)	Abcam
G3BP (mouse monoclonal)	BD Bioscience
Histone H3 (mouse polyclonal)	Abcam
Baf155 (mouse polyclonal)	Abcam
HRP Goat anti-rabbit	Invitrogen
HRP Goat anti-mouse	Invitrogen
Alexa Fluor 555 (polyclonal goat anti-chicken IgG)	Invitrogen
Alexa Fluor 488 (polyclonal goat anti-chicken IgG)	Invitrogen
Alexa Fluor 647 (polyclonal goat anti-chicken IgG)	Invitrogen
Alexa Fluor 555 (polyclonal goat anti-rabbit IgG)	Invitrogen
Alexa Fluor 488 (polyclonal goat anti-rabbit IgG)	Invitrogen
Alexa Fluor 647 (polyclonal goat anti-rabbit IgG)	Invitrogen
Alexa Fluor 555 (polyclonal goat anti-mouse IgG)	Invitrogen
Alexa Fluor 488 (polyclonal goat anti-mouse IgG)	Invitrogen

Alexa Fluor 647 (polyclonal goat anti-mouse IgG)	Invitrogen
MOWIOL mounting medium	Sigma (Merck)
SNAP-TMRStar	New England Biolabs
SNAP-Surface 488	New England Biolabs
TRIBE	(McMahon <i>et al</i> , 2016)
STAR v2.5.3	(Dobin <i>et al</i> , 2013)
HTSeq v0.11.2	(Anders <i>et al</i> , 2015)
DESeq2	(Love <i>et al</i> , 2014)
AREScore	(Spasic <i>et al</i> , 2012)
Guitar	(Cui <i>et al</i> , 2016)
Bedtools	(Quinlan & Hall, 2010)

**The human recombinant α -Synuclein sample was expressed in E.coli, purified and provided by Prof Young Ho Lee, Osaka University, Japan.*

2.2 Instrumentation

Analogue tube roller srt6 (Stuart)

Biosafety cabinet Esco sentinel gold class II (ESCO)

BioScope Resolve atomic force microscope (Bruker)

DD2 Premium 18.8 T (800 MHz) Compact NMR spectrometer (Agilent Technologies)

DNP-10 (Bruker) silicon nitride cantilevers

FP-6200 spectrofluorometer (Jasco)

J-815 circular dichroism spectrometer (Jasco)

JEM 1400 transmission electron microscope (JEOL)

LSM 880 with Airyscan (ZEISS)

MFP-3D-BIO atomic force microscope (Asylum Research)

Microcentrifuges: Galaxy 14D (VWR) & Ministar (VWR)

MiVac Quattro concentrator (SP Scientific)

MST basic C magnetic stirrer with heating (Yellowline, ImLab)

NanoDrop ND-1000 UV-Vis spectrophotometer (Thermo Scientific)

Novex Mini-Cell gel box (Invitrogen)

NuPAGE 8%,10%,12% Bis-Tris protein gels, 1.0 mm, 10-well (Thermo-Fisher Scientific)

Olympus IX61 phase contrast microscope.

pH meter 240 (Corning)

Pipetmans: P20, P200, P1000 (Gilson)

PowerEase500 power supply (Novex)

Rheodyne MXP 10-port 2-position injection valve (Dionex, Thermo Scientific)

SpectraMax® M3 microplate reader, Molecular Devices

Superdex™ 200 10/300 GL, chromatography column, GE healthcare

TSKgel SuperSW 3000 Column, 4.6 mm x 30 cm (Tosoh Bioscience)

TSKgel SuperSW Guard Column, 4.6 mm x 3.5 cm (Tosoh Bioscience)

TTS 2 test tube shaker (Yellowline, ImLab)

UltiMate HPLC 3000 standard system (Dionex, Thermo Scientific)

Vortex-1 genie touch mixer (Scientific Industries)

XCell II blot module (Invitrogen)

2.3 Protein aggregation protocols

Recombinant α SN was provided by Professor Young-Ho Lee, Osaka University. Sample was expressed in *E. coli* and purified *via* Ni-NTA affinity chromatography followed by ion exchange and size exclusion chromatography and lyophilized.

2.3.1 Fibril formation (α SN)

Aliquots of 200 μ l each at 100 μ M recombinant α SN were incubated in 0.1 M Na_2HPO_4 buffer, 100 mM NaCl, pH 6.5. Samples were subjected to gentle agitation on a srt6 analogue tube roller (Stuart) at room temperature and checked daily for fibril formation using thioflavin T (ThT) fluorescence spectrophotometry. Controls were made in the same manner but not subjected to agitation and stored at 4 °C for the duration of the experiment.

2.3.2 Oligomer formation for cytotoxicity & animal experiments (α SN)

A solution of 7 μ M α SN in 50 mM, sterile filtered Na_2HPO_4 buffer, 20% EtOH, pH 7.0 was vigorously shaken for 4 hours on a TTS 2 test tube shaker (ImLab) fitted with a polystyrene centrifuge tube rack secured using masking tape (Danzer *et al.*, 2007). Tubes were left uncapped in order to evaporate the ethanol. Samples were then lyophilized using a MiVac Quattro concentrator (Sp Scientific) rotary evaporator and re-suspended to a α SN concentration of 14 μ M in 50 mM sterile filtered Na_2HPO_4 buffer, 10 % EtOH, pH 7.0. The sample was again subjected to agitation in the same manner, uncapped to evaporate the ethanol, for 24 hours. Samples were then incubated for 6 days, capped lids, no shaking at room temperature. Samples were once again lyophilized and re-suspended at a concentration of 100 μ M.

2.3.3 Oligomer formation for SE-HPLC (α SN)

For SE-HPLC, the oligomer formation protocol was slightly modified to suit experimental needs: Sample was incubated in 50 mM sterile filtered Na_2HPO_4 buffer, 10% EtOH pH 7.0, lids capped to avoid contamination and kept continuously shaking for the duration of each experiment. Lyophilisation steps were also avoided and α SN concentration was kept at 14 μ M.

2.3.4 Hen egg white lysozyme (HEWL) aggregation

Aggregation of HEWL was induced by incubating the protein at a concentration of 1 mM in dH₂O/HCl solution, pH 2.5, at 65 °C. Samples were divided in different aliquots, then quenched in ice and frozen at time points necessary for the various experiments. Samples were thawed and diluted in dH₂O/HCl solution (pH 2.5) to a concentration of 50 µM immediately before injection.

2.4 Size exclusion high performance liquid chromatography (SE-HPLC)

All chromatography experiments were carried out at room temperature using an UltiMate HPLC 3000 standard system (Dionex, Thermo Scientific), running Chromeleon 6 software. Size-exclusion columns used were a TSK gel superSW3000 (Tosoh bioscience; Pore size: 250 Å, Particle size: 4 µm) and Superdex 200 10/300 GL (GE Healthcare; Pore size: 1 – 100 kDa MW range, Particle size: 13 µm). Standard curves were run as a calibration/performance check after any change in buffer set up. Samples were run in 50 mM Na₂HPO₄, 0.1 M Na₂SO₄, pH 7 (TSK), 50 mM Na₂HPO₄, 0.15 M NaCl, pH 7 (Superdex) or 50 mM Na₂HPO₄, 0.15 M NaCl, pH 2.5 (Superdex). Other ionic strengths of the running buffers were tested (0.1 M and 0.3 M Na₂SO₄) to account for potential ion-exchange effects on the stationary phase. The resulting peaks were collected manually, and if necessary further re-chromatographed. There have been cases where the microenvironment of the HPLC column may induce structural perturbations due to the increase in temperature causing in-column shearing (Fekete *et al*, 2014). We are not aware that this is occurring to an extent of varying the distribution of oligomers in the present study, as the particle and pore sizes of the SE-HPLC columns used during these experiments are larger (= 4 µm, 250 Å).

2.5 Gel electrophoresis, western blotting & immunoprecipitation

2.5.1 Gel electrophoresis of α-synuclein

The αSN monomer and oligomer samples were analysed through SDS PAGE. Briefly: two NuPAGE 12% Bis-Tris Protein Gels, 1.0 mm, 10-well (Thermo-Fisher Scientific) were run, one with protein samples ranging 1 - 2.8 µg per well and the second one 3 – maximal µg was loaded. All except the maximal concentration αSN oligomer sample were boiled for 8 min at 82 °C. The αSN oligomer sample was directly dissolved in 20 µL of NuPAGE® LDS Sample Buffer (4X), boiled for 30 min at 65 °C and diluted through the addition of 10 µL of

dH₂O. Lane 1 was not used to avoid edge effects and was loaded with 4x loading buffer and dH₂O (20 µL).

2.5.2 Western blotting & Immunoprecipitation of minimized Atxn2 constructs

Total protein extracts were prepared from opportunistically transfected HEK293T cells were lysed and immunoprecipitation was carried out for each Atxn2-SNAP construct following manufacturer instructions of the SNAP/CLIP-tag®-Trap Agarose (Chromotek) kit. *Briefly:* a 75 cm² flask of cells (~10⁶ – 10⁷ cells) per Atxn2 construct transfection was spun down and washed 3 times in ice cold PBS, the cell pellet was drained and flash-frozen in liquid nitrogen. Cells were lysed by re-suspending the frozen pellet in 200 µl of ice-cold lysis buffer (10 mM Tris/Cl, pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5 % Nonidet™ P40 Substitute, supplemented with protease inhibitor cocktail and 1 mM PMSF). Each tube was then incubated on ice for 30 min, pipetting every 10 min. Lysate was clarified by centrifugation at 17,000x g for 10 min at 4°C and the cleared lysate was transferred to a fresh, pre-cooled Eppendorf. This was then diluted with 300 µl of dilution buffer (10 mM Tris/Cl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, supplemented with protease inhibitor cocktail and 1 mM PMSF). Of this, 50-100 µl per sample were saved as 10-20% input control. For the immunoprecipitation, beads were prepared by washing three times in 500 µl of ice-cold dilution buffer, then centrifuged down at 2,500x g for 5 min. After this, the diluted lysate was added. Lysate and beads were incubated while rotating end over end for 1 h at 4°C. After this, beads were centrifuged down again at 2,500x g for 5 min, then washed twice with ice-cold wash buffer (10 mM Tris/Cl pH 7.5, 150 mM NaCl, 0.05 % Nonidet™ P40 Substitute, 0.5 mM EDTA, supplemented with protease inhibitor cocktail and 1 mM PMSF). After washing, beads were drained, and 50-100 µl of room temperature acidic elution buffer (100 mM citric acid, pH 3.0) was added to the beads and incubated for 60 s while pipetting continuously. Beads were then centrifuged down at 2,500x g for 2 min at 4°C, the supernatant was removed and neutralized using 10 µl of neutralization buffer (1 M Tris, pH 10.4). Samples were then boiled in NuPAGE LDS sample buffer and run on gels as appropriate. For each sample, up to 10 µg of total protein was loaded per well for detecting Atxn2-SNAP constructs, partner proteins and loading controls on 8-12% SDS-PAGE gels and transferred to nitrocellulose membranes. After blocking for an hour in 5% milk in PBS, they were probed using rabbit anti-SNAP (1:1000), rabbit anti-PABP (1:1000), rabbit anti-Lsm12 (1:1000) antibodies, and rabbit anti-histone H3 (1:5000) and rabbit anti-BAF155 (1:2000) loading control antibodies. Corresponding HRP-conjugated secondary antibodies were used between 1:5,000 - 1:10,000 dilution and developed using Pierce ECL western blotting substrate (ThermoFisher) as per the manufacturer's instructions.

For Atx2-SNAP construct immunoprecipitation, transfected cell lysates were normalised to the same volume and concentration, 10-20% of the volume that was previously saved as input control, and Chromotek anti-SNAP-tag conjugated agarose beads and IP kits were used according to the manufacturer's specifications. Pulled-down proteins together with corresponding sample input controls were blotted as described above.

2.5.3 Western blotting of transgenic fly heads

For the fly head WBs, a previously established protocol was adapted (Emery, 2007). Briefly: different *Drosophila* samples (each sample had a different genotype, incubation temperature and age) were collected and frozen at -80° C. Once a sufficient number of flies was collected, 5 individuals per sample were decapitated (flies were kept in dry ice during the procedure) and heads were collected. The heads were immediately squished using pellet pestles (30 s) in 30 µl of extraction buffer (20 mM HEPES, pH 7.5, 100 mM KCL, 5% Glycerol, 10 mM EDTA, 0.1% TRITON, 1 mM DTT, 0.5 mM PMSF, protease inhibitor cocktail) + 10 µl NuPAGE LDS sample buffer. Samples were then boiled at 100 °C for 3 min, spun down and 30 µl of supernatant were removed to a fresh Eppendorf tube. Gel electrophoresis and WB were then carried out on the samples obtained as described previously.

2.6 Spectroscopic techniques

2.6.1 Thioflavin T fluorescence spectrophotometry

Thioflavin T fluorescence spectrophotometry (ThT) was carried out on a FP-6200 Spectrofluorometer (Jasco) (Biancalana & Koide, 2010). A fresh ThT working stock was made for each measurement, by adding 100 µL of stock ThT (10 mM ThT, 10 mM Na₂HPO₄, 150 mM NaCl, pH 7.0, sterile filtered) to 5 mL of 0.1 M Na₂HPO₄ buffer, 100 mM NaCl, pH 6.5. For each measurement 10 µL of protein (100 µM) to 1 mL of ThT working stock in a quartz cuvette. Samples were excited at $\lambda = 440$ nm and spectrums were recorded between 470-500 nm. For plate-reader experiments, a final concentration of 250 µM ThT was used, and measurements were automatically taken every 15 minutes (Ex: $\lambda = 440$ nm, Em: $\lambda = 470$ nm).

2.6.2 Circular dichroism spectropolarimetry

Far UV circular dichroism (CD) spectropolarimetry was carried out using a J-815 CD spectropolarimeter (Jasco). The optical chamber was fully purged with nitrogen gas 30 minutes before use. Far UV CD was performed over 197 - 260 nm (monomers) and 195 - 260 nm (oligomers), both in 50 mM Na₂HPO₄ buffer, pH 7.0, sterile filtered. A second,

fresh monomer sample was also subsequently measured (dissolved in dH₂O rather than buffer), over a range of 185 – 260 nm. The concentrations of monomer, oligomer and monomer in dH₂O solutions were determined spectrophotometrically using a NanoDrop ND-1000 UV-Vis Spectrophotometer (Thermo Scientific) and found to be 0.28 mg/mL, 0.14 mg/mL and 0.21 mg/mL respectively, all within the optimal concentration range for FarUV CD spectropolarimetry (0.1-0.2 mg/mL). A 0.1 cm quartz cuvette was used for all the samples (washed between runs with detergent, dH₂O and EtOH). Suitable buffer blanks were also run and subtracted from all the raw spectrums obtained. The data obtained was transformed in mean residue ellipticity (MRE).

2.6.3 Diffusion-Ordered NMR Spectroscopy (DOSY-NMR)

Background

The diffusion coefficient (D) of a single species of molecule can be extracted from DOSY-NMR data by plotting the Intensity of the peaks (I) obtained against the square of the gradient strength (g^2), according to the Stejskal-Tanner equation (Stejskal & Tanner, 1965):

$$I = I_0 e^{-D\gamma^2 g^2 \delta^2 (\Delta - \delta/3)}$$

Where γ is the gyromagnetic ratio, Δ is diffusion time given, and δ is the time for the gradient pulse. This can then be simplified to a more general Gaussian equation:

$$y = Ae^{-Qx^2}$$

Where Q corresponds to the diffusion coefficient multiplied by ($\gamma^2 \delta^2 (\Delta - \delta)/3$ pulse.) For a mixture of different components, the graph will represent a summation of different Gaussian curves so the equation will become:

$$\sum i = A_1 e^{-Q_1 x^2} + A_2 e^{-Q_2 x^2} + \dots + A_i e^{-Q_i x^2}$$

Finally, from the diffusion coefficient it is possible to calculate the R_h by rearranging the Stokes-Einstein equation from:

$$D = \frac{K_b T}{6\pi\eta R_h}$$

To:

$$R_h = \frac{K_b T}{6\pi\eta D}$$

In which K_b is the Boltzmann constant; T is the absolute temperature and η is the dynamic viscosity. Because of sample-to-sample differences, which may require unique calibration of spectrometer parameters, it is more convenient to use an internal standard as a R_h reference, in these experiments the molecule 1,4 – dioxane.

Diffusion-ordered spectroscopy (DOSY) is a widely used NMR technique for the identification of different chemical moieties/compounds contained in mixtures (Wilkins *et al*, 1999). It has been successfully employed to separate small molecules based on their diffusion coefficients (and thus, from the Stokes-Einstein equation, their R_h values). Change in intensity of peaks from the aliphatic resonance regions as these are not affected by exchange. Attempts were made to expand this technique to determine the size of larger molecules, in particular the early steps of protein aggregation. For a mixture of monomers and aggregates of the same species, chemical shifts are by definition almost completely overlapping, so deconvolution of the data obtained is necessary. A size estimate (ideally a DOSY measurement) of the native protein is also necessary as a curve to subtract from the gaussian of the aggregate mix, in order to estimate the contribution of the aggregates to the overall intensity observed.

All nuclear magnetic resonance (NMR) experiments were carried out on an Agilent Technologies 18.8 T (800 MHz) DD2 Premium Compact spectrometer with a triple-resonance, 5 mm enhanced cold probe. Diffusion-ordered spectroscopy (DOSY) measurements were performed at 293 K. The DgscSL_dpfgsc DOSY pulse program was used, which consists of gradient compensated stimulated echo with spin lock using the excitation sculpting solvent suppression method. A spectral window of 13020 Hz was used, with an acquisition time of 2.46 s with a relaxation delay of 3 s. The FIDs were collected with 32000 complex data points with 64 scans. Logarithmically the gradient pulse strength was increased from 3 % to 86 % of the maximum strength of 32767 G/cm in 60 steps. A diffusion time (Δ) of 100 ms and bipolar half-sine-shaped gradient pulses (δ) of 5 ms were applied. Diffusion coefficients were extracted by plotting the DOSY NMR data according to the Stejskal-Tanner equation. In NMR Spectroscopy experiments, it is

necessary to have a chemical shift reference standard in order to ensure the molecule 1,4-dioxane, which is known to behave independently of protein concentration and the folded state of the protein (Shimizu *et al*, 1994), was used as an internal chemical shift reference and hydrodynamic radius calibration reference (3.75 ppm; $R_H = 2.12 \text{ \AA}$) and the resulting diffusion coefficient (D) values calculated. For the HEWL experiments comparing across samples and using native HEWL itself as a hydrodynamic radius calibration reference was deemed best, as with the particular pulse sequence used the protein sizes across samples exhibited significantly less variance than the dioxane. Peaks were chosen from the aliphatic region of the spectrum, far from the amide region where proton exchange may be occurring and far from the water peak. The data obtained was then fit sequentially using MATLAB, starting from the monomer value and then adding further components to the equation, setting the range of the values. Where this was possible, DOSY curves were obtained for protein in both its native state and at different stages of aggregation, population-averaged diffusion coefficients were compared, and the native diffusion coefficient obtained with DOSY was taken as the initial monomer value.

2.7 Microscopy techniques

2.7.1 Atomic force microscopy

Samples were imaged in air using a BioScope Resolve (Bruker) AFM (tapping mode) and in liquid using an MFP-3D-BIO (Asylum Research) AFM (intermittent contact mode) mounted with DNP-10 (Bruker) silicon nitride cantilevers (tip A, nominal spring constant = 0.35 N/m, resonant frequency = 50 kHz). To prepare the samples for AFM imaging, first mica discs (SPI Chem Mica Grade V1, 12 mm Dia x 0.15 mm) were glued to a glass slide. Prior to depositing the samples, the top layers of the mica surface were exfoliated using adhesive tape. Protein samples made previously and stored at -80°C were thawed, diluted in dH_2O to a concentration of $\sim 7 \mu\text{M}$ and 10 μl of the solution was directly applied on the freshly cleaved mica discs. When imaged in air, samples were directly applied on mica, dried, and imaged immediately. When imaged in liquid, the samples were incubated for 60 s, washed 7 times with dH_2O and topped up with 150-200 μl of dH_2O immediately prior to imaging. The particle size distribution over five areas of each sample were measured.

2.7.2 Transmission electron microscopy

Protein samples were imaged using a JEM 1400 TEM. Prior to imaging, 10 μl of αSN sample (7 or 14 μM) were deposited on 300 mesh carbon coated grids (Ted Pella). The solution was displaced with 0.5% glutaraldehyde and subsequently washed with dH_2O .

The sample was stained using uranyl acetate and allowed to air dry before transferring to the microscope for imaging. *TEM imaging was carried out together with Dr Neal Leddy.*

2.7.3 Immunohistochemistry and imaging of cultured cells

Transfected cells on coverslips were fixed with 4% para-formaldehyde in PBS solution for 15 min, followed by three 5 min washes in PBS. Permeabilization was performed on all cells with 0.5% TritonX100 in PBS solution for 3 min, before three more 5 min washes in PBS. Cells were blocked with 3% BSA in PBS for 1 h 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 1 h 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 >12 h at 4 °C, and imaged on a Zeiss LSM880 Airyscan/AiryscanFast confocal microscope with a 20x air objective.

2.8 Cell culture

2.8.1 Isolation of primary cortical neurons

Primary cortical neurons were isolated from 0-24 hr old neonatal Wister Rats. Neonates were decapitated using a surgical scissors, cortices were removed and chopped using a sterile scalpel blade and incubated at 37° in a 0.2 mg/ml Trypsin/PBS solution for 25 mins to allow sufficient degradation of rat cortices into individual components. After incubation in trypsin solution increasing concentrations of trypsin inhibitor (PBS containing 0.1 mg/ml Soybean Trypsin Inhibitor, 0.06 mg/ml DNase I and 3 mM MgSO₄, followed by 0.1 mg/ml Soybean Trypsin Inhibitor, 0.2 mg/ml DNase and 10 mM MgSO₄) were added to the cortices in order to terminate the proteolytic action of the enzyme. Once the cortices were sufficiently broken up, they were passed through a 40 µm cell strainer (BD) into 50 mL falcon tubes. The cortices were then centrifuged for 3 minutes at 2000 g. Supernatant was then removed and the cell pellets were resuspended in freshly prepared Neurobasal Medium (1x) (NBMA, Invitrogen) supplemented with 10% Horse Serum, 2% Vitamin B27, 2mM L-Glutamine and Penicillin-Streptomycin (100 units/ml Penicillin, 0.1 mg/ml Streptomycin and counted. Typically, preparations yielded 3-5 x 10⁶ cells/ml. Following

isolation cells were plated in clear sterile 96 well plates previously coated with Poly-D lysine (40 µg/ml, M_R 70-150 kDa) at least 2 hrs prior to plating of the cells. After 3 hours media was exchanged for 1% Horse Serum Neurobasal medium (1x) supplemented as described above. Cells were allowed recover for 1 day before treating with 20 µM cytosine beta D-arabinofuranoside (Ara-C) a chemotherapeutic drug which prevents proliferation of glial cells for 24 hrs and then switched to non-serum containing medium (Ara-C free) which they are maintained in until conducting experiments. Media was exchanged every 2-3 days with experiments conducted between DIV 8-11. *Experiments with primary cortical neurons were carried out by Dr Gavin Davis.*

2.8.2 Cell culture, transfection and stress induction of S2R+, HEK293T and U2OS cells

Drosophila S2R+ cells were obtained from the DGRC, Indiana University and were grown in Schneider's S2 media (Gibco) with 10% FBS and 1% penicillin and streptomycin, at 25°C. Transfections were performed using either EugeneHD (Active Motif) or TransIT-X2 (Mirus) reagents at a ratio of 2 µl of reagent per µg of plasmid DNA for 24-72 h depending on downstream use. HEK293T cells were received from Professor Adrian Bracken (Trinity College Dublin) and were grown in Dulbecco's Modified Eagle Media (Gibco) with 10% FBS, 2 mM l-glutamine addition, 1% penicillin and streptomycin, at 37°C and 5% CO₂. U2OS cells were obtained from Professor Martina Schroeder (Maynooth University) and were grown at the same conditions as HEK293T. Mammalian cell transfections performed using 1 mM polyetheleneimine polymer (PEI, Polysciences) solution at a ratio of 2 µl reagent per µg of plasmid DNA for 24-72 h depending on downstream use. When used for confocal imaging applications cells were grown in 24-well plates on glass cover-slips for 24 h before transfection for up to 48 h. For Western blotting and IP, cells were grown in 75 cm² flasks until >80% confluent before transfection for up to 72 h before harvesting. Where relevant, oxidative stress was induced in *Drosophila* S2R+ cells through the addition of sodium arsenite solution to a final concentration of 50 µM in media for 3 h. In mammalian cells, oxidative stress was induced in the same way except for only 1 h.

2.8.2 Cell Viability

Cell Viability was assessed using Alamar Blue® which contains the fluorescent indicator resazurin that is converted to the highly fluorescent compound resorufin which it encounters molecular oxygen. Primary Cortical Neurons were isolated as described and plated into clear 96 well plates at a density of 50×10^6 cells/well in 200 µL of Neurobasal

Medium (NBMA, Invitrogen) supplemented with 10 % Horse Serum, 2 % Vitamin B27, 2 mM L-Glutamine and Penicillin-Streptomycin (100 units/ml Penicillin, 0.1 mg/ml Streptomycin). α SN oligomers were added to each well at varying concentrations and incubated for 24 hrs. 4 hours prior to assessing viability Alamar Blue® was added (10 %) to each well and incubated for 4 hours to allow sufficient detection of viable cells with the probe. Fluorescence was then measured using excitation λ = 570 nm and emission λ = 600 nm. Data was expressed as % Control cell viability. *Experiment carried out by Dr Gavin Davis.*

2.9 *In vivo* electrophysiology

All experiments were carried out in accordance with the approval of the Health Products Regulatory Authority, Ireland, using methods similar to those described previously (Harte *et al*, 2015b). Adult male Lister hooded rats were anaesthetized with urethane (1.5 g/kg, i.p.). Single pathway recordings of EPSPs were made from the stratum radiatum in the CA1 area of the dorsal hippocampus in response to stimulation of the ipsilateral Schaffer collateral/commissural pathways. Test EPSPs were evoked at a frequency of 0.033 Hz and at a stimulation intensity adjusted to give an EPSP amplitude of 50 % of maximum. Adult male Wistar rats were anaesthetized with urethane (1.5 g/kg, i.p.). Single pathway recordings of field excitatory postsynaptic potentials (EPSPs) were made from the stratum radiatum in the CA1 area of the dorsal hippocampus in response to stimulation of the ipsilateral Schaffer collateral/commissural pathways. Test EPSPs were evoked at a frequency of 0.033 Hz and at a stimulation intensity adjusted to give an EPSP amplitude of 50 % of maximum. To inject samples into the rat brain a cannula was implanted in the lateral cerebral ventricle (coordinates: 1 mm lateral to the midline and 4 mm below the surface of the dura) just prior to electrode implantation. Injections (5 μ l over 5 min) were made via a Hamilton syringe connected to the internal cannula. *Experiments carried out by Dr Igor Klyubin.*

2.10 Statistical treatment of the data

For statistical analysis of α SN oligomer cytotoxicity on primary cortical neurons, data was expressed as mean $\pm$ SEM % control viability and analyzed through one-way ANOVA & Dunnett's multiple comparison test ($F = 325.9$, $df = 17$). For statistical analysis of *in vivo* electrophysiology experiments, the strength of synaptic transmission was expressed as a percentage of EPSP amplitude recorded over a 30 min period. The magnitude of changes

was measured at 90 min post-injection and expressed as the mean $\pm$ SEM % pre-injection baseline. Paired Student's t-tests were used to compare the changes within one group. A p value of <0.05 (two-tailed) was considered statistically significant. Regarding the TRIBE analysis, the sequencing reads obtained had a mean quality score (Q-Score) ≥ 37 . Analysis of the TRIBE data was performed as described previously (McMahon et al., 2016, Singh et al., 2021). Briefly, the reference genome and gtf file of *Drosophila melanogaster*, version dm6, were downloaded from the UCSC genome browser. Raw sequencing reads were mapped using TopHat2 (Trapnell *et al*, 2009) with the parameters '--library-type fr-firststrand -m 1 N 3 -read-edit-dist 3 p 5 g 2 -l 50000--microexon-search --no-coverage-search -G dm6_genes.gtf'. Only uniquely mapped reads are considered for editing analysis. A table of raw and mapped reads is included in Supplementary Table 1. A threshold file was created by ensuring only edits with coverage of at least 20 reads and 15% edits were retained. All the TRIBE experiments were performed in duplicates, and only the edits identified in both replicates above the edit threshold are reported. Regarding the climbing assay the average of 3 trials were used for each biological replicate. 7-10 biological replicates were used for each genotype.

3. TIME RESOLVED MONITORING OF α -SYNUCLEIN OLIGOMERS

3.1 Introduction

3.1.1 Parkinson's disease

Parkinson's disease (PD) is a neurodegenerative disorder characterized by the death of dopaminergic neurons in the *substantia nigra pars compacta* (SNpc), which causes dopamine deficiency in the nigro-striatal tract. Affected neurons exhibit intracellular inclusions known as Lewy bodies (LB). These contain the protein α -synuclein (α SN) in an aggregated state (Poewe *et al*, 2017). Clinically, it's primary symptoms include bradykinesia (slowed movements), resting tremors and rigidity, although a variety of non-motor function symptoms appear much earlier (Armstrong & Okun, 2020) (**Figure 3.1**). It is the second most common neurodegenerative disease after Alzheimer's (AD), with a worldwide incidence of 5 – 35 cases every 100'000 individuals per year. It is rare in individuals < 50 years of age but it's incidence increases between 5 – 10 fold in individuals 60 – 90 years old, affecting 2 – 3 % of people > 65 years of age (Savica *et al*, 2013; Twelves *et al*, 2003). Interestingly, it is twice as common in males than in females, although it's unclear if this is due to (i) protective effects of female sex hormones, (ii) a sex-associated genetic mechanism or (iii) sex-specific environmental risk factors. Other environmental factors have been shown to increase (exposure to pesticides) or decrease (smoking, consumption of caffeine) the risk of developing PD (Ascherio & Schwarzschild, 2016; Kouli A, 2018).

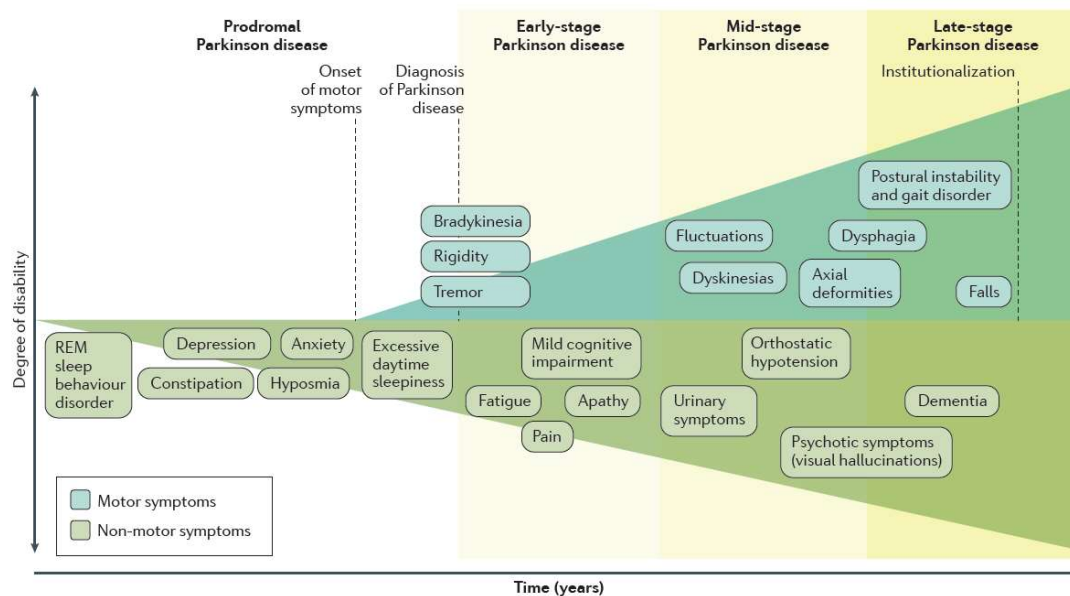


Figure 3.1: Clinical progression of Parkinson's disease. While a PD diagnosis typically occurs after the appearance of its classical motor symptoms (bradykinesia, rigidity and resting tremor), prodromal symptoms usually appear decades earlier. These include altered sleep patterns, constipation and loss of sense of smell. (*Adapted from: Poewe et al, 2017.*)

Most PD cases (~85%) are sporadic, with the remaining 15% being familial. However for PD it is hard to trace a clear line between the two, as it appears that multiple genetic risk factors act synergistically, affecting the likelihood of an individual to develop both sporadic or familial PD, depending on the environmental factors he/she encounters. So far, no single-gene mutation has been discovered that has 100% penetrance for PD (Ascherio & Schwarzschild, 2016; Tran *et al*, 2020).

While familial cases make up the minority of PD cases, the mutations identified have helped reveal which molecular mechanisms, when perturbed, can give rise to PD. Taking these as a starting point, different processes were identified that seem to play a driving role in PD (**Figure 3.2**).

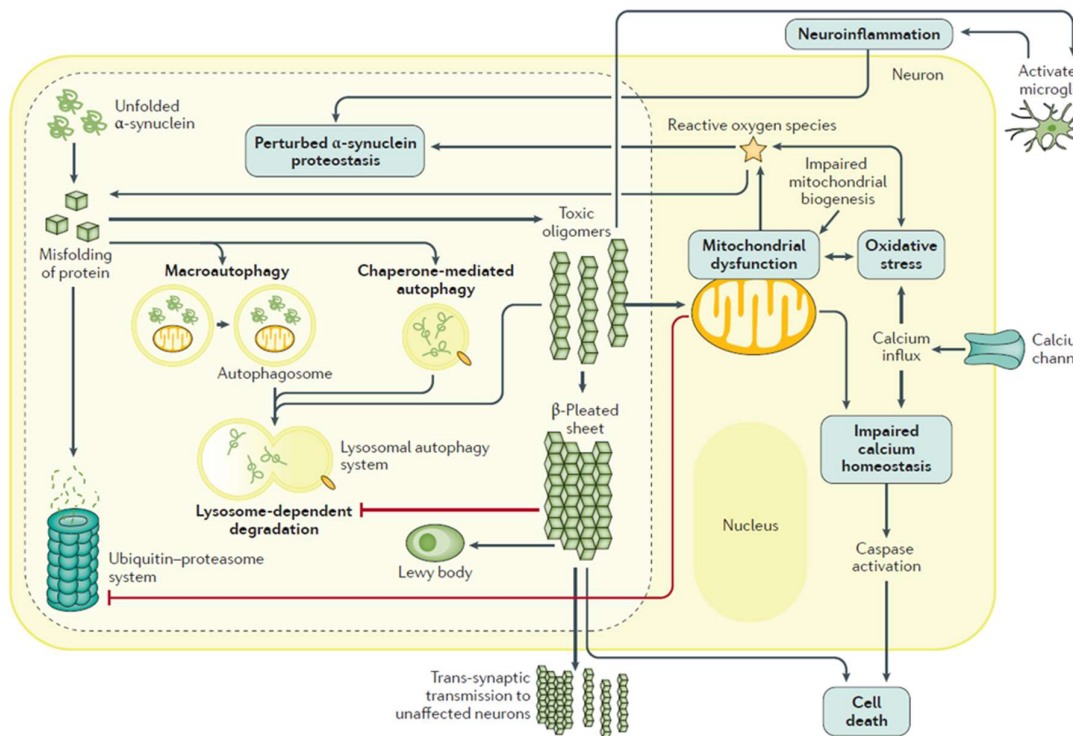


Figure 3.2: Schematic diagram showing different cellular processes implicated with Parkinson's pathogenesis. Many of these appear to be interrelated, and act in a synergistic manner to cause the disease phenotype. (Adapted from: Poewe *et al*, 2017.)

The different molecular pathways that become dysfunctional in PD are in many points tightly interrelated. Briefly, the most important mechanisms seem to be:

- Proteostasis, degradation and spread of α SN.** Genetic mutations involving proteostasis of α SN, including more generally the ubiquitin-proteasomal system (UPS) and the lysosomal autophagy system (LAS) are strong drivers of disease. This is consistent with age being the biggest risk factor for PD, as both these systems appear to become less functional with age (Malik *et al*, 2019; Rao *et al*, 2015). Accumulation of aggregated α SN is thought to play a role in over-burdening cellular proteostatic mechanisms (Fusco *et al*, 2017; Thibaut *et al*, 2018). Misfolded or soluble aggregates of α SN seem to also be capable of spreading from one neuron to another, which might explain the patterns of disease spread that are observed (Braak *et al*, 2003; Luk *et al*, 2012b; Recasens & Dehay, 2014; Volpicelli-Daley *et al*, 2011).

- **Mitochondrial dysfunction.** Activity of complex I of the mitochondrial electron transport chain is reduced in PD. This might be due to α SN aggregates accumulating in the mitochondrion. Impaired mitochondrial function would cause lesser available energy, which in turn may be the root cause of early PD events where dopaminergic neuron axons become dysfunctional (Chan *et al*, 2010; Poewe *et al.*, 2017).
- **Oxidative stress.** Tightly related to mitochondrial dysfunction. There is much evidence of oxidative stress being upregulated in PD patient brain tissue. Dopaminergic neurons might be particularly susceptible to oxidative stress due to their long unmyelinated axons and multitude of synapses, which in turn have high energy requirements to be maintained (Chan *et al.*, 2010; Giguère *et al*, 2018; Surmeier *et al*, 2017).
- **Neuroinflammation.** Inflammatory reactions of glial cells surrounding PD affected neurons (possibly in response to secreted misfolded α SN or increased expression of immune receptors) may exacerbate pathology. It is unlikely though these inflammatory reactions are the initial triggers of PD, but rather a consequence of other disease-driving phenomena (Hirsch *et al*, 2012; Troncoso-Escudero *et al*, 2018).

While there are various treatments available to manage the symptoms of PD and improve patient quality of life, notably supplementation of dopamine (DA) or its precursor L-Dopa, and more recently deep brain stimulation (Limousin & Foltynie, 2019). No treatments targeting the causative agents of PD currently exist, and PD is at this time incurable (Armstrong & Okun, 2020; Kouli A, 2018).

3.1.2 Lewy bodies and α -synuclein aggregates

A defining trait of all PD patients is the intraneuronal presence of LBs mostly made up of aggregated α SN (Spillantini *et al*, 1997). Encoded by the *SNCA* gene which is located on the 4th chromosome, α SN is a small (140 amino acids, ~14.5 kDa) acidic (pI = 4.7) protein, found to be intrinsically disordered in aqueous solution but capable of forming α -helical structures upon binding to lipids (Fink, 2006). Single point mutations and gene multiplications of the *SNCA* gene can cause familial PD, implying a strong causative role of α SN in PD (Poewe *et al.*, 2017). The native functions of α SN are still poorly understood, but it seems to be broadly involved in synaptic vesicle dynamics and trafficking (Bellani *et al*, 2010; Man *et al*, 2021), mitochondrial functions (Faustini *et al*, 2017), possibly may work as a chaperone and has shown evidence of auto-proteolytic activity (Rekas *et al*, 2012; Vlad *et al*, 2011). In particular, it has been shown to associate with highly curved

phospholipid membranes, such as those of small vesicles involved in synaptic transmission (Man *et al.*, 2021; Middleton & Rhoades, 2010). Its sequence can be roughly divided in three domains (**Figure 3.3, A & B**): the amphipathic N-terminal domain, which has α -helical propensity upon phospholipid binding and probably plays an important role in the native function of α SN, the central non-A β -component domain (NAC) which is aggregation-prone and can also form α -helices, and the acidic, negatively charged C-terminal domain, capable of Ca^{2+} binding. Across the N-terminal and the NAC domains there are eleven KTKEGV repeats, which play a role in α -helix fold formation (Poewe *et al.*, 2017; Ulmer *et al.*, 2005). Amyloid fibrils of α SN, as well as proto-fibrils and early oligomers, have been structurally characterized to a certain extent (Li *et al.*, 2018); two helically winding filaments, each arranged at its core in a characteristic greek-key-like motif (**Figure 3.3, C**), make up the α SN fibril. Regarding the oligomers, a variety of different morphologies have been described, some of which are thought to be cytotoxic and initiators of pathology spread (Chen *et al.*, 2015; Danzer *et al.*, 2007; Fusco *et al.*, 2017; Rekas *et al.*, 2010). Interestingly, most of the pathogenic mutations in the *SNCA* gene associated with early-onset familial PD are situated at the dimer interface of the fibril and appear to have a stabilizing effect on the structure (Li *et al.*, 2018).

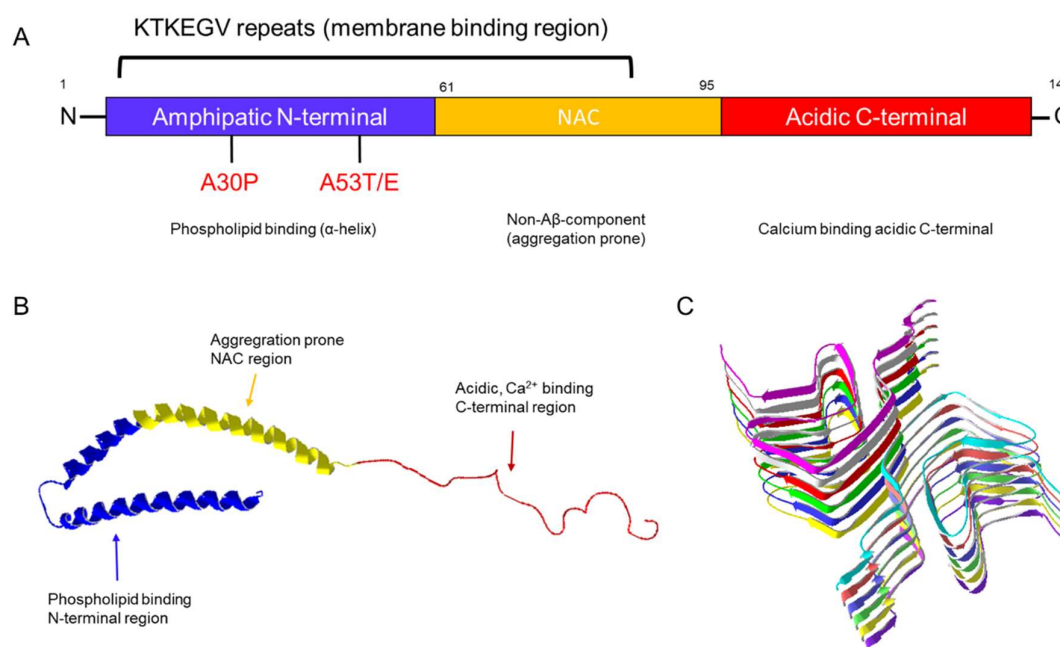
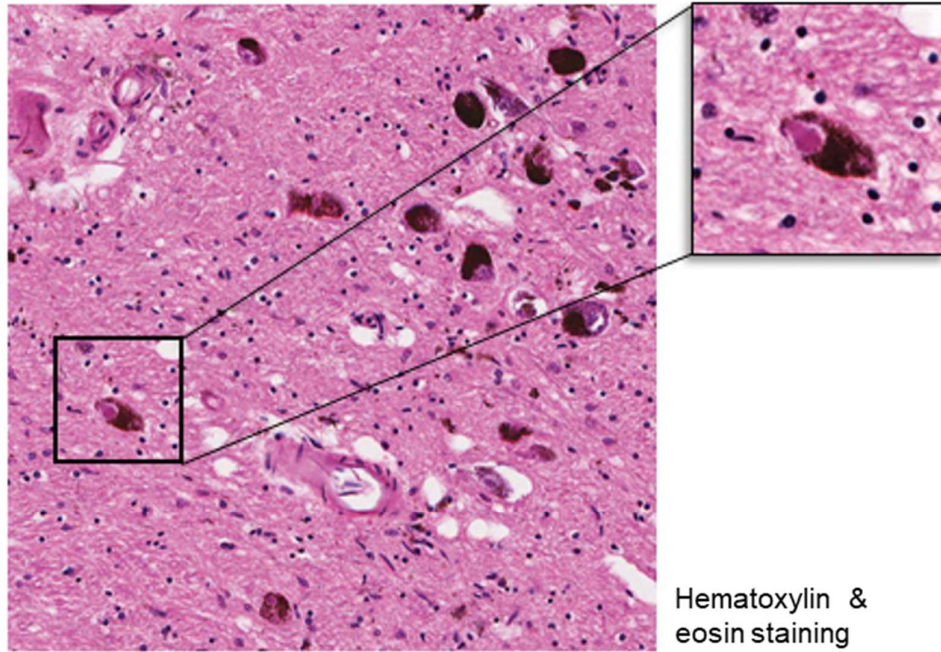


Figure 3.3: Domains and structural features of α SN. (A) The protein α SN can be structurally subdivided into three domains: the amphipathic N-terminal domain, which has α -helical propensity and is capable of binding phospholipids (residues 1 - 60), the central non-A β component (NAC) which plays a role in α SN aggregation (residues 61 - 95) and finally the

negatively charged C-terminal domain, which is capable of binding Ca^{2+} ions. The N-terminal and central regions contain 11 interspersed repeats of the KTKEGV sequence, which are involved in helix formation. Several pathogenic mutations involved in familial PD (or risk factors in sporadic PD) are found across the N-terminal and the NAC domain (B) Structure of micelle-bound αSN as determined by NMR spectroscopy (PDB ID: 1XQ8). Upon phospholipid binding, the N-terminal and NAC region seem to undergo a disorder-to-order transition and acquire an α -helical structure (Ulmer *et al.*, 2005). (C) Cryo EM structure of the αSN fibril core (PDB ID: 6A6B). Fibrils form as a dimer of two twisting filaments, the fibril core arranging in a characteristic greek key-like motif (Li *et al.*, 2018).

Fibrillar αSN is the major component of LBs. Apart from αSN , these structures contain a variety of other proteins, including ubiquitin, heat shock proteins (HSPs), tau, cytoskeletal components and lysosomal markers (Poewe *et al.*, 2017). LBs can range in size between 5 – 30 nm, they appear as granules in affected neurons (Kouli A, 2018). Upon immunohistochemistry staining for αSN a characteristic αSN “halo” can be observed, spreading around the compact core (**Figure 3.4**).

A



B

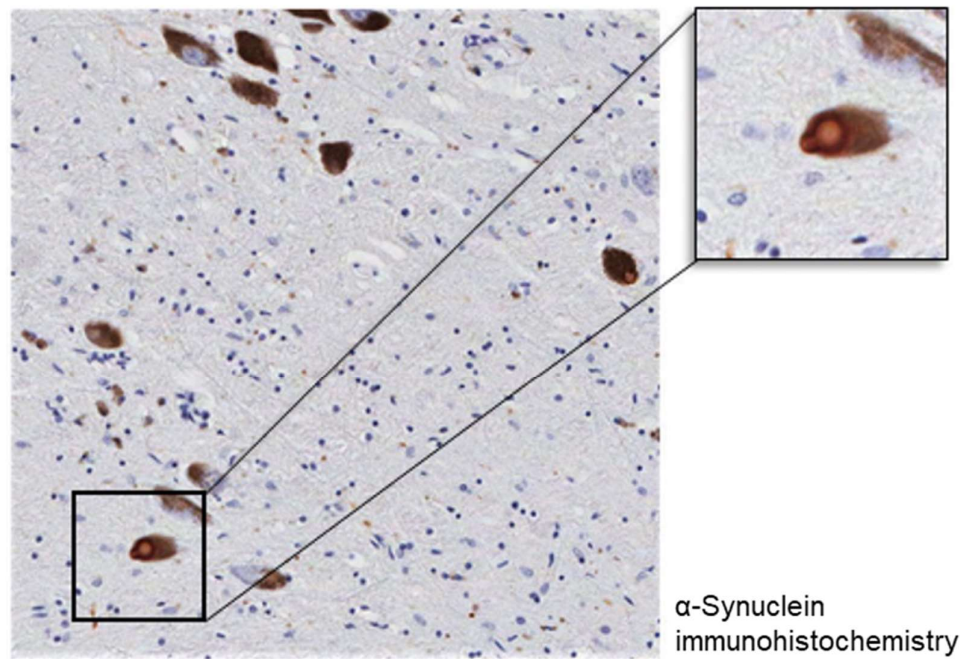


Figure 3.4: Lewy body stainings in neurons of the nigrostriatal pathway from PD patients. (A) Hematoxylin & eosin staining of neuromelanin-containing dopaminergic neurons from the nigrostriatal pathway of a PD affected patient. LBs can clearly be observed. (B) αSN immunohistochemistry staining of a similar sample. This staining method is much more sensitive and the characteristic LB halo of αSN can be observed (*Adapted from: Kouli A, 2018*)

Many lines of inquiry suggest that α SN aggregation is a central cause of PD, and seems to play some role in the molecular pathways perturbed in PD patients. Various potential mechanisms of α SN toxicity have been suggested and experimentally tested to variable extents:

- Aggregation of α SN might disrupt its native function, which could be related to DA vesicle trafficking and release.
- Mutant/aggregated/post-translationally modified forms of α SN might impair mitochondrial function in general, or directly inhibit complex I activity.
- Aggregation of α SN might drive ER Stress by disrupting ER-Golgi trafficking.
- Aggregation of α SN may overload protein clearance pathways such as the UPS and the LAS, causing a breakdown in normal cellular functions.
- Misfolded/aggregated α SN, when released in the intercellular space may trigger inflammatory reactions from nearby glial cells, in turn causing neuronal injury and death.

3.1.3 The role of pathology spreading & selective vulnerability in PD

There are many reasons why dopaminergic neurons in the SNpc might be particularly vulnerable to the adverse effects caused by α SN aggregation (**Figure 3.5**), such as:

- DA SNpc neurons have an exceptionally long and branched unmyelinated axonal arbor with a large amount of synapses. This in turn causes these cells to be highly metabolically active, increasing the risk of free radical oxidative damage (Chan *et al.*, 2010).
- The large amount of synaptic terminals also means higher vesicular trafficking activity, which might mean endogenously high levels of α SN, increasing the likelihood of aggregation.
- They exhibit autonomous pacemaking, so consistently undergo slow oscillations of intracellular Ca^{2+} concentration. These are thought to drive levels of mitochondrial oxidative phosphorylation, which in turn is thought to cause higher levels of free radicals that would cause oxidative damage (Poewe *et al.*, 2017).

In general, a general breakdown of cellular bioenergetics associated with malfunctioning proteostasis machinery in DA SNpc neurons seems to be driving cell death in PD (Surmeier *et al.*, 2017).

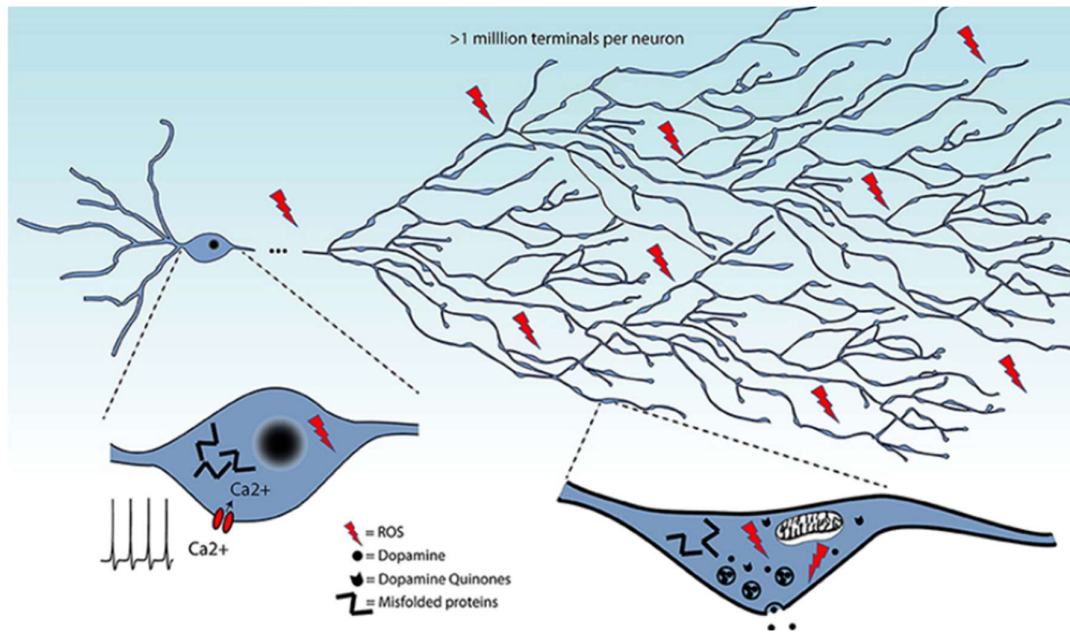


Figure 3.5: Summary of putative PD vulnerability determinants of dopaminergic neurons from the *substantia nigra pars compacta*. DA SNpc neurons have an extremely large and branched axonal arbor. This is associated with high energy demands, meaning these cells are very metabolically active, with high levels of oxidative phosphorylation, which causes an increased presence of ROS. Furthermore, DA SNpc neurons exhibit autonomous pacemaking, which mean continuous changes in intracellular Ca^{2+} levels, also driving mitochondrial activity, as well as potentially influencing the phase separation propensity of αSN . Due to large amounts of vesicular trafficking in these neurons, local concentrations of αSN are probably higher than in other neuron types. Some intermediate steps in dopamine synthesis produce highly reactive dopamine quinones, which may increase oxidative damage. (Adapted from: Giguère *et al.*, 2018).

While there is a strong body of evidence that PD pathology initially affects DA SNpc neurons rather than other cell types because of their specific cellular physiology and biochemistry, there is also evidence supporting the idea of PD pathology spreading, specifically through the prion-like transmission of misfolded or aggregated αSN (Goedert, 2015). This hypothesis postulates that following the formation of αSN aggregates in a specific neuron, these (or misfolded monomers) can spread inter-axonally, possibly through exosomes (Tsunemi *et al.*, 2021), into the extracellular space and then be taken up via endocytosis by connected or neighbouring neurons. This misfolded/aggregated form of αSN would then act as a seed for aggregation in these cells, propagating the disease. This model is consistent with the Braak staging hypothesis (Braak *et al.*, 2003; Giguère *et al.*, 2018) which postulates the progressive spread of disease (possibly *via* misfolded αSN) initiating from the gut enteric nerves and the olfactory bulb, which would

explain some of the initial prodromal symptoms (such as loss of sense of smell and constipation). PD pathology would travel through different brain regions, causing motor dysfunction only once the *substantia nigra* is reached. This was based on a series of autopsies of PD patient brains, assessing the presence of LBs in different areas of the nervous system. This hypothesis has been contested, mainly on the grounds that this sort of spreading pattern is not always observed in PD patients (Kouli A, 2018).

While the clear route of disease staging proposed by Braak remains controversial, there is instead very strong evidence of misfolded/aggregated α SN cell-to-cell transmission. This was famously observed in tissue grafted in the brains of PD patients during an experimental clinical trial, which upon dissection, over ten years later, exhibited evidence of LB pathology (Li *et al*, 2008; Recasens & Dehay, 2014). This, together with evidence in different cultured cell types that misfolded α SN can be transmitted from one cell to another suggests that prion-like behaviour of α SN may play a causative role in PD pathology.

3.1.4 Why study α SN oligomers?

The study of early aggregates of proteins known to form amyloid fibrils can be challenging due to (i) the kinetics of amyloid fibril formation, (ii) the intrinsically transient nature of these structures, (iii) the tendency of oligomers to re-equilibrate with monomers upon purification, and (iv) consequent low enrichment of purified samples (Chiti & Dobson, 2017; Fink, 2006). The past decade has seen a rise in interest towards these non-fibrillar aggregates as potential drivers of disease. A prominent example of this being α -synuclein (α SN) in Parkinson's disease (PD), dementia with Lewy bodies, multi system atrophy, and a variety of other neurodegenerative diseases, including Alzheimer's disease (Chen *et al.*, 2015; Chiti & Dobson, 2017; Danzer *et al.*, 2007; Fusco *et al.*, 2017; Goedert, 2015; Levin *et al*, 2014). Previous work from this laboratory has shown that even oligomeric aggregates of innocuous proteins can have cytotoxic effects and inhibit long-term potentiation *in vivo* (Harte *et al.*, 2015b).

Recent investigations have shown that oligomers of α SN can be generated *in vitro* from wild type α SN using a variety of different methods, such as: incubation in artificial solvent systems (Danzer *et al.*, 2007; Munishkina *et al*, 2008; Munishkina *et al*, 2003), chemical modification (Rekas *et al.*, 2010), physical agitation (Kumar *et al*, 2020), and through lyophilization (by increasing the solvent/air interface and consequently decreasing the intermolecular distances between monomers) (Chen *et al.*, 2015). The dimensions of the oligomers (from these studies) show significant polydispersity, ranging from 160 up to even 1300 kDa (as estimated using techniques such as analytical ultracentrifugation, atomic force microscopy, blue native PAGE and cryo-EM) (Chen *et al.*, 2015; Danzer *et al.*, 2007;

Fusco *et al.*, 2017; Ruggeri *et al.*, 2018). Due to the heterogeneity found, it appears that the aggregation pathway is strongly affected by the conditions used, highlighting the need for validated methods to characterize the species produced.

When studying these early aggregates, the first problem encountered is that classical techniques used to monitor amyloid fibril formation such as thioflavin T (ThT) fluorescence spectrophotometry or Congo red birefringence assays are not suited for observing pre-fibrillar species (Fink, 2006). Size exclusion high performance liquid chromatography (SE-HPLC) is a technique commonly employed to visualize, separate and purify molecules according to their hydrodynamic radius (R_H). It is also part of the standard battery of methods used during protein analysis in the biopharmaceutical industry, for example in the characterization of bio-therapeutics and any eventual aggregates occurring during their production (Burgess, 2018; Fekete *et al.*, 2014; Hong *et al.*, 2012). Despite this, it has not been a central analytical technique for classical aggregation studies of pathologically misfolded proteins due to the initial focus on mature amyloid fibrils, which are inaccessible with chromatographic methods. Applied instead to early aggregation stages, SE-HPLC makes it possible to capture the formation of transient pre-fibrillar protein aggregates as a function of time while also providing separate, quantifiable measurements of different sized species at physiological concentrations. In contrast to FRET/FCS methods (Iljina *et al.*, 2016; Nath *et al.*, 2010) it does not require the addition of fluorescent tags or other protein modifications. Finally, SE-HPLC has the potential to separate and isolate different aggregate populations, an important first step towards gathering high-resolution structural data. Provided the kinetics of interconversion is slow, a chromatogram obtained would represent a snapshot of the distribution of particles in a heterogeneous mixture at the time-point of the injection. Aggregates larger than the pore size will be excluded, and any differences in total absorbance intensity can be used to estimate sample loss. Similar problems are encountered when using transmission electron microscopy (TEM) and atomic force microscopy (AFM), as fixing an aggregating protein sample to TEM grids or drying on mica slices for AFM will almost inevitably affect its composition, although for AFM a microfluidic spray device ameliorating this issue has recently been developed (Ruggeri *et al.*, 2018).

In contrast, techniques that can measure the composition of the sample directly in solution without perturbing the aggregation environment come with their own set of difficulties. Various kinds of light scattering measurements are commonly used to track aggregation progress in solution, but require large, non-physiological concentrations, and depending on the intrinsic or extrinsic properties of the particles, may emphasize a certain distribution, or if not, output strictly population-averaged values. Diffusion-ordered NMR

spectroscopy (DOSY-NMR) also requires non-physiological concentrations, but could in theory separate different populations by translational diffusion rates, provided that the issues of overlapping resonance frequencies and short relaxations times are resolved (Morris, 2009; Price, 2002). Recent work from this laboratory proposes an approach that could partially circumvent these problems (Kandiyal *et al*, 2019) and here we show that direct analyses of the decay curves can lead to size estimations of protein oligomeric forms.

In order to probe the structure and stability of early oligomers of α SN, protein samples at physiological concentration were injected in the SE-HPLC at different time points of the aggregation protocol (Danzer *et al.*, 2007) to capture the formation of transient pre-fibrillar protein aggregates in real-time. As a control, this approach was repeated using the non-disease related protein hen egg white lysozyme (HEWL) which generated a very different aggregation profile. HEWL is similar in size to α SN (14.3 KDa) but in contrast to α SN is not intrinsically disordered, having a stable and well-characterized native fold (Weiss *et al*, 2000). Proteins of the lysozyme family are important mediators of the innate immune response, acting as hydrolases that cleave the bacterial peptidoglycan wall (Chipman David & Sharon, 1969; Ragland & Criss, 2017). Although natively well folded, HEWL can also form amyloid fibrils under certain conditions, such as a combination of low pH and high temperatures (Harte *et al*, 2015a). Due to its size similarity and ability to form amyloid fibrils without these being involved in disease, it was chosen as an interesting aggregation kinetics comparison for α SN. Steps of HEWL aggregation have been previously investigated, and it was not expected to form the same sort of oligomer intermediates observed during α SN aggregation (Chaari *et al*, 2015; Kandiyal *et al.*, 2019). The aggregation protocol selected for α SN was chosen over the protocol described by Chen (Chen *et al.*, 2015) as it was better suited to follow the progression of aggregation over time. Furthermore, the initial aim was to use SE-HPLC to isolate different sized oligomers at different time points and test their cytotoxic/LTP effects, and this protocol allowed physiological concentrations and physiological pH. The protein α SN has been shown to participate in LLPS and its phase separation competency has been suggested to play a role in seeding α SN aggregation (Brodin *et al*, 2022; Ray *et al*, 2020; Sawner *et al*, 2021).

3.2 Results

3.2.1 Characterization of α -synuclein monomer and aggregates

Thioflavin T fluorescence spectrophotometry was used (**Figure 3.6**) to monitor a well-established fibrillation protocol at non-physiological concentration (100 μ M) for α SN (Fink, 2006). This showed that fibrillation roughly occurred after 5 days. Based upon this, it was decided to adopt an oligomerization protocol which leads to semi-stable oligomeric aggregates rather than fibrils at physiological concentrations (14 μ M) (Danzer *et al.*, 2007). In both protocols, physical agitation was the main driver of aggregation. Agitation increases the overall enthalpy of the solution, causing a

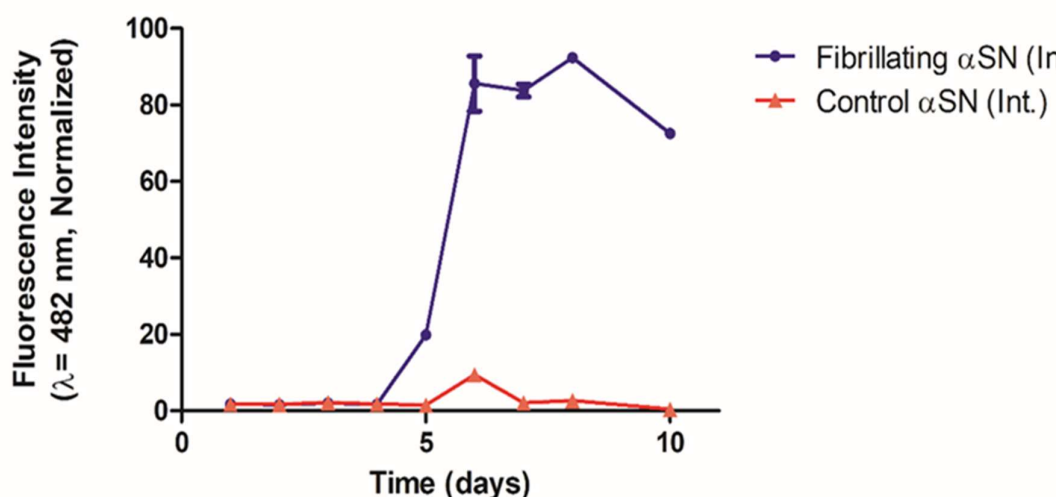


Figure 3.6 Fibrillation of α -synuclein. Thioflavin T fluorescence spectrophotometry was used to monitor fibrillation progression of α SN ($n = 3$). Maximum fibrillation was detected after 5 days.

Earlier events in the progression of α SN aggregation were then investigated using SE-HPLC injections at different time points. SE-HPLC experiments, to periodically monitor column performance over time, calibration standards (**Figure 3.7, Table 3.1**) were injected and the reproducibility of retention times checked. From the initial injections of the non-aggregated α SN sample, the well-documented (Chen *et al.*, 2015) unusual R_H behavior of the apparent α SN monomer was immediately noted, exhibiting an $R_H = 37.4 \pm 1.7$ Å corresponding to an apparent M_R of 65.7 ± 8.1 kDa (**Figure 3.8**), much larger than the calculated M_R of 14.46 kDa. In addition, a minor peak at an apparent M_R of 8.4 ± 2.2 kDa was also observed (Ehrnhoefer *et al.*, 2008; Zhu *et al.*, 2004), and interpreted as proteolytic fragments.

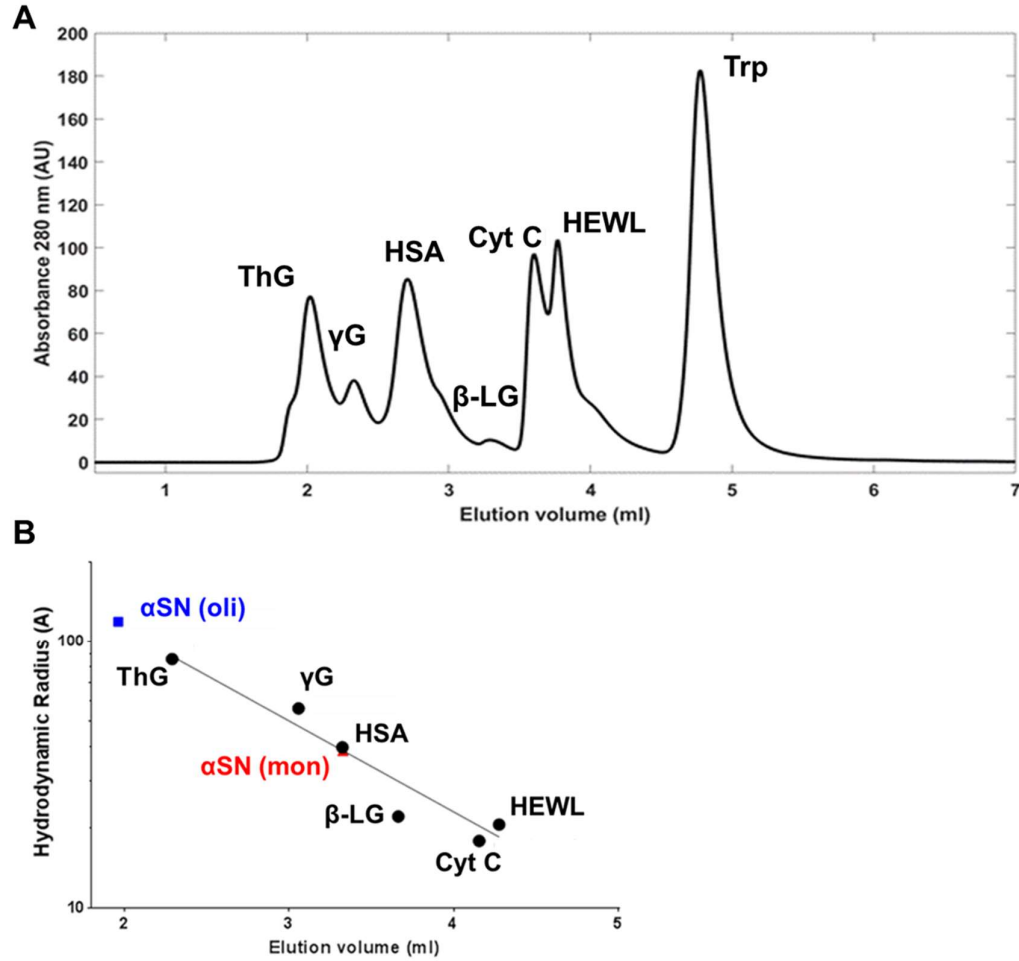


Figure 3.7: SE-HPLC standard curve for protein size determination. (A) Chromatogram of protein mix used to calibrate the column and obtain the standard curve. (B) Standard curve mix periodically utilized to assess column performance. Standards are, in order: Thyroglobulin (ThG, 660 kDa), γ -globulins (γ G, 200-150 kDa, average of two peaks observed), Human serum albumin (HAS, 66.5 kDa), β -lactoglobulin (β -LG, 18.4 kDa), Cytochrome C (Cyt C, 12.3 kDa), hen egg white lysozyme (HEWL, 14.3 kDa). Standards were run individually and as a mixture to ensure correct peak identification.

Table 3.1: Hydrodynamic radii of proteins used for the standard curve

Protein	Number of residues	R_H (Å)	Elution volume (mL)	Source of R_H value
Thyroglobulin	2769	85.8	2.29	(Edelhoch, 1960)
γ -globulin	1364	56	3.06	(Nag <i>et al</i> , 2015)
Human serum albumin	603	40	3.32	(White <i>et al</i> , 2015)
β -Lactoglobulin	162	22	3.66	(Parker <i>et al</i> , 2005)
Hen egg white lysozyme	129	20.5	4.27	(Wilkins <i>et al.</i> , 1999)
Cytochrome C	105	17.8	4.15	(Wilkins <i>et al.</i> , 1999)

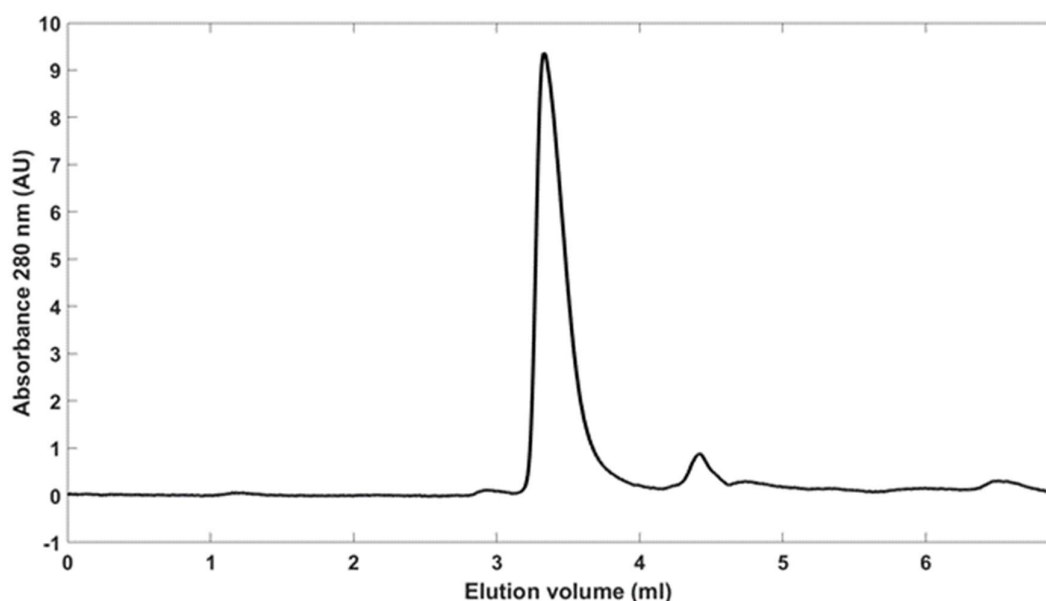


Figure 3.8: Chromatogram of native α SN. Full chromatogram of freshly dissolved α SN in aggregation buffer. A minor low molecular weight peak eluting between 4-5 min was interpreted as a proteolytic degradation product of the protein.

The hypothesis that native α SN exists in a tetrameric state (Bartels *et al*, 2011; Wang *et al*, 2011) was tested to ensure that the higher M_R peak truly corresponded to a α SN monomer and not a tetramer or trimer. Thus the protein was injected in the SE-HPLC column under strong denaturing-disaggregation conditions of 8.0 M urea (**Supplementary Figure 6.1, Supplementary Table 6.1**), the rationale being that the protein would be in a completely denatured and unfolded state under these conditions, and the retention time of the protein (relative to the buffer system) would either increase (if the protein were indeed in a tetrameric state under non-denaturing conditions), decrease (if the protein were in a monomeric but partially folded state) or stay the same (which would mean that α SN is

natively in a fully unfolded state). The SE-HPLC chromatograms appeared similar in both solvent systems, the retention time decreasing slightly, suggesting that there was no further dissociation upon denaturant treatment and that at physiological conditions α SN is found in the previously-characterized intrinsically-disordered conformational monomeric state (Uversky, 2003). The α SN monomer identity was further confirmed through far UV circular dichroism (CD) spectropolarimetry (**Figure 3.9**), showing a classical random coil profile rather than the α -helical structure proposed for the tetrameric form (Bartels *et al.*, 2011).

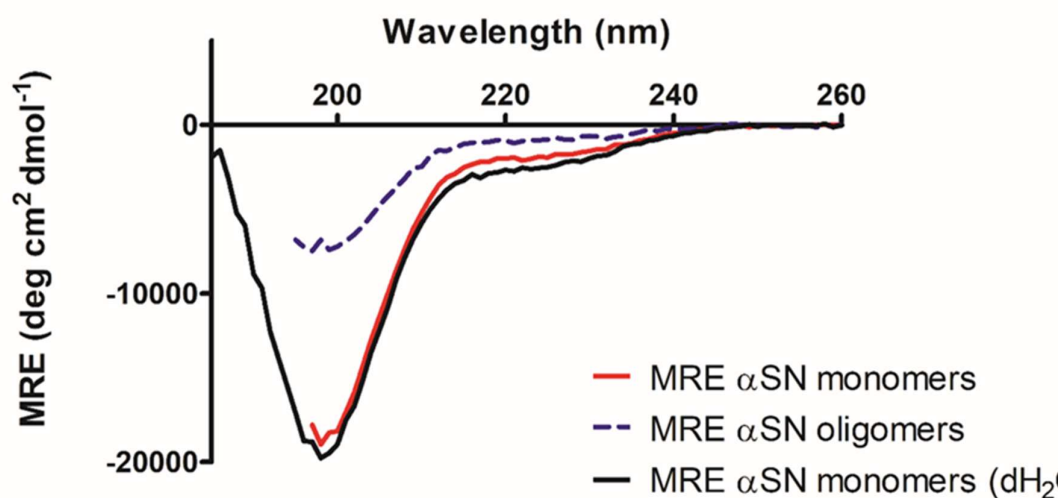


Figure 3.9: Far-UV circular dichroism spectropolarimetry spectra of monomeric and oligomeric α SN. Both monomers and oligomers exhibited a classical random coil signal, consistent with an intrinsically disordered structure. No evidence of α -helical structure was observed.

SDS PAGE analysis (**Figure 3.10**) was also carried out and a fragment corresponding to the lower M_R peak found in the SE-HPLC chromatograms was not detected.

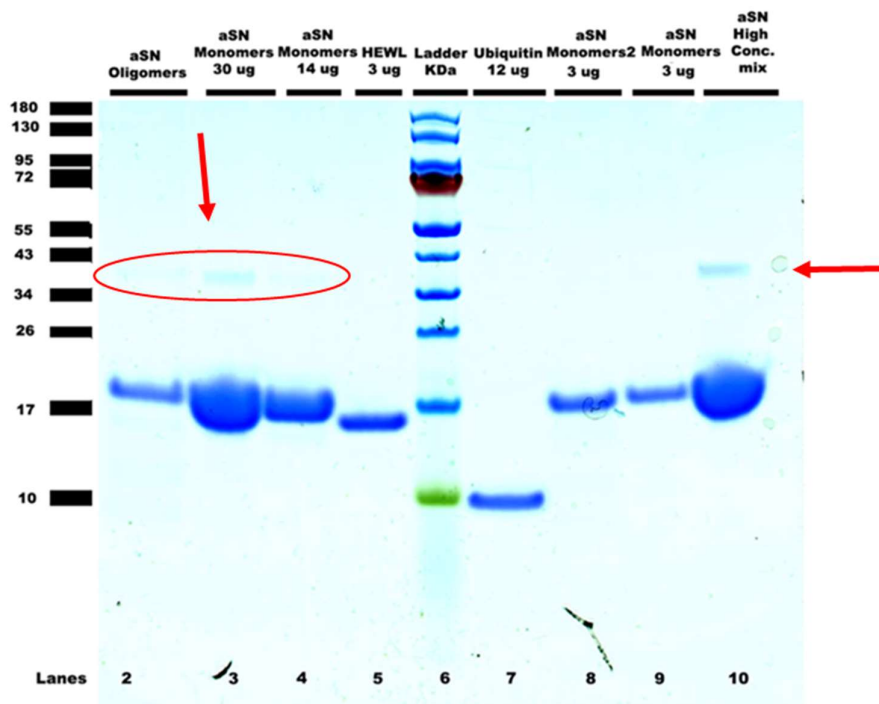


Figure 3.10: SDS PAGE analysis of α SN sample. Following note of an unidentified peak during the SE HPLC experiments, the α SN received was analysed through gel electrophoresis to verify no impurities were present. No smaller fragments were detected using SDS PAGE. A sample of oligomerized α SN was also run, to see if any SDS-resistant oligomers were formed; a faint band is present in the oligomerized and the high concentration samples (marked with red arrows/circled), at a height consistent with a trimeric/tetrameric specie.

3.2.2 Time-resolved monitoring of α -synuclein aggregation

For real time monitoring, different time points of the oligomerization of the α SN solution were injected into the SE-HPLC column directly without any dilution. Chromatograms were obtained at multiple UV-Vis wavelengths, including the monitoring of static light scattering ($\lambda = 3 \text{ nm}$). It was noted that the 280/220 nm signal ratio increased between monomer and oligomer peaks. Decrease in the monomer peak without detection of any other peaks was interpreted as higher order aggregates too large to enter the guard column. During these experiments, aggregates of different sizes were captured beginning 24 h into the aggregation protocol and particularly substantial from 33 h onwards. The two most consistently observed species were found to have a $R_H \sim 150 \text{ \AA}$ and $R_H \sim 110 \text{ \AA}$ (**Figure 3.11**). A very small yet consistently present peak at a $R_H \sim 59 \text{ \AA}$ was also observed. These R_H values closely agree to those obtained by Chen (Chen *et al.*, 2015), and assuming a

sphere, would respectively correspond to 66-mer, 26-mer and tetrameric species (calculated using the formula $V = (4/3)\pi r^3$, with $r = R_H$). Likely, in the protocol used (Danzer *et al.*, 2007) the initial formation of oligomer seeds is very slow/kinetically unfavourable. Their formation is driven by the increase in enthalpy of the system due to the continuous agitation. Thus, the question arose to what would happen to these oligomers the moment agitation was stopped.

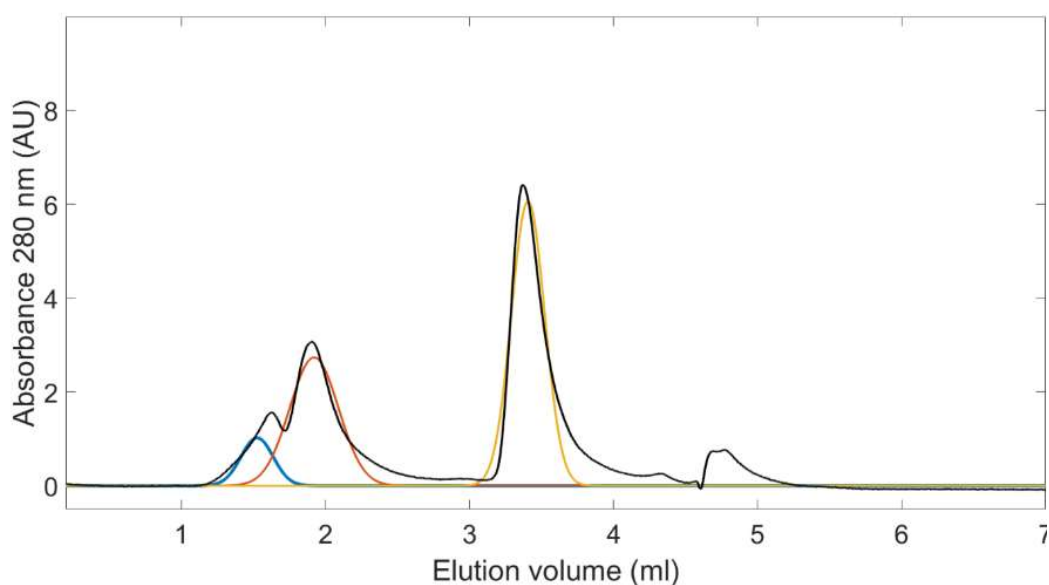


Figure 3.11: Chromatogram of oligomeric aggregates of α SN.) Representative chromatogram ($t = 33$ h) of the most consistently observed oligomer peaks during the aggregation procedure.

In order to assess if agitation was a key requirement in keeping the integrity of the oligomers, it was decided to halt agitation in one of the aggregation experiments at the 36 h time point. Somewhat surprisingly, it was observed that the samples reverted to their monomeric form. Upon resuming agitation, it was possible to detect aggregates about ~ 16 h later (**Figure 3.12**). Interpretation of this observation was that halting agitation causes the enthalpy of the system to drop, making the entropic component dominant, thus driving the kinetics of the reaction in the opposite direction, towards monomeric forms of α SN. The observation that at this early stage, and at physiological concentration/pH aggregation is reversible, suggests that what is being captured in the chromatograms obtained is the transition between physiological LLPS and formation of aberrant soluble aggregates. Attempts to isolate these early aggregates by collecting fractions and then re-

chromatographing failed, most likely due to their reversion towards an equilibrium with the monomeric state. This behavior highlights even more the transient and unstable nature of these aggregates at this early stage. Interestingly, components of even-larger hydrodynamic volumes were observed (**Figure 3.12, B**), suggesting that there may be diverse and larger oligomers possibly formed *en route* to fibrillation.

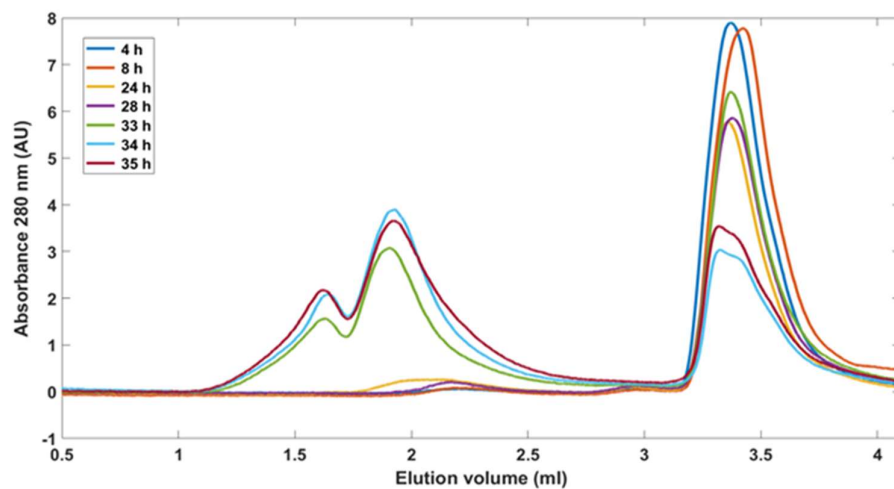
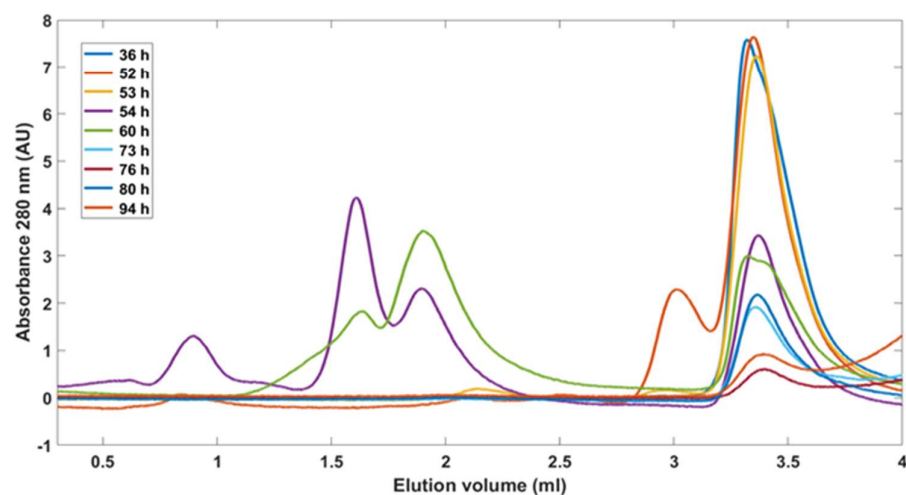
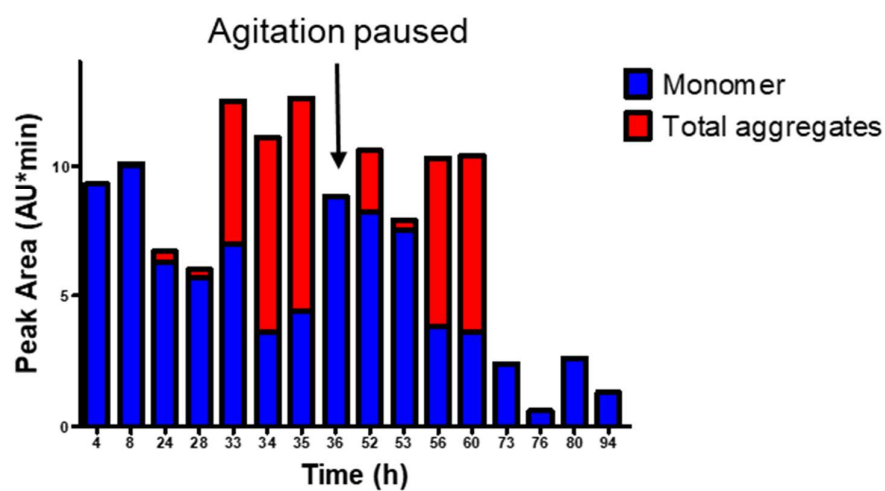
A**B****C**

Figure 3.12: Time resolved monitoring of early α SN aggregation. (A) Overlay of chromatograms obtained during the aggregation procedure (up to when agitation was halted to test stability of the aggregates obtained at this stage of aggregation). (B) Chromatogram overlay

obtained during the aggregation protocol after agitation was restarted. (C) Change in composition of the α SN sample over time. Following the injection at $t = 35$ h, agitation was paused, then restarted after the injection at $t = 36$ h

3.2.3 Time-resolved monitoring of the aggregation of a non-disease related protein

As with the α SN experiments, ThT fluorescence spectrophotometry was initially used to obtain a general idea of the aggregation timeframe of HEWL; fibrils detectable using this method could be observed after ~ 7 days (**Supplementary Figure 6.2**, *experiment carried out by Dr JiYoon Kim*). In this case much higher concentrations (1 mM) were used, and the main drivers of aggregation were low pH and high temperature, following a protocol previously used in this laboratory (Harte *et al.*, 2015b). Samples were diluted before injection but the mobile phase buffer was maintained at the same pH (~ 2.5) that was used during the aggregation protocol. It was immediately evident from the initial injections that differences between samples would be much subtler in this case, so it was deemed necessary to test the robustness of both the aggregation protocol used and the SE-HPLC detection method. Samples of native HEWL and after 72 h were injected in a randomized order, and unlike the α SN cases, were not baseline-separated, yet showed clear differences in hydrodynamic volumes (**Figure 3.13**).

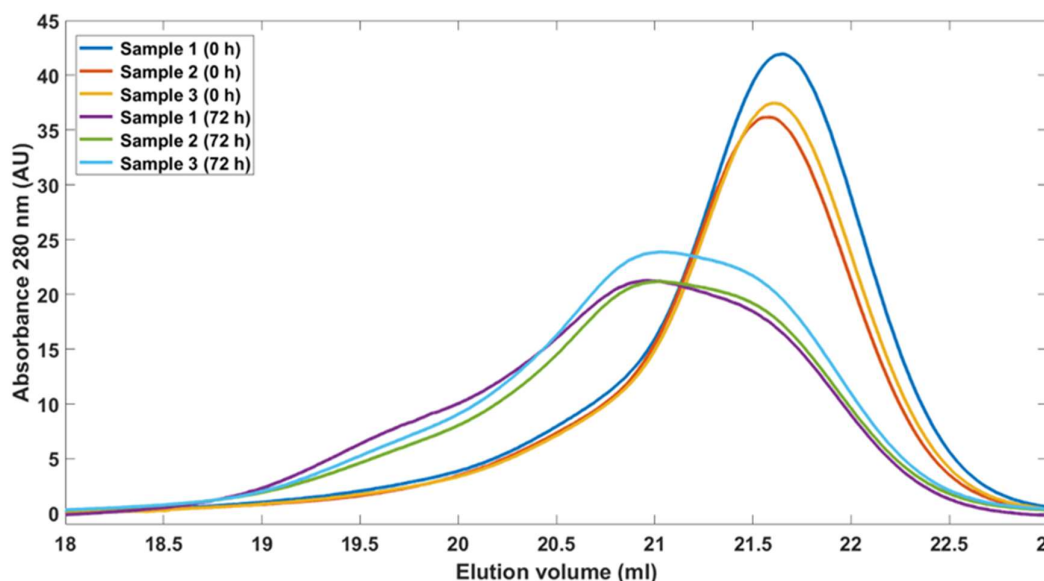


Figure 3.13: Early aggregation of HEWL. Overlay of injections of HEWL aggregates at $t = 0$ h and at $t = 72$ h. Each sample was aggregated independently to assess robustness of aggregation

procedure and SE-HPLC detection methodology. Injection order was randomized to avoid time series effects.

Due to this, peak deconvolution was carried out to quantify the different components and estimate their sizes (**Figure 3.14**). A three-component poly-gaussian fit resulted in obtaining populations which consistently corresponded to monomer ($20.1 \pm 0.2 \text{ \AA}$), dimer ($23.7 \pm 0.2 \text{ \AA}$) and trimers ($28.2 \pm 0.2 \text{ \AA}$).

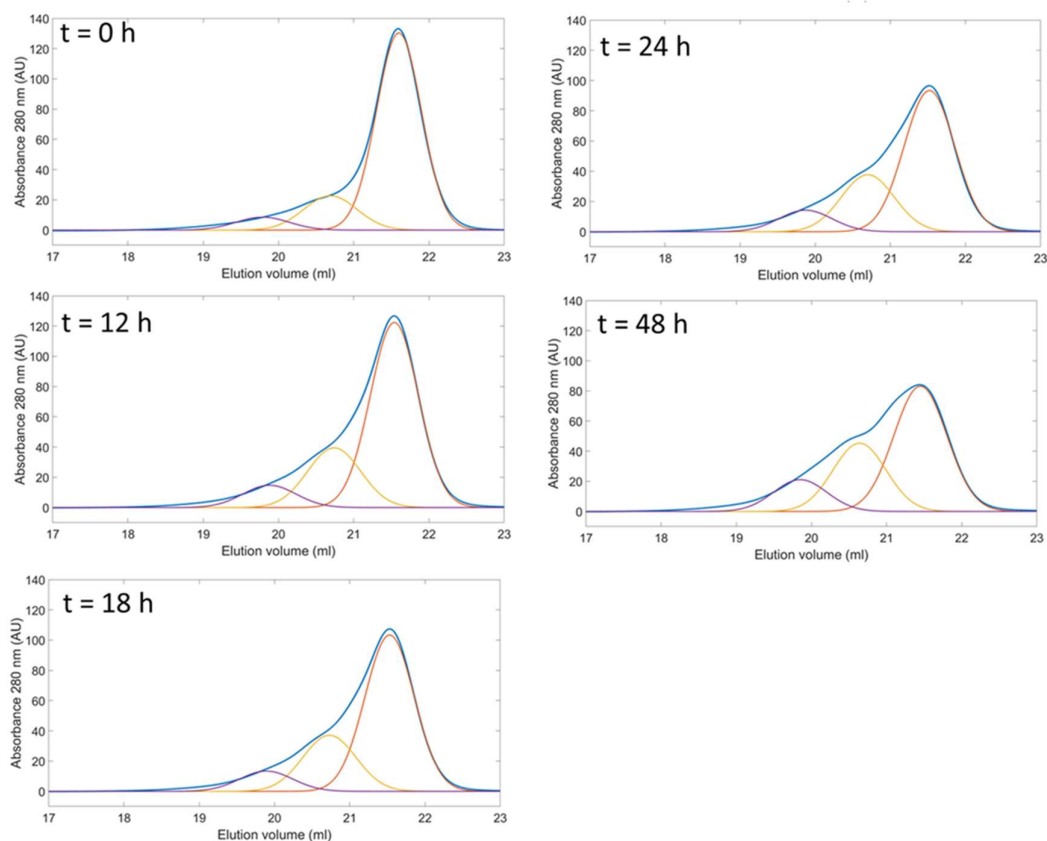


Figure 3.14: Aggregating HEWL peak deconvolution over time. Representative SE-HPLC injection time points of aggregating HEWL. Three component gaussian peak deconvolution was used to estimate the amount of separate components underneath the shoulder of the monomer peak.

Early aggregation of HEWL was measured for a duration of 72 h, exhibiting a progressive shift in composition towards the two larger species (**Figure 3.15**).

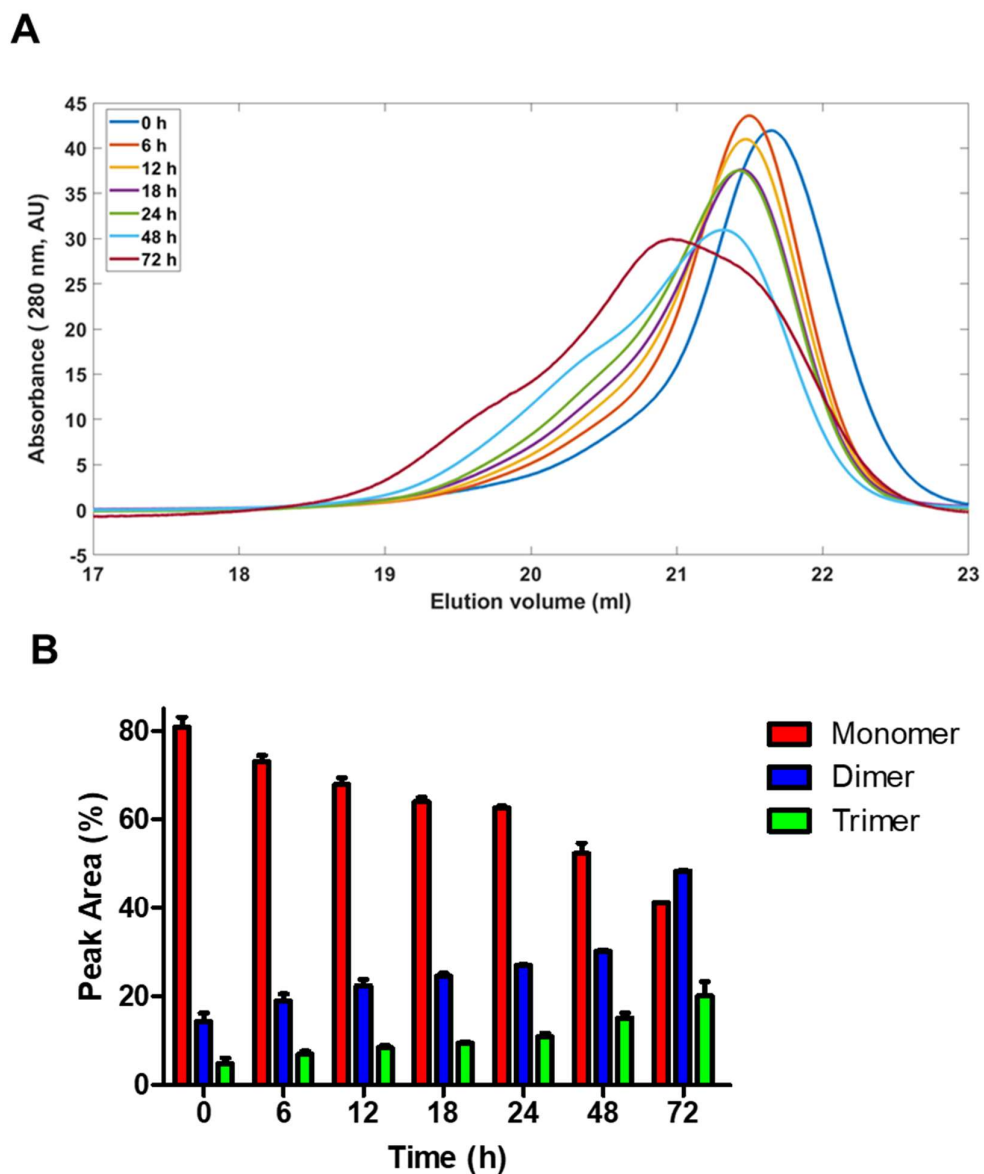


Figure 3.15: Quantification of HEWL peak components at different time points. (A) Area-normalized overlay of chromatograms obtained during the aggregation procedure. (B) Change in composition of the sample over time, as determined from the deconvolution of the chromatograms obtained ($n = 2$; mean $\pm$ SD).

3.2.4 Cytotoxicity and electrophysiological effects of the α -synuclein aggregates

Aggregates produced following the full oligomer formation protocol (Danzon *et al.*, 2007) were used to test cytotoxicity in primary cortical neuron cells. Dramatic cytotoxicity was observed down to α SN concentrations as low as 200 nM (**Figure 3.16**, *experiment performed by Dr Gavin Davis*). Cytotoxicity was assessed through an alamarBlue assay, which is a fluorescence-based assay that detects and quantifies cellular respiration through the fluorescence of the dye resazurin. As the sample applied to cells is shown to be a heterogeneous mixture of monomeric and oligomeric α SN, and there have been reports of monomeric α SN toxicity (Dettmer *et al.*, 2015; Wong & Krainc, 2017), meaning that α SN monomers may be contributing to the effect observed.

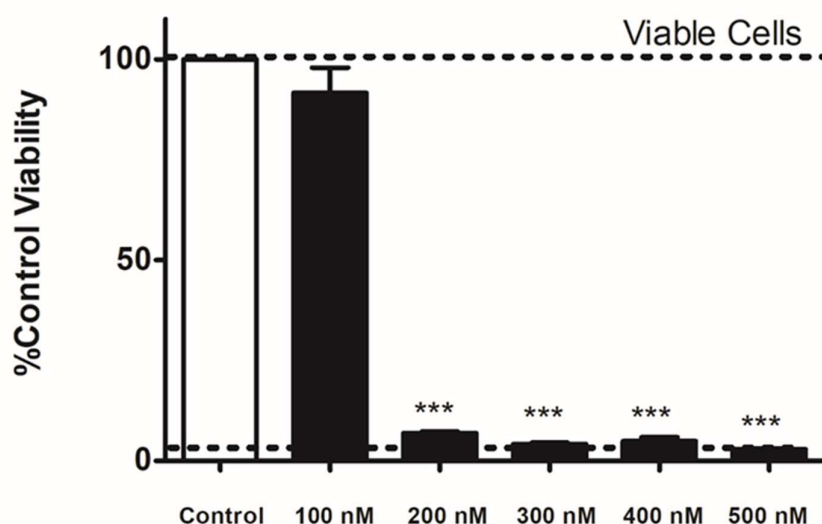


Figure 3.16: Cytotoxic effect of α SN oligomers on primary cortical neurons. Cytotoxicity of α SN oligomers on primary cortical neuron cultures *in vitro* after 24 h incubation ($n = 3$, mean $\pm$ SD). Oligomers exhibited significant cytotoxicity ($p < 0.0001$) down to concentrations as low as 200 nM. *Experiment performed by Dr Gavin Davis.*

Aggregates made using the same protocol were also injected *in vivo* in rats to obtain effects on baseline synaptic transmission (**Figure 3.17**, *experiment performed by Dr Igor Klyubin*), and were found to have significant electrophysiological effects down to concentrations as low as 135 pM. Aggregates were injected in the CA1 region rather than the nigrostriatal tract (primary region affected initially in PD) as the original experimental plan was to carry out full long-term potentiation (LTP) experiments (which proved

impossible due to the extreme effect observed on baseline synaptic transmission). LTP experiments are traditionally carried out in this brain region (Bliss & Lomo, 1973; Lomo, 2003; Walsh *et al*, 2002). Due to the nature of the experimental protocol, any non-soluble fibrillar/amorphous aggregates would be automatically excluded from the injection, meaning that the effect observed is due to a mixture of soluble monomeric, oligomeric and possibly proto-fibrillar species of α SN (not detected by AFM during imaging in liquid, but seen in TEM with dried sample).

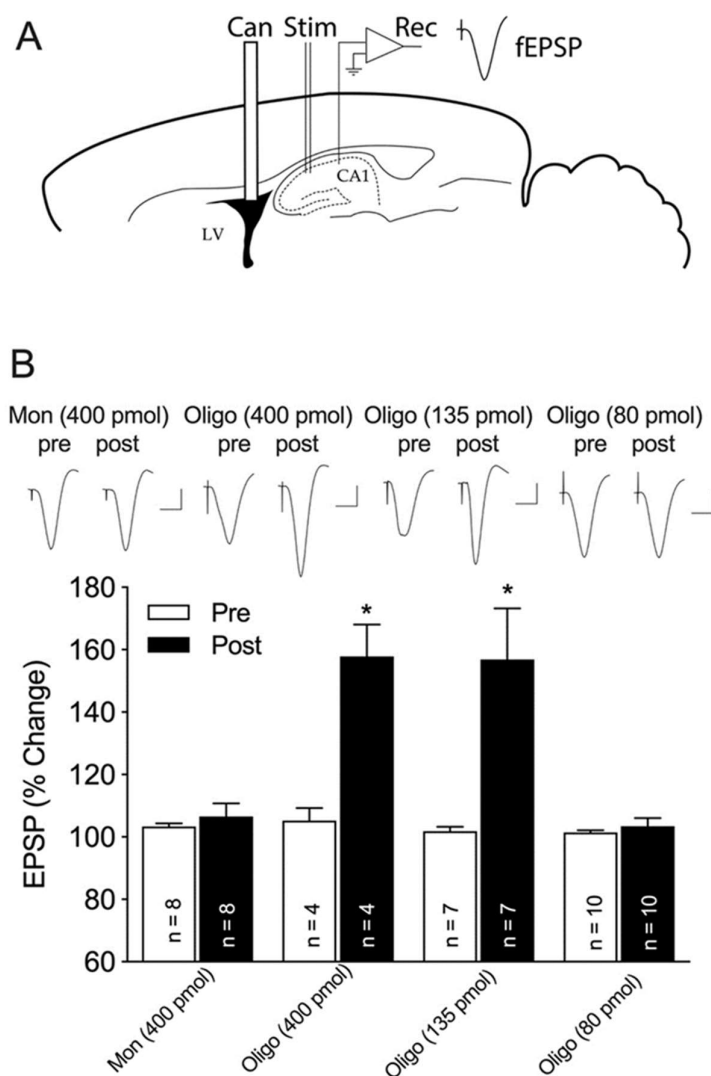


Figure 3.17: Effects of α SN oligomers on baseline synaptic transmission in rats. A schematic diagram of location of cannula (Can), stimulation (Stim) and recording (Rec) electrodes. Field excitatory postsynaptic potential (fEPSP) was recorded in the stratum radiatum of area CA1. The samples were injected into the lateral ventricle (LV). (c) Summary bar chart of changes in EPSP amplitude before (pre) and 90 min after (post) i.c.v. injection (5 μ l) of α SN

monomers and oligomers at indicated doses. Inserts show representative EPSP traces. Calibration bars: vertical, 1 mV; horizontal, 10 ms. * $P < 0.05$ (paired Student's t-test) compared with the same animals before injection. Values are the mean $\pm$ S.E.M. % pre-injection EPSP amplitude. *Experiment performed by Dr Igor Klyubin.*

3.2.5 Size and morphology of aggregating protein samples using atomic force microscopy

The size and morphology of the aggregates produced were also monitored using atomic force microscopy (AFM) and transmission electron microscopy (TEM). For α SN, most of the imaging was carried out on the samples used for cytotoxicity and animal experiments (**Figure 3.18**), which had undergone the full oligomerization protocol (Danzer *et al.*, 2007).

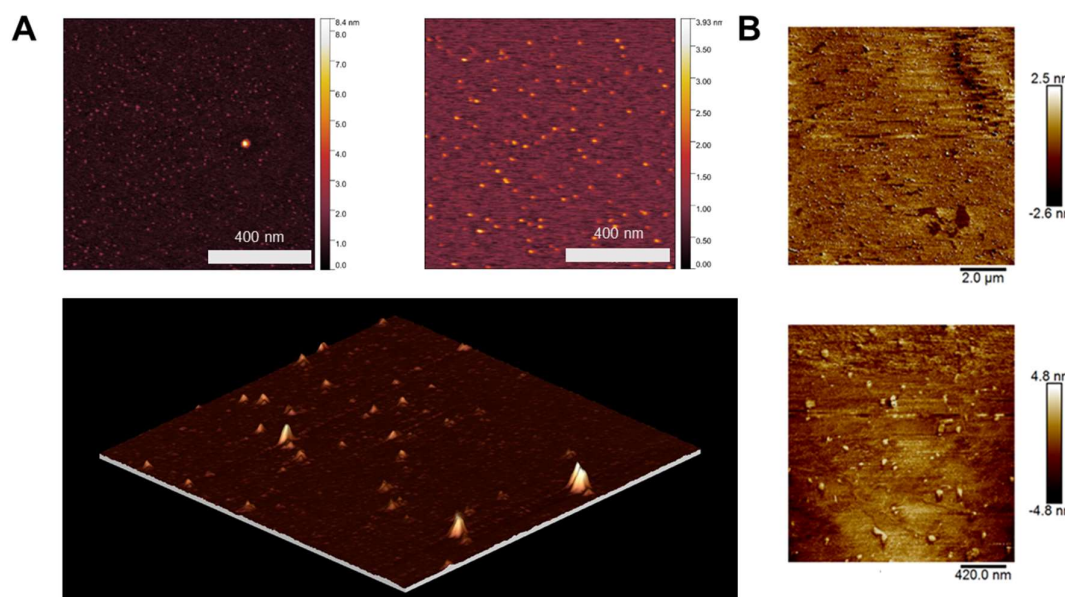


Figure 3.18: AFM imaging of α SN following the full oligomerization protocol. (A) AFM images of the α SN oligomers imaged in liquid: 2D and 3D representations (3D image scale = 500 x 500 nm). (B) AFM images of the α SN oligomers in air. Species imaged in air appear larger than when imaged in liquid, in part due to drying effects

A heterogeneous mixture of aggregates was observed, with samples imaged in air appearing significantly larger than when imaged in liquid due to further aggregation occurring when drying on the substrate. Samples imaged in liquid using AFM (**Figure 3.18, A**) exhibited smaller and more uniform particles, $R_{eq} = 40 - 150 \text{ \AA}$., while samples imaged using AFM in air (**Figure 3.18, B**) exhibited a less uniform distribution of larger particles of

$R_{eq} = 50 - 400 \text{ \AA}$. In the sample imaged in liquid a large proportion of monomer was also observed ($R_{eq} = 20 - 30 \text{ \AA}$). Quantification of the sample showed an overall heterogeneous mixture of aggregates of different sizes with a significant monomer proportion (**Figure 3.19**).

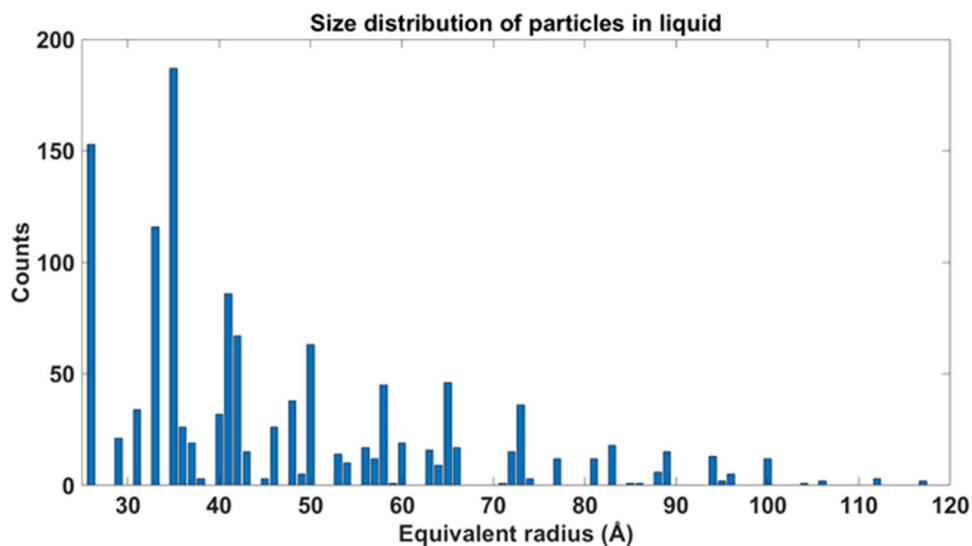


Figure 3.19: Size distribution of α SN aggregates imaged in liquid. Particle size distribution across samples ($n = 5$).

The sample was also imaged using TEM, through which proto-fibrils were also observed (**Figure 3.20, A**). TEM was also carried out on HPLC purified oligomer fractions, showing a heterogeneous mix of particles similar to what observed in the AFM images carried out in air (**Figure 3.20, B**).

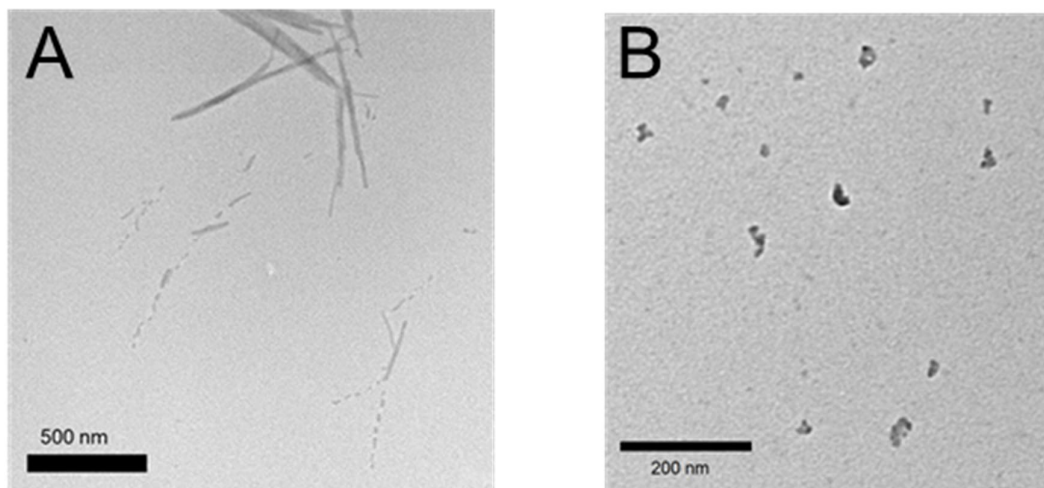


Figure 3.20: TEM imaging of α SN aggregates. (A) Small proto-fibrils observed through TEM imaging of sample at endpoint of oligomerization protocol. (B) TEM images of oligomer fraction purified using SE-HPLC. *TEM imaging carried out together with Dr Neal Leddy*

AFM measurements on the HEWL sample were carried out at similar time points as the SE-HPLC measurements; using AFM, samples appeared as heterogeneous mix of monomers and oligomers for the first 72 h (**Figure 3.21, A-C**), but after 144 h very large oligomers and proto-fibrils could be observed (**Figure 3.21, D**).

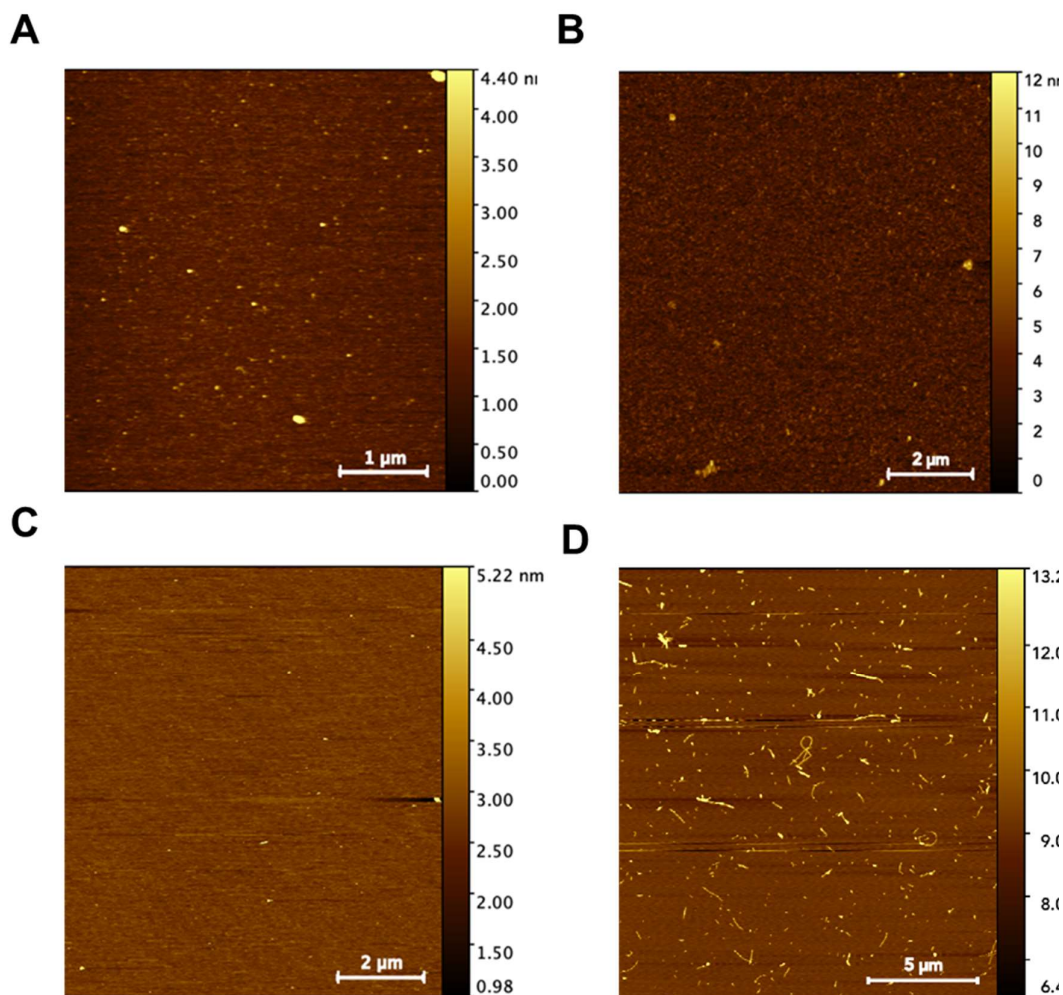


Figure 3.21: AFM imaging of HEWL aggregation at different time points. (A-D) AFM images of HEWL samples aggregating for (A) 24 h, (B) 48 h, (C) 72 h and (D) 144 h. The 144 h sample is the only one that is markedly different, with aggregates exhibiting clear fibrillary morphology. *Experiment carried out together with Dr Ji Yoon Kim.*

In general during the first three time points, protein species appeared roughly similar (20 – 200 Å), albeit in different proportions (**Figure 3.22**).

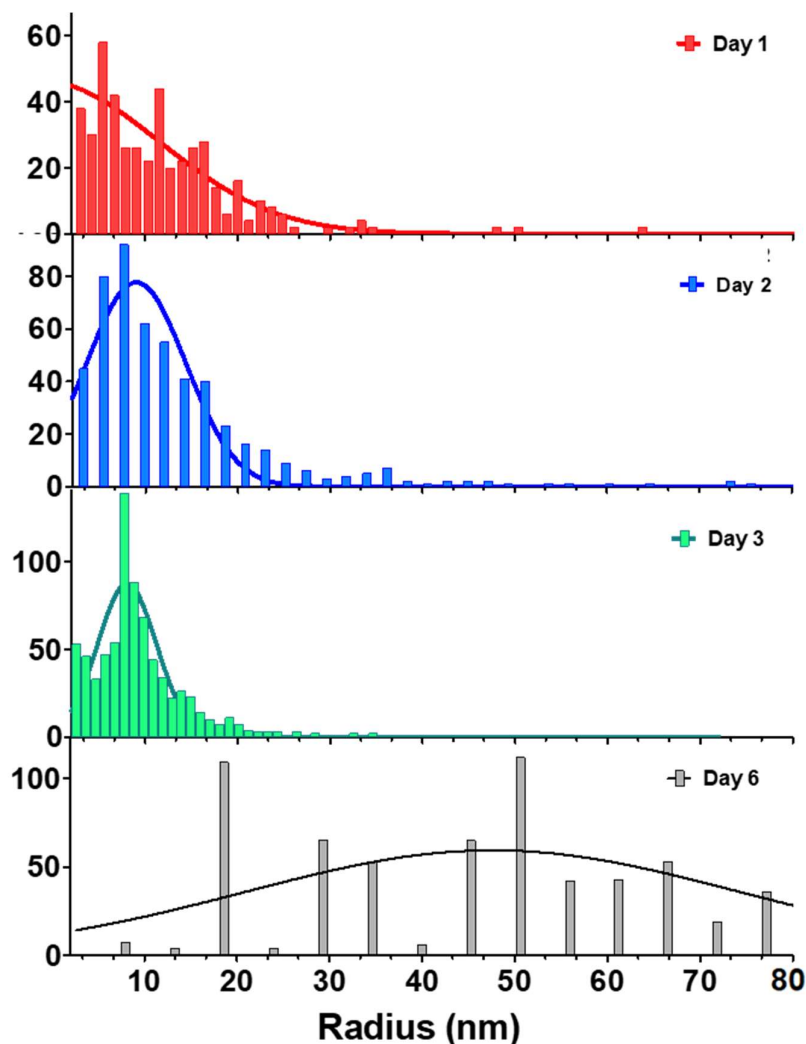


Figure 3.22: Aggregating HEWL sample composition at different time points.

Particle distribution across AFM images of aggregating HEWL samples at different incubation times ($n = 5$). Data analysis carried out together with Dr Ji Yoon Kim.

3.2.6 DOSY-based size estimations of aggregating samples

Size-estimation of protein aggregates was also attempted using DOSY-NMR spectroscopy. The signal intensity decay upon gradient strength change depends on the diffusion speed of the molecule (**Figure 3.23**), according to the Stejskal-Tanner equation (Stejskal & Tanner, 1965) $I = I_0 e^{-D\gamma^2 g^2 \delta^2 (\Delta - \delta/3)}$, where γ is the gyromagnetic ratio, Δ is diffusion time given, and δ is the time for the gradient pulse. By plotting signal intensity over gradient strength and fitting this to a gaussian curve ($y = Ae^{-Qx^2}$, where Q corresponds to the diffusion coefficient (D) multiplied by $\gamma^2 \delta^2 (\Delta - \delta/3)$) an estimate of the diffusion coefficient of a molecule can be obtained. The R_H can then be obtained by

rearranging the Stokes-Einstein equation $R_H = \frac{K_b T}{6\pi\eta D}$, where K_b is the Boltzmann constant, T is the absolute temperature and η is the dynamic viscosity of the solution.

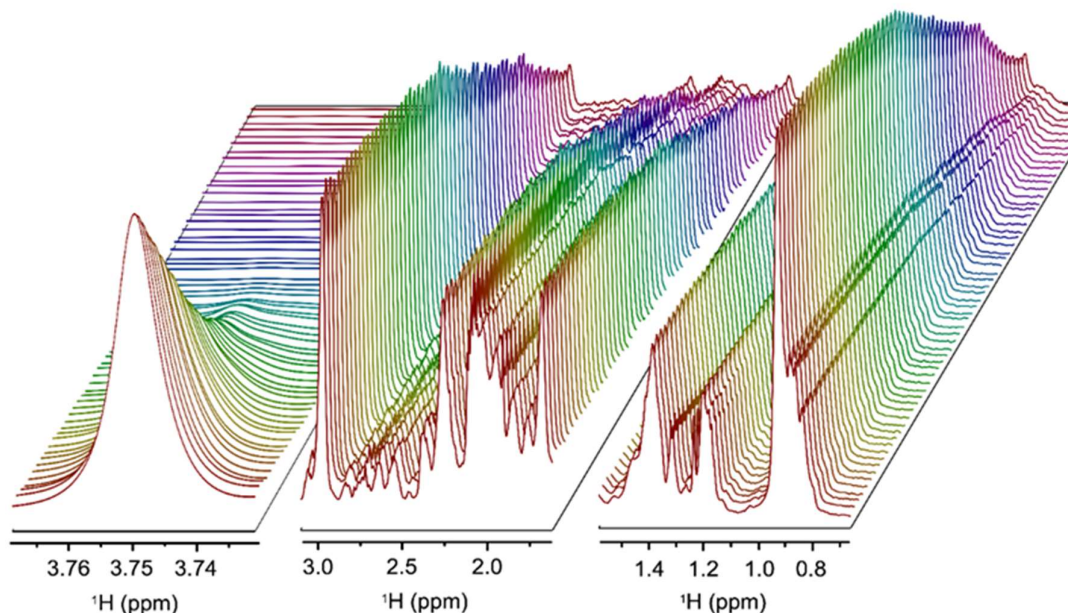


Figure 3.23: Diffusion projection of NMR spectrum from oligomerized α SN. Diffusion projections obtained from DOSY NMR spectroscopy analysis of aggregated α SN sample used for cytotoxicity and animal experiments stages (endpoint of aggregation protocol). The peak at 3.75 ppm is the reference standard dioxane, which size was used to calculate the diffusion coefficient of the aggregated protein species. *Image processed by Dr Ji Yoon Kim.*

For the oligomerized α SN sample, the 2D DOSY plot (**Figure 3.24, A**) obtained from the experiment showed that most peaks appeared at a diffusion coefficient value roughly consistent with literature values of its monomeric size (Araki *et al*, 2016; Chen *et al.*, 2015), 29.1 ± 2.2 Å (between 26.9 and 31.3 Å). Other, very weak peaks were also observed in the aromatic region, diffusing more slowly. Peak intensities from both the aliphatic (**Figures 3.24, B & E**) and aromatic (**Figure 3.24, C & F**) resonance frequency regions were used to calculate the diffusion coefficient.

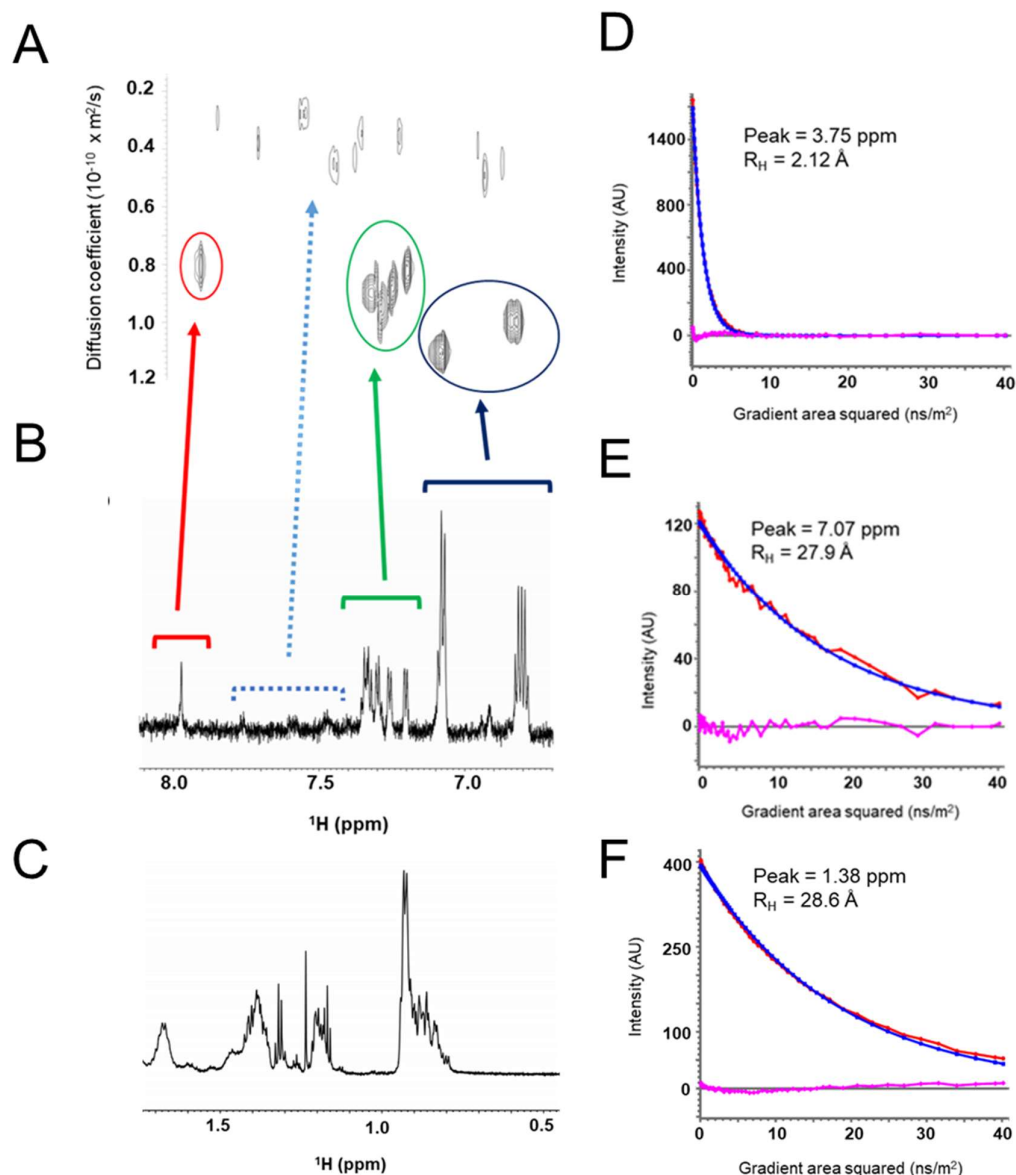


Figure 3.24: Oligomerized α SN DOSY NMR measurements.. (A) 2D DOSY plot of aromatic resonance frequency region of the aggregated α SN sample. The more prominent peaks correspond to aromatic residues histidine (red), phenylalanine (green) and tyrosine (blue). The less intense, slower diffusing peaks may be due to exchanging backbone amides. (B) Aromatic resonance frequency region of the 1D ^1H NMR spectrum obtained from the aggregated α SN sample. The less intense peaks observed in the 2D DOSY plot are also faintly visible. (C) Aliphatic resonance frequency region of the 1D ^1H NMR spectrum obtained from the aggregated α SN sample. (D) Single component DOSY-NMR signal decay fit for the dioxane peak at 3.75 ppm. The diffusion value obtained was used to calculate the apparent R_H of α SN following the proportion: $R_{Hp} = (D_d \times R_{Hd}) / D_p$. (E) Representative single component DOSY-NMR signal decay fit for the α SN peak in the aromatic resonance region at 7.07 ppm. The R_H obtained was 27.9 Å.

(F) Representative single component DOSY-NMR signal decay fit for the α SN peak in the aliphatic resonance frequency region at 1.38 ppm. The R_H obtained was 28.6 Å. This procedure was repeated on six different peaks and an average $R_H = 29.1 \pm 2.2$ Å was obtained. *Experiment carried out together with Dr Pancham Kandiyal & Dr Ji Yoon Kim*

In all cases, R_H values were determined based on 1,4-dioxane, the internal chemical shift reference and hydrodynamic radius calibration reference (3.75 ppm; $R_H = 2.12$ Å) (**Figure 3.24, D**) (Jones *et al*, 1997; Shimizu *et al.*, 1994). For HEWL, samples were run before aggregation was started (native condition) and after 72 h of the aggregation protocol. These were also applied to SE-HPLC analyses. From the single-component fit of the native aliphatic resonance region (**Figure 3.25, A**), an estimate for the diffusion was obtained, which was then set to its literature R_H . Any deviation from the native diffusion value observed in the aggregating sample was taken as due to aggregation (**Figure 3.25, B**).

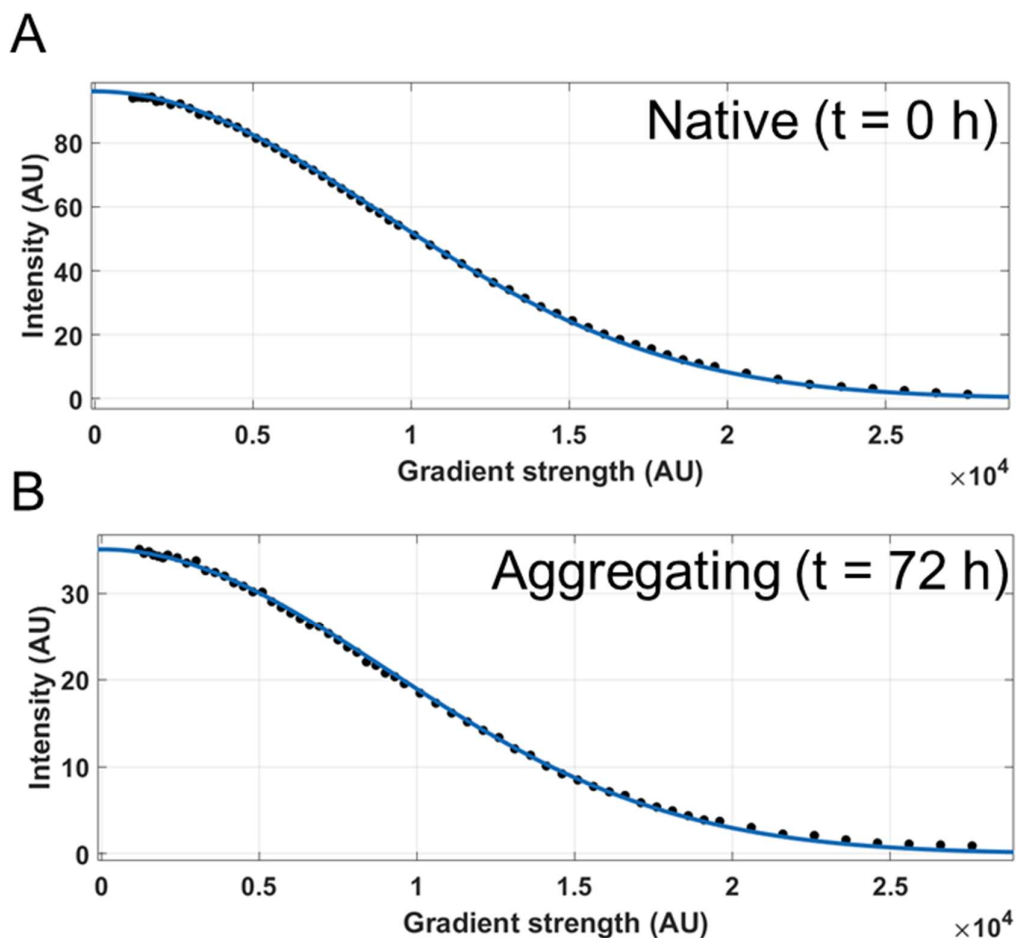


Figure 3.25: Single component fitting of diffusion projections from native/aggregating HEWL. (A) Representative single-component DOSY-NMR signal decay fit of the average aliphatic resonance frequency region of native hen egg white lysozyme. The average diffusion value obtained from a triplicate of three samples was set to the literature R_H of HEWL (Wilkins *et al.*, 1999), that is $R_H = 20.5 \pm 0.36 \text{ \AA}$, CI = 20.14 to 20.86 $\text{\AA}$ ($n = 9$, mean $\pm$ SD). (B) Representative single-component DOSY-NMR signal decay fit of the aliphatic resonance frequency region of HEWL aggregating for 72 h. The average diffusion value obtained from a triplicate of three samples was compared with the value obtained from the native sample experiments, and an $R_H = 21.48 \pm 0.29 \text{ \AA}$, CI = 21.19 to 21.73 $\text{\AA}$ was obtained ($n = 9$, mean $\pm$ SD).

As with the SE-HPLC analysis, differences were very subtle, the native exhibiting an R_H value between 20.14 and 20.86 $\text{\AA}$ ($n = 9$, mean $\pm$ SD) while from the sample aggregating for 72 h an R_H value between 21.19 and 21.73 $\text{\AA}$ was obtained ($n = 9$, mean $\pm$ SD) (**Figure 3.26**).

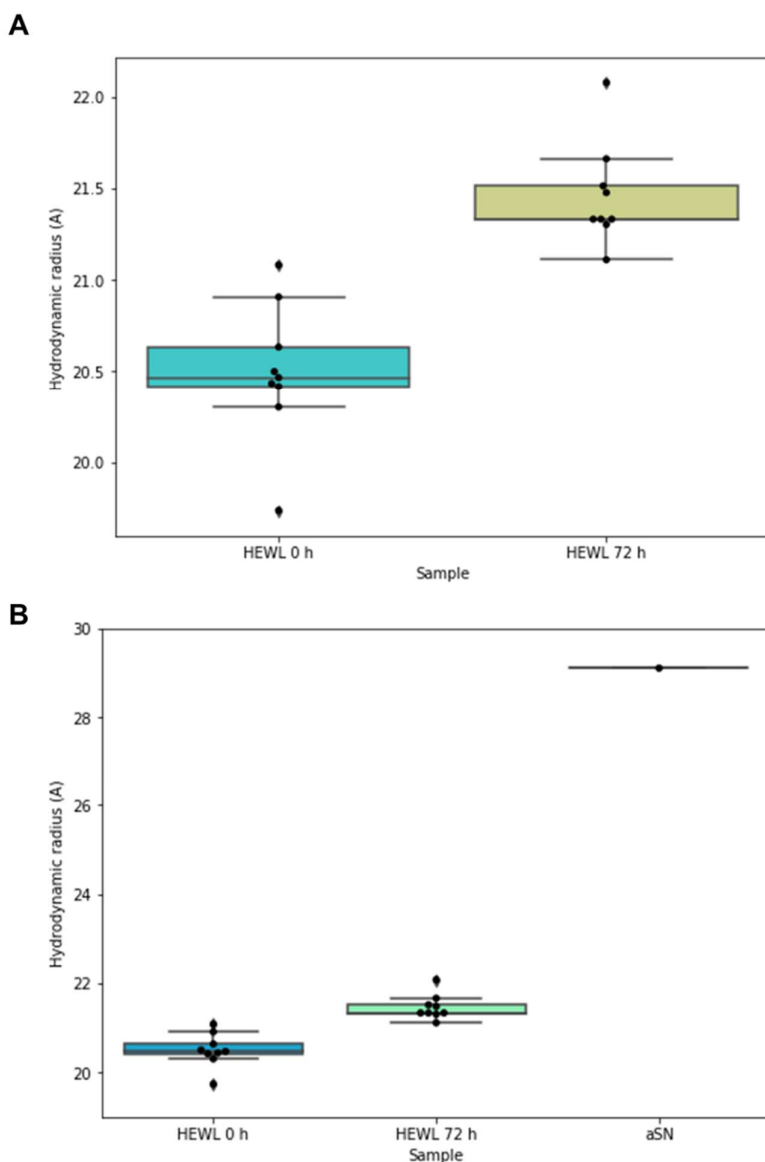


Figure 3.26: Differences between protein R_H values obtained during the DOSY experiments. (A) Boxplot exhibiting the difference in population averaged size between native and 72 h aggregated HEWL, as measured by DOSY NMR. Although small, difference in size is clearly observable. (B) Boxplot exhibiting the difference in population averaged size between native, 72 h aggregated HEWL and α SN, as measured by DOSY NMR. The α SN R_H obtained is much larger than both that of native and aggregating HEWL. A single DOSY experiment was run on the α SN sample due to sample quantity constraints (DOSY NMR requires a large amount of protein).

A further step was taken with the 72 h aggregating sample, following a procedure previously developed in this laboratory (Kandiyal *et al.*, 2019): by using the intensities

obtained from the SE-HPLC chromatogram deconvolution from the 72 h time point, and inputting the native (monomer) diffusion value, a three-component poly-gaussian fitting following the equation form $\sum i = A_1 e^{-Q_1 x^2} + A_2 e^{-Q_2 x^2} + \dots + A_i e^{-Q_i x^2}$ was carried out (Figure 3.27).

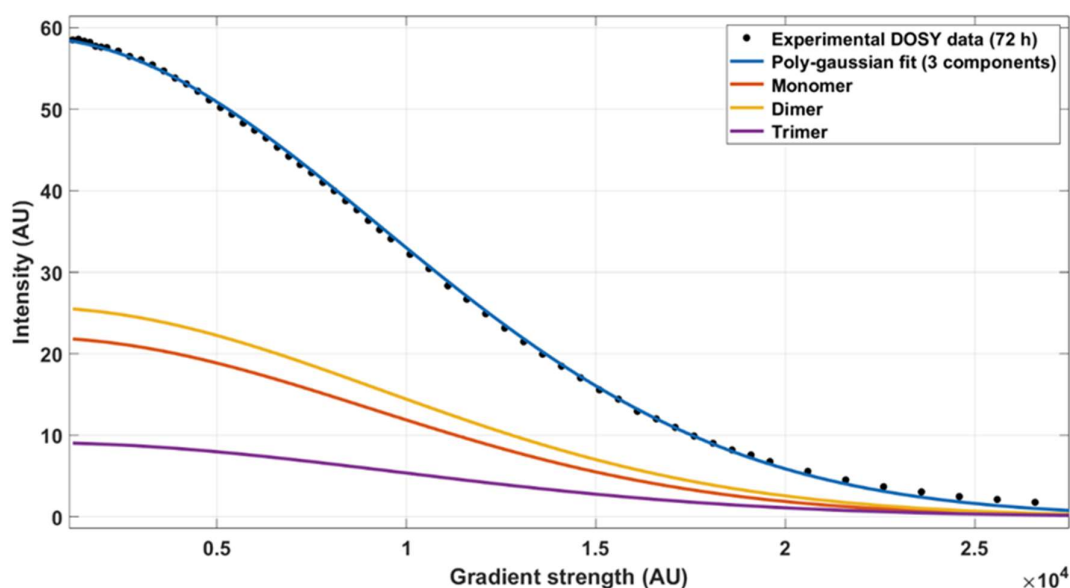


Figure 3.27: Three-component poly-gaussian fit of aggregating HEWL DOSY results.

Representative poly-gaussian fitting (three components) of the DOSY-NMR signal decay of the aliphatic resonance frequency region of HEWL aggregating for 72 h. The diffusion coefficient for the monomer was fixed (using the value obtained from the native HEWL DOSY-NMR experiments), as were the intensity values, obtained from the deconvolution of the SE-HPLC data (Monomer $R_H = 20.5 \text{ \AA}$, Dimer $R_H = 21.89 \text{ \AA}$, Trimer $R_H = 24.94 \text{ \AA}$).

This approach to DOSY data processing was previously validated for this pulse sequence on a sample containing a mixture of small molecules (Figure 3.28): as expected, the trend of the diffusion values was in close inverse correlation with their size (Figure 3.28, A). More importantly, this sample offered the opportunity to compare the diffusion coefficients obtained from single-component gaussian fittings of fully separated peaks against two-component gaussian fittings of overlapping peaks (Figure 3.28, B). The resulting poly-gaussian deconvolution agreed closely with single peak fittings (Table 3.2).

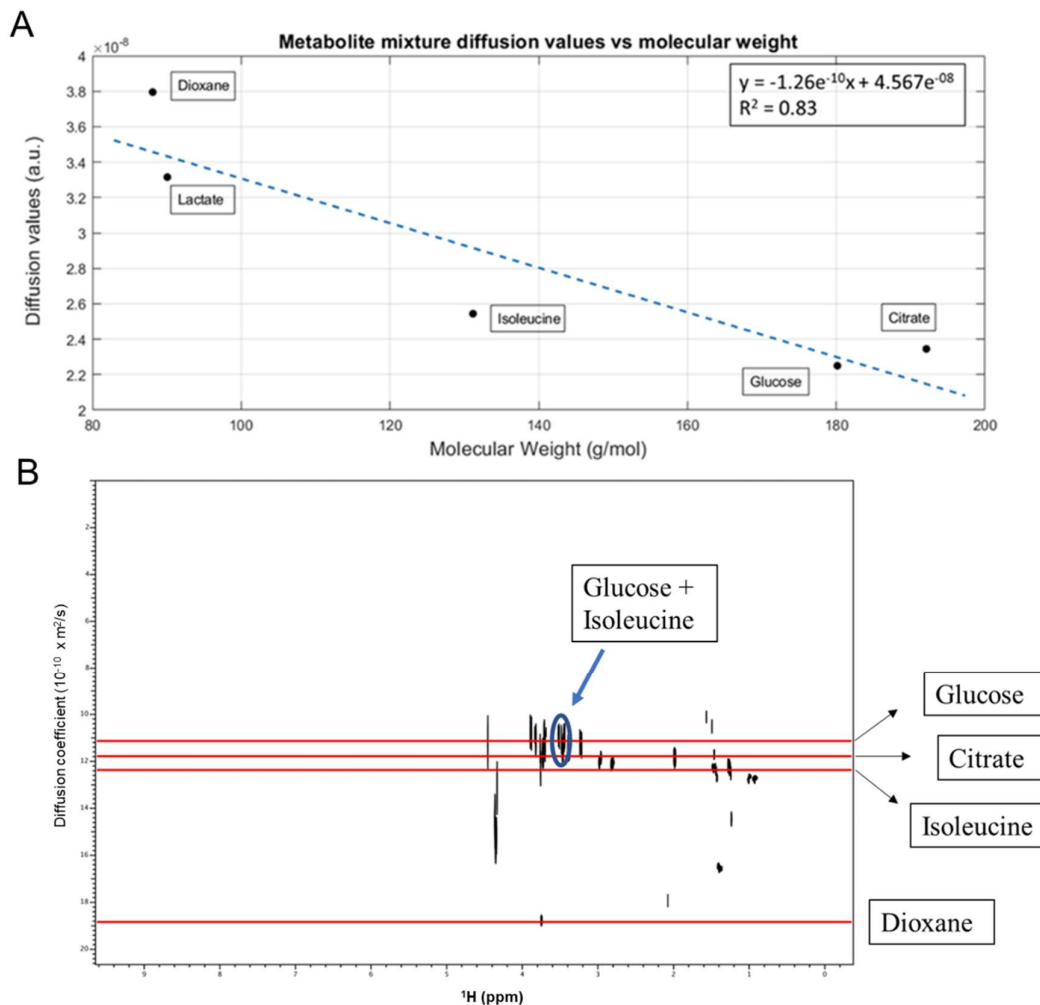


Figure 3.28: Results of DOSY NMR experiment on a mixture of small molecules. (A) Standard curve obtained from small molecule DOSY experiments. As expected, diffusion values are inversely correlated with molecule size. (B) Two dimensional DOSY plot obtained from a metabolite mixture. In the case of small molecules, the peak dispersion is sufficient to allow diffusion coefficients to be assigned unambiguously, allowing immediate verification of values obtained through poly-gaussian fitting of overlapping peaks. Diffusion curves were obtained by plotting intensity as a function of gradient strength following the Stejskal-Tanner equation ($I = I_0 e^{-D\gamma^2 g^2 \delta^2 (\Delta - \delta/3)}$). A two component fit ($I = I_1 e^{-D_1 \gamma_1^2 g^2 \delta^2 (\Delta - \delta/3)} + I_2 e^{-D_2 \gamma_2^2 g^2 \delta^2 (\Delta - \delta/3)}$) was used for the overlapping glucose and isoleucine peaks.

Table 3.2: Validation of HEWL deconvolution approach using a mixture of small molecules.

Compound	Molecular weight (g/mol)	Diffusion values (a.u.)	Estimated R_H (Å)
Dioxane	88.11	3.80E-08	2.12
Lactate	90.08	3.32E-08	2.42
Isoleucine	131.17	2.54E-08	3.15
Glucose	180.16	2.25E-08	3.57
Citrate	192.12	2.35E-08	3.42
Deconvolution from Glu + Ile peak (Glucose)		2.21E-08	3.63
Deconvolution from Glu + Ile Peak (Isoleucine)		2.39E-08	3.34

3.3 Discussion

A battery of different techniques was used to probe the composition and stability of the α SN aggregate mixture at early stages of aggregation. Despite each technique (SE-HPLC, AFM, DOSY-NMR) being based on different underlying equations and assumptions, the size estimates are in good agreement. Possibly the most striking result observed was the difference in SE-HPLC profiles between aggregating HEWL and aggregating α SN: the α SN samples exhibited consistent, baseline separated peaks (**Figure 3.11**) corresponding to large oligomers of sizes similar to those previously reported in the literature (Chen *et al.*, 2015), while injections of the HEWL samples showed a progressive increase in non-baseline separated peaks corresponding to species much closer in R_H to the HEWL monomer.

It is worth noting that in the present experiments the existence of two different oligomeric species was immediately and consistently revealed by the SE-HPLC measurements, without need for further separation procedures (*i.e.* analytical ultracentrifugation). Compared to techniques commonly used to monitor changes in sizes of aggregating proteins, SE-HPLC offers a variety of advantages (and some disadvantages):

Pros

- Relatively cheap and widely accessible.
- More sensitive than most spectroscopic/light-scattering techniques, allowing the use of protein aggregation protocols carried out at physiological concentrations, without the use of fluorescent dyes.
- Nearly all spectroscopic/light-scattering techniques, when used to estimate the size of a heterogeneous mixture of molecules, output a population-averaged value, requiring separate populations in the mixture to be estimated through mathematical deconvolution approaches. In comparison, SE-HPLC measures the composition of a heterogeneous mixture as individual populations (within the limits of the column capabilities).
- The time required to make a chromatographic measurement is relatively short, making it suitable for time series analysis.
- Potentially could be used to purify different-sized oligomeric species.

Cons

- Aggregate composition may be affected by injection in the HPLC system.
- There are limitations to the aggregate size that can be monitored using SE-HPLC, making it useful only for the initial aggregation steps.

- Injection of aggregating protein will affect column quality, making it necessary to clean and eventually substitute columns relatively often.

With the abundance of recent reports describing the pathophysiological relevance of soluble oligomer species of α SN (Aulić *et al.*, 2017; Chen *et al.*, 2015; Danzer *et al.*, 2007; Fusco *et al.*, 2017; Heras-Garvin *et al.*, 2018), as well as the cytotoxicity and electrophysiology measurements obtained presently (**Figures 3.16, 3.17**). These findings are interesting when compared with recent work suggesting fibrillar species denoted as sarkosyl-insoluble α SN might be most competent in seeding aggregation (Morgan *et al.*, 2020), as well as attempts at selecting aggregation inhibitors raised against fibril cores (Sangwan *et al.*, 2020). It is likely that the oligomers observed are *en-route* to fibril formation. The seeding pathway from unfolded protein to the final steps has recently been carefully dissected, showing the final steps of LB formation to be caused by a complicated interplay of cellular organelles and proteostatic pathways (Mahul-Mellier *et al.*, 2020). Given this information, the consistent observation of these oligomer peaks rather than a progressive increase in intermediate species (as seen with the HEWL sample) points towards a kinetic aggregation plateau directly relevant to disease progression which occurs before the fibril stage and is already a driver of pathology.

From the AFM and TEM results obtained, we can observe that the Danzer oligomerization protocol (Danzer *et al.*, 2007) leads to a heterogeneous aggregate population (**Figures 3.18, 3.19**) containing a large monomer population, as well as some sporadic proto-fibrils (**Figure 3.21**). This suggests that these aggregates are most likely *en route* towards amyloid fibrils (the aggregate forms of α SN directly observed in Lewy bodies), and not off-pathway oligomers. For the monomeric α SN, the average R_H value obtained from the DOSY-NMR confirms its expanded nature when compared to proteins of similar M_R such as HEWL, corroborating literature values of α SN size obtained using other techniques. The diffusion coefficient obtained for the monomeric α SN ($1.18 \pm 0.08 \times 10^{-6} \text{ cm}^2/\text{s}$; **Figure 3.24**) is markedly different from the values obtained from the tetrameric form ($3.07 \pm 0.06 \times 10^{-5} \text{ cm}^2/\text{s}$) (Wang *et al.*, 2011), contradicting the notion that α SN exists natively as a tetramer in cells. The resulting $R_H = 29.1 \pm 2.2 \text{ Å}$ is consistent with a small angle X-ray scattering (SAXS) study estimates of an R_G of $27.2 \pm 4.4 \text{ Å}$ (Araki *et al.*, 2016), which corresponds to a R_H of $35.1 \pm 5.7 \text{ Å}$ (using the relationship $R_H = R_G/0.775$ (Kohn *et al.*, 2004)). During the DOSY-NMR of the α SN sample, larger aggregate species were not observed, most likely due to the large τ_c values associated with the $> 100 \text{ Å}$ species found in the AFM and SE-HPLC results.

Regarding the HEWL oligomers, we were able to deconvolute the SE-HPLC chromatograms into three components (monomer, dimer, and trimer) and estimate their hydrodynamic volumes. HEWL dimers and trimers are known to exist and considered relevant to the early steps of HEWL aggregation (Sharma *et al.*, 2016). The sizes and proportions of these three components are echoed in the DOSY-NMR results, which enabled us to distinguish native from aggregating protein (**Figure 3.25**). A general method to identify an arbitrary distribution of signal intensity as a function of diffusion coefficients using inverse Laplace transforms is not currently available. As a result, while separating components in a 2D DOSY plot where resonance frequencies are moderately overlapping is possible (**Figure 3.28**), it is relatively challenging for extreme levels of overlap (termed “low-resolution DOSY” cases) such as those found in an aggregating mixture of a pure protein (Morris, 2009). In contrast, using independently obtained data as constraints (in this case the SE-HPLC measurements), and fitting the data as multi-component gaussians, DOSY NMR can provide very strong corroborating evidence, as can be observed from the excellent fit of the diffusion data obtained for the 72 h HEWL aggregate sample with the proportions of species observed during SE-HPLC (**Figure 3.27**). The radical differences observed during DOSY-NMR between the α SN and the HEWL samples suggest completely different aggregation pathways for the two proteins: HEWL going through small, stable intermediates (dimers – trimers) detectable using this technique, while for α SN the first semi-stable species seem to be much larger (as shown by SE-HPLC, putatively ~20-70 mers) and consequently not detectable by DOSY-NMR due to large τ_c values. A possible explanation for this behavior, consistent with other recent observations (Brodin *et al.*, 2022; Huang *et al.*, 2022; Ray *et al.*, 2020) is that α SN aggregation occurs, even *in vitro*, by transitioning from an LLPS state (where it would still be seen as monomeric by NMR), and very rapidly jumps into semi-stable, soluble oligomeric intermediates, that are still in a state of dynamic equilibrium with the LLPS separated monomer. This is also consistent with the SE-HPLC results showing the reversibility of α SN aggregation using this protocol (**Figure 3.12**), suggesting a sequence of events in which: (i) α SN LLPS may occur in a localized fashion at synaptic terminals, possibly mediating synaptic vesicle re-uptake (Brodin *et al.*, 2022; Burré, 2015; Man *et al.*, 2021; Milovanovic & De Camilli, 2017) (ii) in the presence of pathogenic mutations and/or in neurons with very high levels of synaptic activity, LLPS seeds the formation of soluble α SN oligomers which (iii) can then spread to through the synaptic cleft, either by punching holes through the lipid bilayer or through a receptor mediated mechanism (Aulić *et al.*, 2017; Fusco *et al.*, 2017; Rösener *et al.*, 2020; Urrea *et al.*, 2017).

4. DIVERSE ROLES OF STRUCTURED & DISORDERED REGIONS OF ATAXIN-2 IN mRNP GRANULE FORMATION

4.1 Introduction

4.1.1 Spinocerebellar ataxias

Spinocerebellar ataxias (SCAs) are a group of rare, genetically heterogeneous, autosomal dominant inherited progressive neurodegenerative disorders. They are mostly adult onset and their most prominent clinical hallmarks are loss of balance (ataxia literally means “lacking coordination”) and slurred speech (Klockgether *et al*, 2019). There are currently no effective treatments for any of the SCAs, and all generally lead to a premature death (Schöls *et al*, 2004). In each SCA form, the causative mutations affect a different protein, and can be roughly subdivided in (i) repeat expansion mutations, characterized by an increased number of amino acid repeats in certain protein domains (usually poly-glutamine repeats, also denoted as CAG repeats) and (ii) non-repeat expansion mutations (specific point mutations in the affected proteins) (Durr, 2010). Regardless of this difference in protein targets, which unsurprisingly causes a wide degree of heterogeneity of clinical features across SCAs, in most SCAs the affected areas are mainly the cerebellum (in particular the Purkinje neurons) as well as the spinal cord (although the extent of damage varies and other parts of the nervous system can be affected across SCAs). Molecular determinants of disease are also widely shared across SCAs, and include ion channel dysfunction, proteotoxicity and RNA toxicity (**Figure 4.1**) (Durr, 2010; Schöls *et al.*, 2004). In the case of repeat expansion SCAs, severity of symptoms is correlated to length of the amino acid repeats; these repeats are unstable, and tend to increase in size from one generation to the next (Klockgether *et al.*, 2019).

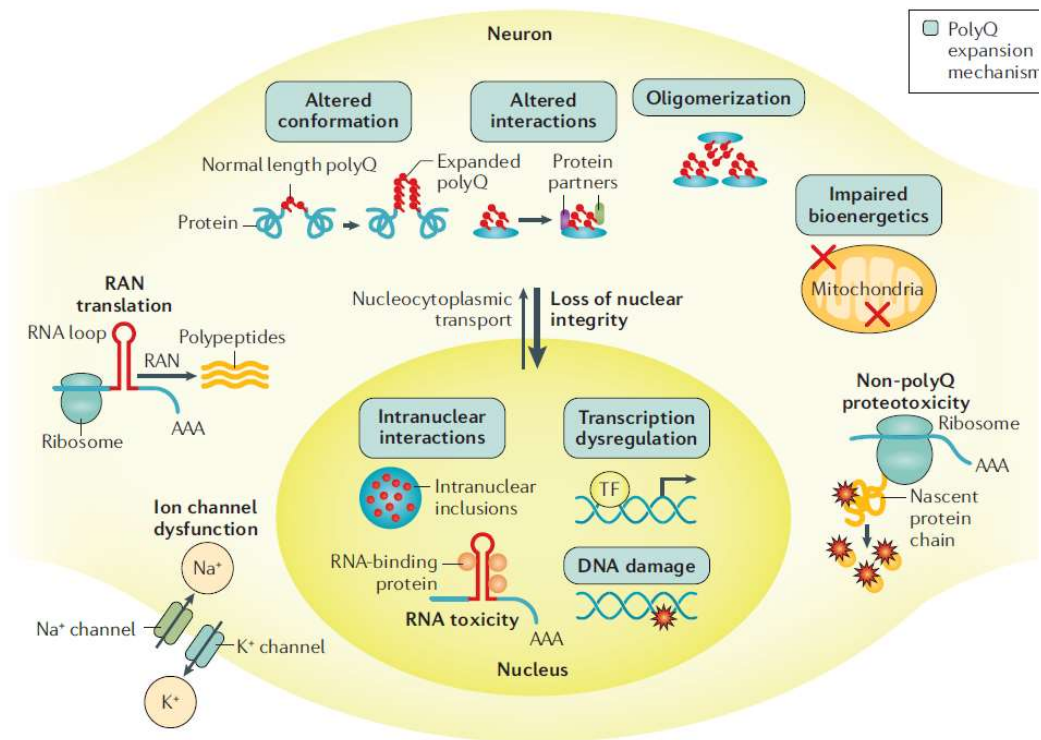


Figure 4.1: Molecular determinants of pathology observed through different types of spinocerebellar ataxias. Repeat expansions (for example poly-glutamine expansions) as well as certain point mutations, can cause protein misfolding that can cause the affected protein to (i) lose its ability to perform its native function, (ii) increase its aggregation propensity. Protein aggregates can then damage the cell in a variety of ways, for example by impairing mitochondrial or ion channel function, or by causing RNA toxicity if associated in protein-RNA condensates. Misfolded nuclear proteins may also cause DNA damage and dysregulation of nucleocytoplasmic transport (*Adapted from Klockgether et al., 2019*).

In the poly-glutamine (poly-Q) repeat expansion SCAs (SCA1, SCA2, SCA3, SCA6, SCA7 and SCA17) the poly-Q stretches are thought to directly drive pathogenesis. These expansions are known to drive protein aggregation, forming intracellular inclusions in affected cells (Paulson *et al*, 2017). The precise pathological role of the inclusions themselves is debatable: it is still unclear if these are directly responsible for cytotoxicity or rather they are the result of a cellular response aimed to process/clear misfolded protein (there is some evidence of a protective effect in mouse and cellular models (Klockgether *et al.*, 2019)). In any case, protein aggregation disrupts cellular trafficking, and may trap protein and mRNA binding partners into these inclusions, perturbing normal cellular

function and homeostasis. Furthermore, the continuous production, management and degradation of non-functional misfolded protein may overload the cells energy production and proteostatic machinery, a burden especially on cells with endogenously high energy demands (as would be Purkinje neurons) (Klockgether *et al.*, 2019). In some forms of SCAs, the function of the pathogenic poly-Q expanded protein has been well elucidated and some interesting implications for potential treatments have emerged. In the SCA1 manifestation for example, the protein Ataxin-1 (Atxn1) acts a transcriptional co-factor, which can bind different partner proteins, such as the transcriptional repressor Capicua or the spliceosome component Retinoid binding protein 7 (RBP7) (Rousseaux *et al.*, 2018). The poly-Q expansion makes the interaction with Capicua more favourable, disrupting gene expression and splicing, which is thought to drive cytotoxicity. Interestingly, toxicity of the poly-Q expanded mutant Atxn1 was found to also depend on the phosphorylation of a serine residue situated far from the poly-Q expanded region of the protein (S776). Deletion of S776 inhibits toxicity of mutant Atxn1, and pharmacological intervention in mouse models aimed at reducing phosphorylation of S776 was also found to have protective effects (Duvick *et al.*, 2010; Pérez Ortiz *et al.*, 2018; Rousseaux *et al.*, 2018).

This example highlights that while the poly-Q expansions are essential disease drivers, the specific function of the protein, as well as the structure of its other domains, are important to understand mechanisms of pathogenicity. This is especially relevant in the context of drug development, as potential drug targets might be hidden in regions far away from the disease-causing-expansion. The current work aspires to take steps towards this direction, studying Atxn2, the protein underlying SCA2 pathogenicity.

4.1.2 Ataxin-2 and neurodegeneration

Similarly to SCA1, SCA2 is also an autosomal dominant glutamine-repeat expansion SCA, involving the gene *ATXN2* (Klockgether *et al.*, 2019). This gene encodes the protein Ataxin-2 (Atxn2), which contains a poly-Q repeat near its N-terminal region. In SCA2 patients, this repeat (which normally is 22-24 amino acids long) is expanded to over 34 glutamines. Intermediate poly-Q expansions of Atxn2 (27-33 glutamines) are instead correlated with a higher risk of a variety of neurodegenerative conditions, most importantly ALS, which is the most common kind of motor neuron disease in adults (Elden *et al.*, 2010). In SCA2, the poly-Q expansion is thought to cause a toxic gain-of-function in Atxn2, possibly by increasing its aggregation propensity, although the exact cytotoxic mechanisms of mutant Atxn2 in SCA2 have not been fully elucidated (Ostrowski *et al.*, 2017). Atxn2 is highly expressed in cerebellar Purkinje neurons, metabolically active cells with a long and highly branched axonal arbor (Laffita-Mesa *et al.*, 2021).

In ALS pathology, motor neurons progressively die, which causes a variety of motor symptoms (such as muscle weakness and loss of coordination), eventually leading to paralysis, inability to breathe, and death (Kim *et al*, 2020). A variety of mutations in different genes have been shown to drive this process. The genes encoding TAR DNA-binding protein 43 (TDP-43) and fused-in-Sarcoma (FUS) are examples of genes implicated in ALS pathology, and are involved in the formation of inclusion bodies in motor neurons (Neumann *et al.*, 2006; Patel *et al*, 2015; Vance *et al.*, 2009). Poly-Q expanded Atxn2 has been shown to interact with both TDP-43 and FUS, contributing to their aggregation and cytoplasmic translocation (Elden *et al.*, 2010). Antisense oligonucleotide therapies against Atxn2 in mutated TDP-43 mouse ALS models have been shown to have a beneficial effect on pathology (Becker *et al*, 2017). Depletion of Atxn2 has also been shown to decrease cytotoxic effects in a *Drosophila* model of C9ORF72 driven ALS (Bakthavachalu *et al.*, 2018). In this case it is unclear if the driver of disease consists in repetitive RNAs or translated dipeptide repeats of this; whichever the case, Atxn2 appears to interact and possibly stabilize the toxic product of the C9ORF72 gene (van Blitterswijk *et al*, 2014).

Structurally, Atxn2 is a large (1313 amino acids long) protein, with three well-annotated structured domains (LSm, LSm-AD and PAM2 domains) interspersed amongst long intrinsically disordered regions (**Figure 4.2, A**). Atxn2 also contains a large number of potential phosphorylation sites, as well as some putative acetylation and ubiquitination sites, in particular around its structured domains (PAM2 in particular), suggesting that post-translational modifications of the protein might be involved in regulating its function – most likely by changing the accessibility to its well structured domains. Its structure is highly conserved throughout most eukaryotes (Ostrowski *et al.*, 2017). Excluding birds, all vertebrates contain two homologues of the protein, denoted Ataxin-2 and Ataxin-2-like (Atxn2L). Atxn2L does not contain the poly-Q regions. In *Drosophila* (and putatively in all organisms without this duplication) deletion of Ataxin-2 (Atx2), and indeed of any of its structured domains, is lethal (**Figure 4.2, B**) (Bakthavachalu *et al.*, 2018).

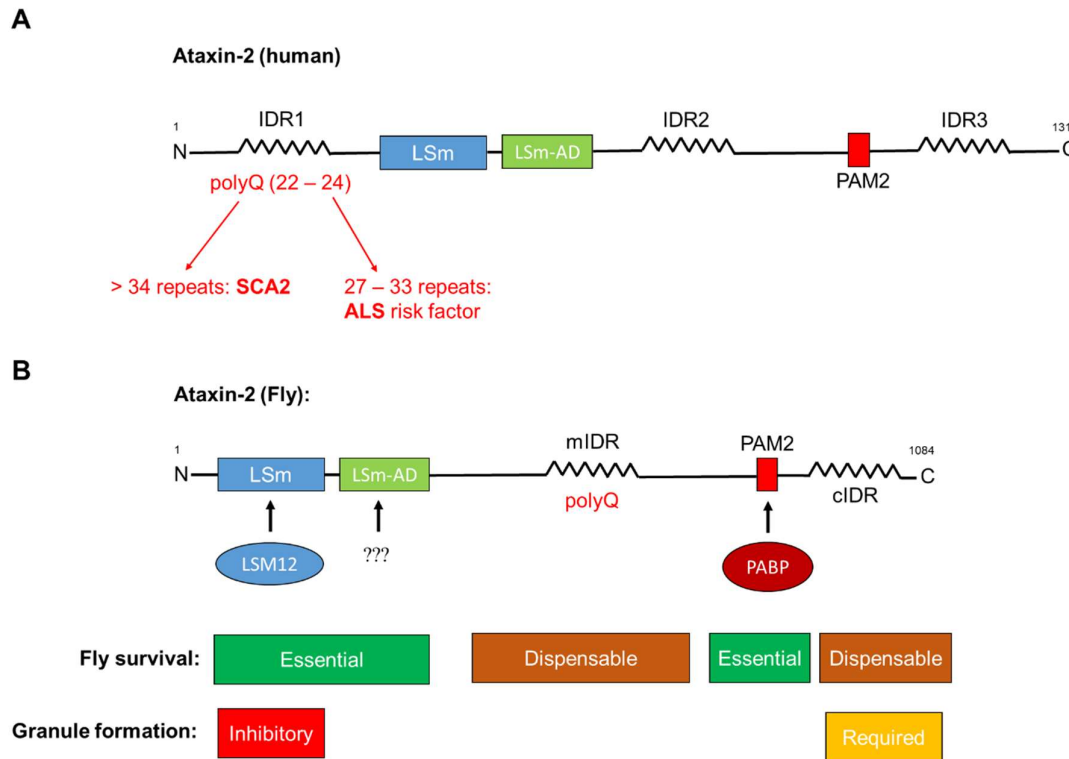


Figure 4.2: Domain structure schematic of Ataxin-2, human and fly. (A) Domain structure of human Atxn2. The poly-Q repeats are found in the first IDR region at the N-terminal region; if these are expanded to > 34 glutamine repeats, a SCA2 phenotype is observed, while medium expansion (27 – 33 glutamine repeats), leads to a higher risk of developing ALS or other neurodegenerative pathologies. (B) Domain structure of fly Atx2. Human and fly protein are highly conserved but fly Atx2 is slightly shorter and lacks the first IDR domain. Poly-Q repeats can be found in the middle IDR (mIDR) of the protein. Experiments with *Drosophila* have demonstrated that the structured domains are essential for fly survival, while the intrinsically disordered regions are not. *In vitro* and *in vivo*, overexpression of fly Atx2 is enough to cause the formation of Atx2 condensates. Deletion of the Atx2 C-terminal IDR (cIDR) blocks fly Atx2 ability to form condensates upon over-expression.

Although the precise cellular & molecular functions of Atxn2 have yet to be fully elucidated, it has been shown to be essential for fly survival (Bakthavachalu *et al.*, 2018). It is well accepted that it acts as an RNA binding protein, and has been shown to associate with a variety of different-RNA-binding-protein-partners, suggesting it may be playing different roles depending on the proteins it associates with (Lee *et al.*, 2017; Lim & Allada, 2013; Satterfield & Pallanck, 2006). At a cellular level, Atxn2 is an important component of the stress response pathway through its recruitment to SGs, cytoplasmic protein-RNA condensates observed upon subjecting cells to stress (Bakthavachalu *et al.*, 2018; Nonhoff

et al, 2007). At a molecular level, experiments depleting Atxn2 suggest it has a stabilizing effect on its target mRNAs, but it also appears to have a role in miRNA silencing, so its specific effect most likely depends on the nature of the target mRNA and/or on the protein complex that is formed (**Figure 4.3**). Atxn2-target mRNAs have specific roles in multiple neuronal functions, suggesting that in neurons translational regulation by Atxn2 can have downstream effects on a variety of complex functions of the organism, like circadian rhythm regulation and memory (Lee *et al*, 2018; Sudhakaran *et al.*, 2014).

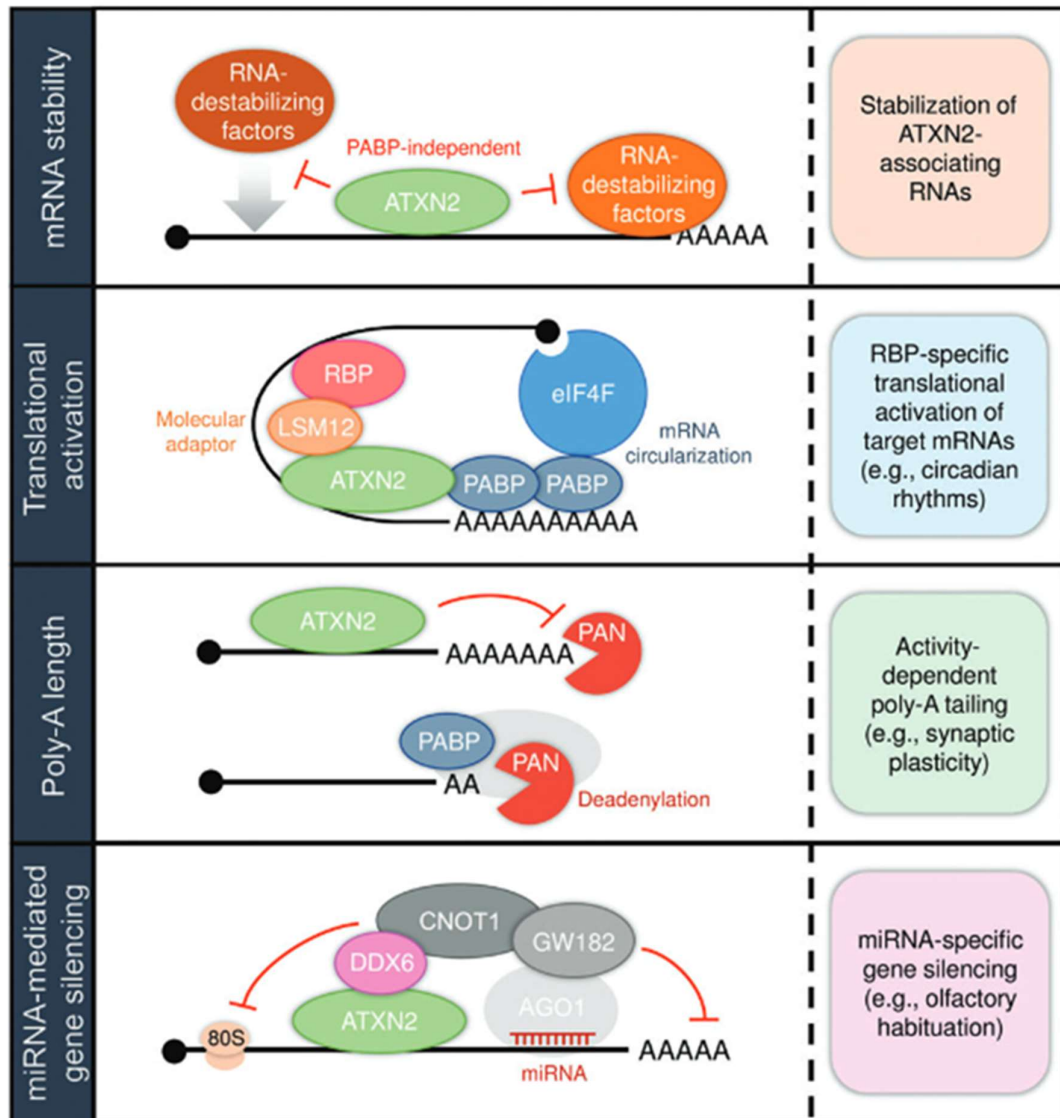


Figure 4.3: Atxn2 can have different functions depending on the mRNA transcript/protein complex it is associated with. While overall the association of Atxn2 to its mRNA targets seems to have a stabilizing effect, the identity of the mRNA transcript and/or the rest of the associating protein complex determine the specific molecular effect associated with Atxn2. Examples in

which Atxn2 acts as a translational activator, as an activity dependent regulator of poly-A tail length and as a participant in miRNA gene silencing are depicted. (*Adapted from: Lee et al, 2018*)

4.1.3 Atxn2 domains and their functions

The structured domains of Atxn2 are the LSm, the LSm-AD and the PAM2 domains. These structured domains are highly conserved across organisms, while the intrinsically disordered regions are less conserved. The LSm and the PAM2 domains are found in many other proteins and their structure and function has been characterized in detail (Xie *et al*, 2014; Zhou *et al*, 2014). However, little is known about the LSm-AD domain, other than the fact that it exhibits significant homology with the anticodon-binding domain of the prolyl tRNA synthetase protein (Albrecht & Lengauer, 2004; Yaremchuk *et al*, 2001). Due to this, the LSm-AD domain is thought to be involved in RNA binding and specific nucleotide recognition.

The IDRs. The long intrinsically disordered regions (IDRs) are thought to play a mostly non-sequence-specific physical role, allowing Atxn2 to form phase-separated condensates (Bakthavachalu *et al.*, 2018). Some recent findings suggest though that the IDRs may be more complex than previously thought, and play more specific roles in Atxn2's native function (Boeynaems *et al.*, 2021). As shown above (**Figure 2.4, B**), previous work from the lab (Bakthavachalu *et al.*, 2018) demonstrated that both IDRs are dispensable for fly survival (while a full Atx2 KO is lethal), but that the cIDR domain is instead indispensable for fly cells to form SGs. The fact that its deletion is not embryonically lethal suggests that the ability of Atx2 to drive granule formation is – at least to a certain extent – uncoupled from the resto of its cellular functions.

The LSm domain. Structurally, the LSm domain has a highly conserved characteristic cross β -sheet fold (**Figure 4.4, A**). Most Sm or Sm-like proteins are essential RNA binding proteins, found in all domains of life. The majority of known LSM proteins are small and assemble into ring-like hetero-heptameric/hexameric structures. These rings perform RNA processing tasks such as splicing in the nucleus (Sm ring, LSM2-8 ring) or mRNA degradation in the cytosol (LSM1-7 ring). Crystal structures of these rings have been solved, showing the canonical way LSm proteins interact with each other, as well as how they bind RNA (Zhou *et al.*, 2014) (**Figure 4.4, B**). LSm:LSm interactions occur by interfacing the β 4b strand of one unit with the β 5 strand of the next. Amino acids exposed on the inside of the LSm rings grant specificity to the protein-RNA interaction. Subsequently, a variety of other LSm domain-containing proteins was discovered, including larger proteins containing the conserved LSm domain but also other domains (Albrecht & Lengauer, 2004) (**Figure 4.4, C**). These proteins are still presumably capable

of binding to RNA/nucleotides and have been shown to interact with a variety of other proteins. Atxn2 falls in the latter category. The Atxn2 LSm domain has been predicted and subsequently demonstrated to bind another LSM protein from the latter group, LSM12, forming a complex that induces transcriptional activation of specific mRNA targets (Albrecht & Lengauer, 2004; Lee *et al.*, 2017). Interestingly, LSM12 was subsequently identified as a Nicotinic acid adenine dinucleotide phosphate (NAADP) receptor, activating two pore channels (TPCs) and mobilizing Ca^{2+} stores from lysosomes (Zhang *et al.*, 2021).

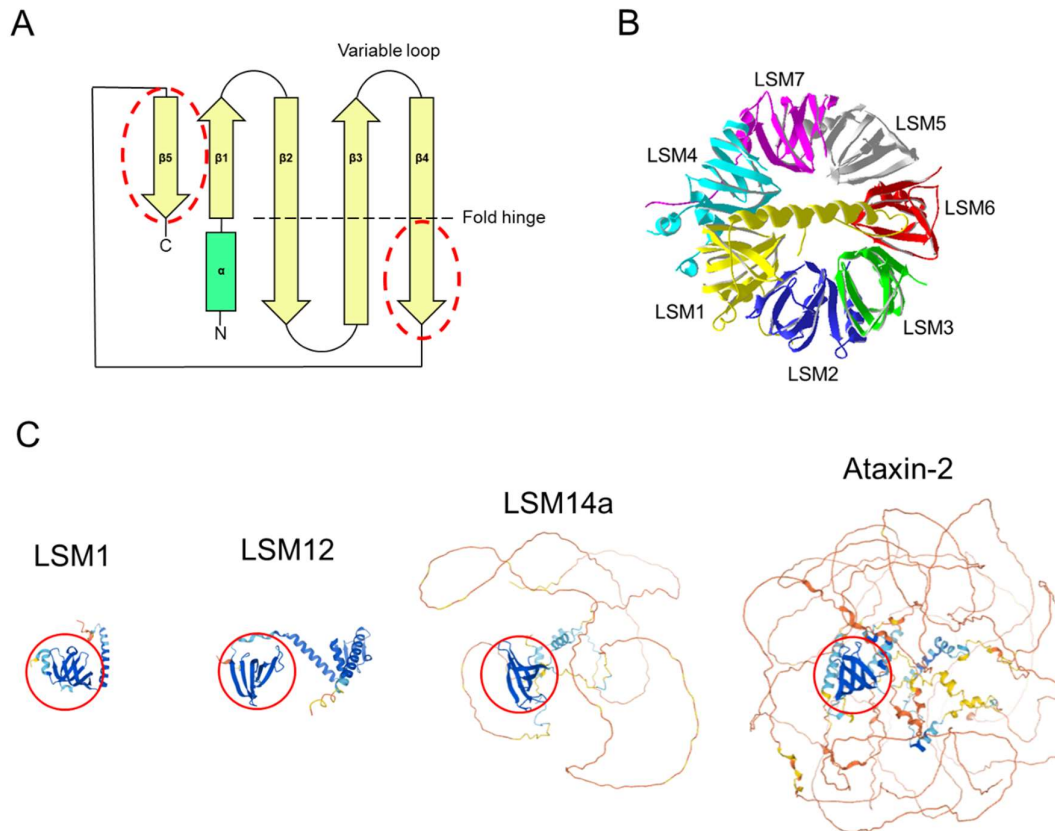


Figure 4.4: Structure of LSm domain and mechanism of LSm:LSm binding. (A) Structure schematic of a generic small LSm-domain containing protein. The LSm fold is highly conserved and consists in a five strand antiparallel cross β -sheet, arranged in a 51234 sequence. Three of the strands are characteristically bent at a 45° angle in the middle (sometimes referred to as a β -sheet sandwich structure). LSm:LSm interactions occur between the $\beta 4b$ strand of one LSm protein and the $\beta 5$ strand of the next LSm protein (circled in red). (B) Crystal structure of the cytosolic LSM1-7 complex (PDB ID: 4M75), involved in mRNA degradation (Zhou *et al.*, 2014). Predicted structures of different LSm-domain containing proteins from the AlphaFold database, with the LSm domain circled in red. LSM1 is an example of a small LSm-domain containing protein which functions in cells as a hetero-heptamer. Most of its sequence is the LSm domain itself. LSM12, LSM14a and Ataxin-2 are examples of LSm-domain containing proteins with long C-terminal tails. These have been shown to perform a variety of RNA/nucleotide binding related

functions and appear to be still capable of LSm:LSm binding (Jumper *et al*, 2021; Varadi *et al*, 2022).

The PAM2 domain. The PAM2 domain is a short (~ 10 - 14 amino acids long) domain, sometimes referred to as a short linear motif (SLiM) in recent literature. Its consensus sequence (L/P/F)X(P/V)XAXX(F/W)XP binds to a highly conserved region in the C-terminal of cytosolic forms of poly-A binding protein (PABPC1) known as the MLLE domain (Kozlov *et al*, 2001; Roy *et al*, 2002; Xie *et al.*, 2014). The PABPC1 protein is a key translational machinery component. It associates with the poly-Adenylate (poly-A) tails of mRNAs, bridging with the eIF4F complex at the 5' methylguanosine cap, inducing mRNA circularization and kick-starting translation initiation (Deo *et al*, 1999; Kahvejian *et al*, 2005; Safaee *et al*, 2012).

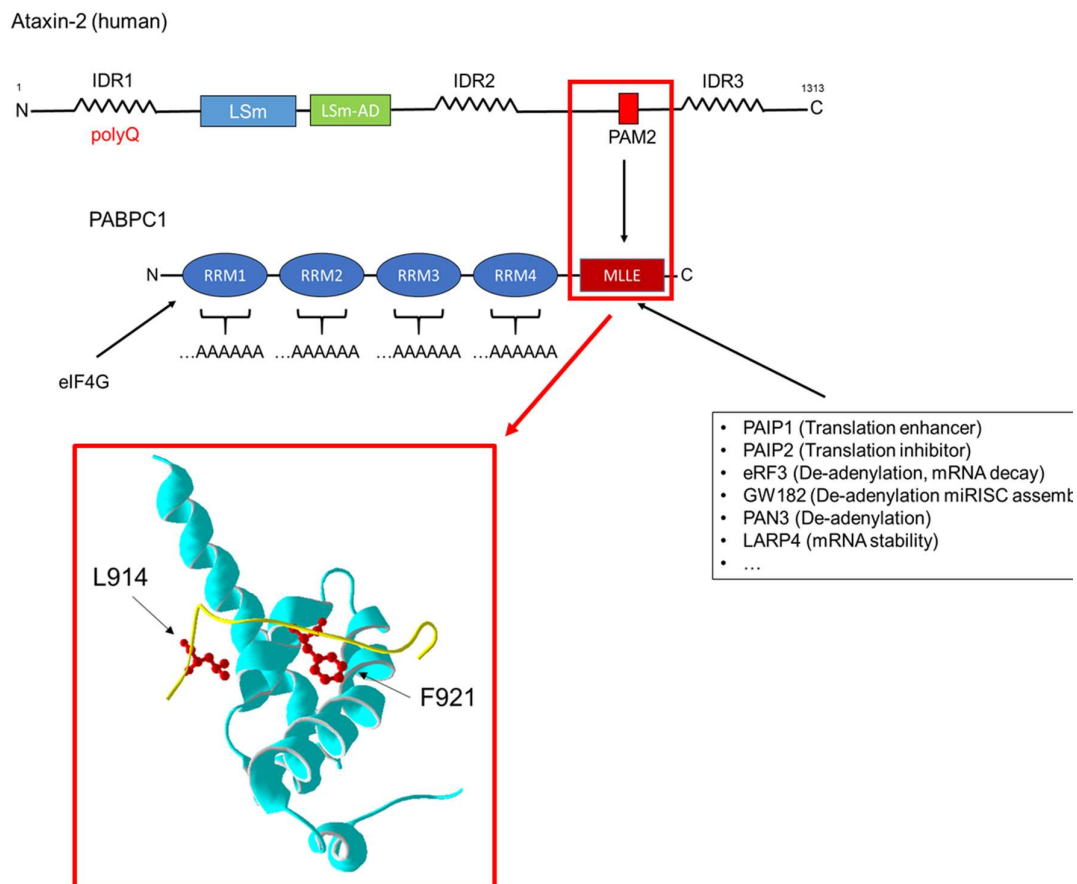


Figure 4.5: Ataxin-2 interacts with PABPC1 via the PAM2:MLLE interaction. The PAM2:MLLE binding is highly conserved and regulates the interaction of PABPC1 with a variety

of RNA binding proteins. The PAM2:MLLE binding is mostly determined by the interaction between two highly conserved hydrophobic residues on the PAM2 domain (leucine 914 and phenylalanine 921) with hydrophobic patches in the MLLE domain helices, shown in the crystal structure (PDB ID:3KTR) inside the red box (Kozlov *et al*, 2010). Over twenty known PAM2 domain containing proteins bind PABPC1, mostly playing RNA processing roles, while the translation initiation factor eIF4G binds PABPC1 at its RRM domains.

Over twenty proteins containing the PAM2 domain have been identified in humans (Xie *et al.*, 2014) (**Figure 4.5**). These perform a variety of functions, mostly related to mRNA processing. Sometimes, like in the case of Poly-A binding protein interacting proteins 1 and 2 (PAIP1- PAIP2), these functions oppose each other, and suggest sequential binding of different interactors with PABPC1 depending on the stage of a process (Roy *et al.*, 2002; Yoshida *et al*, 2006). PABPC1 is also a component of both PBs and SGs (Ivanov *et al.*, 2019). Binding of Atxn2 to PABPC1 is thought to direct Atxn2 to different mRNA and protein targets (Boeynaems *et al.*, 2021).

4.1.4 Atxn2, mRNP granules and disease

As mentioned earlier, most IDRs have the ability to transition from liquid-like states into solid amyloid fibrils with a characteristic cross- β structure *in vitro*, especially if incubated at high concentrations (Alberti *et al*, 2019; Li *et al.*, 2013; Murray *et al*, 2017; Patel *et al.*, 2015; Ramaswami *et al*, 2013). This information, when taken together with studies showing the protective effect in animal models of neurodegenerative diseases of eIF2 α kinase inhibitors or inhibitors of other downstream events (including SG formation) (Chou *et al*, 2017; Halliday *et al*, 2017; Sidrauski *et al.*, 2015; Wong *et al.*, 2018; Zyryanova *et al*, 2021) lead us to the following ideas:

- (1) mRNP granules provide a microenvironment where pathogenic protein seeds can form and grow (Bakthavachalu *et al.*, 2018; Mandrioli *et al*, 2020; Patel *et al.*, 2015).
- (2) Increased misfolded-protein loads can lead to inclusion formation, chronic stress signalling and reduced protein translation (Hetz *et al.*, 2020; Preissler & Ron, 2019).
- (3) This heavier load on protein handling systems results in multiple cellular defects, notably in the functions of membrane-less organelles (Alberti *et al*, 2017; Azkanaz *et al*, 2019; Jiang *et al*, 2020; Latonen, 2019; Schuller *et al*, 2021). In particular, aberrant SG formation also results in nuclear transport defects which may contribute to cell death and toxicity (Hochberg-Laufer *et al*, 2019; Zhang *et al*, 2018).

Atxn2 offers significant support to the idea that RNP granule formation plays a central role in promoting disease. Loss of Atxn2 has been shown to be cytoprotective in yeast TDP-43 and *Drosophila* TDP-43 or C9ORF72 or Tau models of cytotoxicity (Bakthavachalu *et al.*, 2018; Becker *et al.*, 2017; Elden *et al.*, 2010; Huelsmeier *et al.*, 2021; Kim *et al.*, 2014; Lee *et al.*, 2016; Shulman & Feany, 2003). In mouse models for SCA2 or ALS, either genetic loss of *ATXN2*, or delivery of antisense oligonucleotides (ASOs) targeting *ATXN2* in the central nervous system, was shown to reduce TDP-43 aggregation, which in turn increased animal survival and improved motor function (Becker *et al.*, 2017; Scoles *et al.*, 2017). It must be noted though, that in these diseases, the insoluble aggregated protein found in inclusion bodies is found in “surviving” cells, adding to the controversial question of if these inclusions are actually the end result of an overloading of the proteostatic machinery and amyloid fibrils not being the main cytotoxic component in these diseases. These observations have led to ASOs against human *ATXN2* being developed and approved for clinical trials (Biogen, 2021).

Given Atxn2's therapeutic significance and multiple roles in biology, it is particularly important to determine which molecular activities of the protein are relevant to disease and to its various biological functions (Kim *et al.*, 2020). Previous work in *Drosophila* has shown that a C-terminal IDR of Atx2 is selectively required for mRNP assembly into granules (Bakthavachalu *et al.*, 2018). Parallel experiments showing that the IDR is also required for cytotoxicity in *Drosophila* FUS, C9ORF72 and Huntington's disease models suggest RNP-granule formation to be a significant mechanism by which Atx2 promotes neurodegeneration (Bakthavachalu *et al.*, 2018; Huelsmeier *et al.*, 2021). A recent discovery that the Atx2-LSm domain antagonizes IDR-function has led to a model in which the Atx2 cIDR:

- (1) Does not support mRNP assembly when Atx2 is associated with actively translating mRNAs through Atx2-LSm domain interactions.
- (2) Becomes accessible and active in mediating mRNP assembly when LSm-domain interactions break and mRNA translation stalls (Boeynaems *et al.*, 2021; Singh *et al.*, 2021).

Following are a series of experiments further elaborating the mechanisms by which Atxn2 functions in mRNA translation and mRNP assembly. These show that the PAM2 motif of Ataxin-2 and its interactions with PABP are not essential for granule assembly but are required to efficiently recruit Atx2-target mRNAs and specific protein components into Atxn2 granules. When taken together with other findings (Boeynaems *et al.*, 2021; Kim *et*

et al., 2014; Satterfield & Pallanck, 2006; Singh *et al.*, 2021), these observations indicate that PAM2 binding to PABP on the poly-A tail of mRNAs helps specify the composition of Atxn2 granules. An early role for PAM2:PABP interactions working in coordination with the LSm domain to support mRNA translation and thereby oppose mRNP formation is also supported by the findings presented here (Boeynaems *et al.*, 2021); as well as a later role in escorting translationally-stalled PAM2:PABP linked mRNAs into mRNP granules. Regarding the LSm domain, it is demonstrated that the interaction between Atxn2 and LSM12 is a canonical LSm:LSm interaction; this interaction also preliminarily appears to be specific for this protein pair. *In vivo* experiments analysing motor decline in transgenic *Drosophila* indicate that the PAM2:PABP interactions also support the progression of neurodegenerative process. New evidence is provided as well as fresh insight into the enigmatic role of mRNP assembly in neurodegeneration.

4.2 Results

4.2.1 The structured PAM2 domain of Atx2 is necessary for the correct protein and mRNA content in Atx2 granules

Targets of RNA-Binding Proteins Identified by Editing (TRIBE) technology was used to identify Atx2 mRNAs associating with Atx2 and test if the PAM2 domain played a role in recognition & selection of these mRNA targets in *Drosophila* brains. Similar experiments had been recently carried out, targeting the LSm and LSm-AD domains, which didn't significantly affect mRNA target recognition (Singh *et al.*, 2021). *All TRIBE experiments & data analysis were carried out by Dr Amanjot Singh & Prof Baskar Bakthavachalu.*

Gal80^{ts}- controlled *elav-Gal4* was used to express Atx2ΔPAM2 conjugated to the catalytic domain of Adenosine deaminase acting on RNA (ADARcd) protein in brains of adult *Drosophila* for 5 days. Poly-A positive brain RNAs edited were identified using RNA sequencing (RNA-seq), and compared to those present in Atx2-ADARcd, using previously described methods (McMahon *et al.*, 2016). Briefly: the ADARcd domain converts Adenosine nucleotides to Inosines in RNAs and are identified as A to G transitions in TRIBE analyses (**Figure 4.6**). Samples were sequenced to obtain 20 million reads (**Supplementary table 6.2**), and edits were considered only from transcriptome regions containing upwards of 20 reads. High-confidence, true targets were selected if edits were identified at a threshold above 15% in two biological replicates. Edit frequency and edited gene identity were compared between brains of flies expressing Atx2ΔPAM2-ADARcd and brains of flies expressing Atx2-ADARcd.

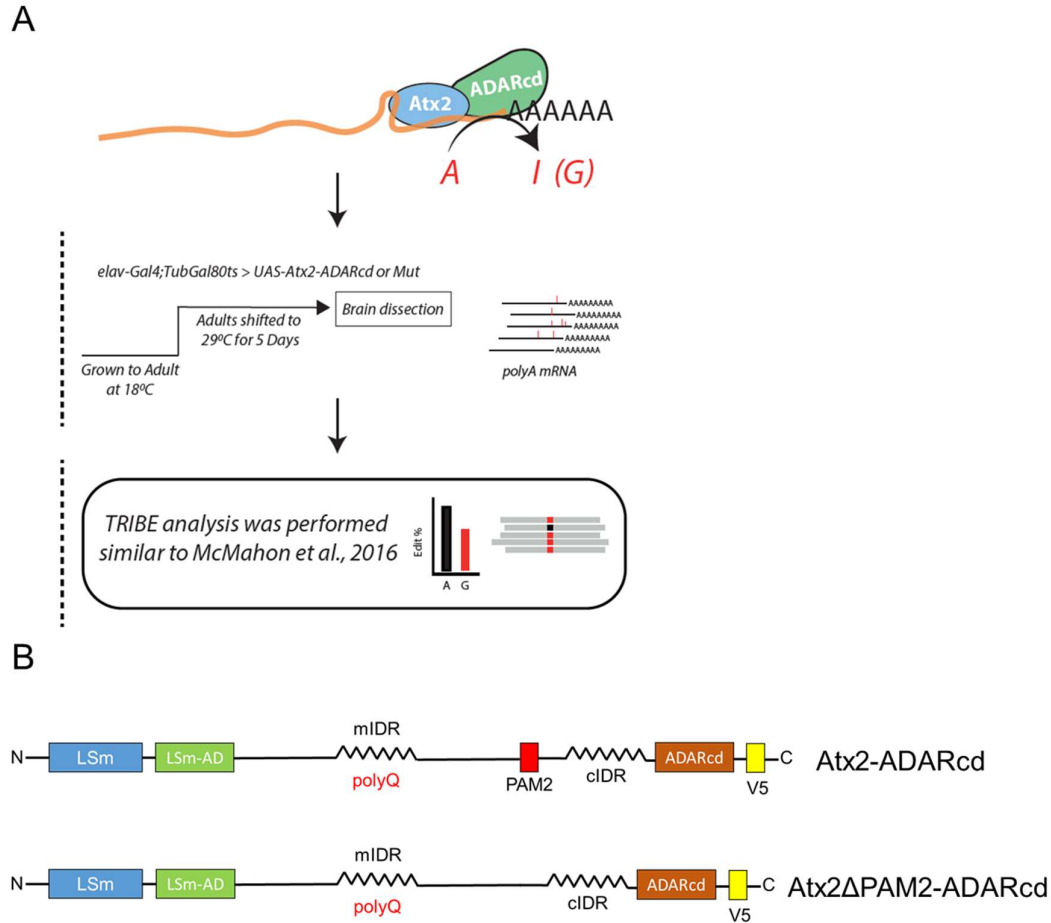


Figure 4.6: Summary of the TRIBES analysis pipeline and ADARcd constructs used. (A) Flowchart depicting the TRIBES analyses pipeline. Atx2ΔPAM2-ADARcd was expressed in adult *Drosophila* brains. Total brain RNA was isolated and RNA edits were identified and compared to Atx2-ADARcd, similar to Singh et al 2021. *Image courtesy of Prof Baskar Bakthavachalu.* (B) Domain map of Atx2-ADARcd (top) and Atx2ΔPAM2-ADARcd (bottom) constructs used for the TRIBES analysis.

From the TRIBES experiments, it was found that Atx2ΔPAM2-ADARcd edited significantly less mRNA targets than wild-type (WT) Atx2-ADARcd (108 genes, 165 edits vs 256 genes, 317 edits, **Figure 4.7, A & B, Supplementary table 6.3**) Additionally, the set of mRNA targets edited by the ΔPAM2 mutant was substantially different from the set edited by the wild type form of Atx2-ADARcd. Of the 108 genes edited by Atx2ΔPAM2-ADARcd, only 36 were also targets of Atx2-ADARcd (**Figure 4.7, A & B, Supplementary table 6.3**). Regarding the edit sites, 50 targets were found to be in common between the wild type

and the Δ PAM2 mutant; These were edited with much lesser efficiency by the Δ PAM2 mutant.

Location of the edits relative to the coding sequence of target mRNAs made by the Atx2 Δ PAM2 mutant was also different from WT Atx2-ADARcd and Δ LSm Atx2-ADARc which had been tested in previous work (Singh *et al.*, 2021). For both wild type and Δ LSm in fact, edits were mostly found in the 3' UTR of the mRNAs, while for the Δ PAM2 mutant there is no apparent preference for a given region of its target mRNAs (**Figure 4.7, C**). Taking all this information together, it appears that the PAM2 motif is necessary for Atx2 to bind its correct mRNA targets. The PAM2 motif binds PABP, which in turn binds poly-A tracts at the 3' end of mRNAs (Deo *et al.*, 1999) – hence this structured PAM2:PABP interaction plays a directing role for Atx2 to bind its target mRNAs, which could then proceed to be included in Atx2-positive granules.

Regarding the targets obtained with TRIBE, some particularly interesting hits emerged and are summarized in the supplementary information (**Table 6.3**). Targets common across wt Atx2 and Δ PAM2 Atx2 are shown in bold. Some interesting common targets – so probably recruited either *via* IDR or LSm mediated interactions – are:

Vmat: an essential monoamine transporter, responsible for neurotransmitter packing into presynaptic vesicles.

hipk: A Ser/Thr protein kinase that regulates homeodomain transcription factors.

Tob: A kinase involved in the regulation of mRNA turnover. It contains a PAM2 domain so could be a direct competitor to Atx2:PABP binding.

Mei-P26: A protein involved in a variety of processes. It has been shown to play a poorly characterized role in the nervous system, acting in seizure suppression (Glasscock *et al.*, 2005).

An interesting interaction which seems exclusive to wt Atx2 is with the **Atx2** gene itself. This suggests a possibly dose-dependent role in regulating levels of Atx2 itself in the cytoplasm.

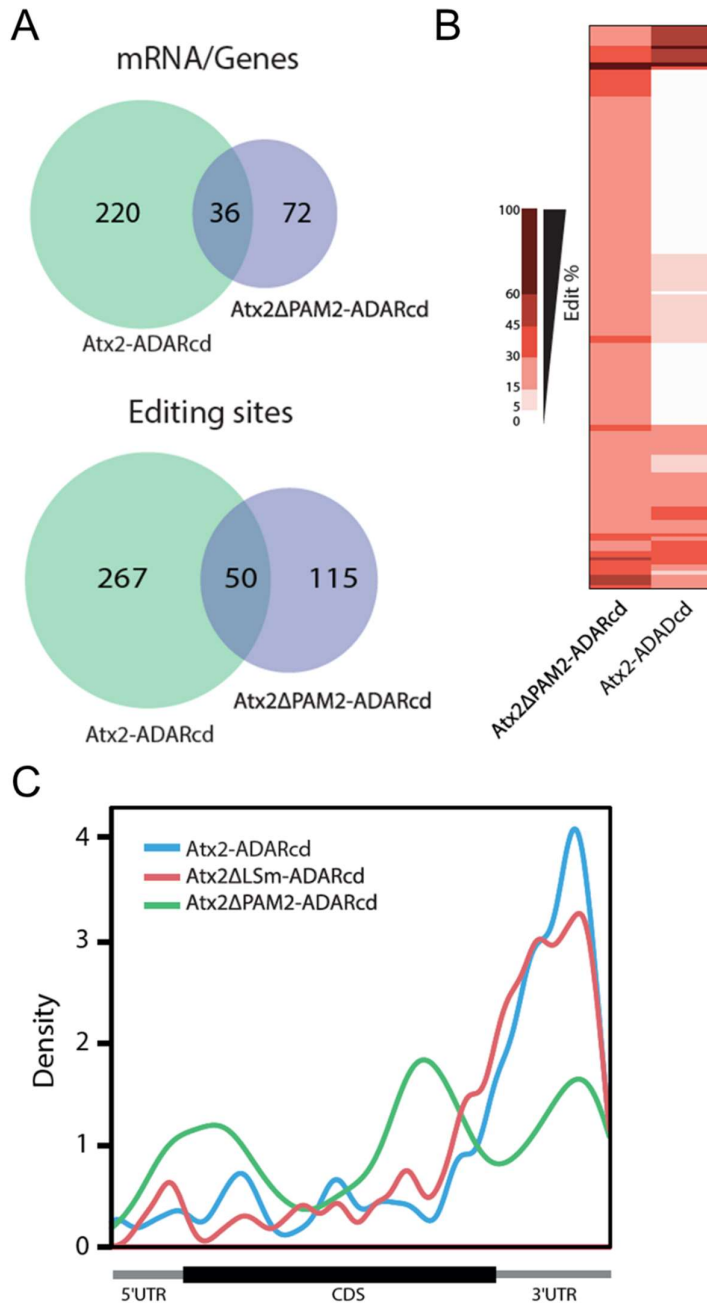


Figure 4.7: The PAM2 domain plays a directing role in Atx2 RNA target selection. (A) Comparison of genes and edits identified by TRIBE between Atx2-ADARcd and Atx2ΔPAM2-ADARcd targets. (B) Targets of the ΔPAM2 mutant that were identified by TRIBE are different from those of wild type Atx2; the lack of edits by part of the wild type, coupled with the lower levels of editing seen in the PAM2 targets suggests these to be caused by non specific of Atx2ΔPAM2 with non-target RNAs. (C) Loss of the PAM2 motif causes 3'UTR edit specificity to disappear, which was instead observed with ΔLSm mutant and the wild type protein (Singh *et al.*, 2021). Image courtesy of Prof Baskar Bakthavachalu.

Assuming that Atx2 interacts with mRNAs exclusively if present in protein-RNA condensates (Singh *et al.*, 2021), then the ability of the Δ PAM2 ADAR mutant to still edit some of the target mRNAs suggests that the PAM2 motif, while clearly important for mRNA targeting, is not essential for the formation of Atx2-positive mRNP granules. This was tested by expressing WT and Δ PAM2 mutant forms of GFP-tagged Atx2, controlled using the native genomic promoter in *Drosophila* S2R+ cells. WT Atx2-GFP overexpression in S2 cells caused the formation of mRNP granules closely resembling SGs, positive for Atx2 as well as for a battery of previously described SG markers, as reported previously (**Figure 4.8, A**) (Bakthavachalu *et al.*, 2018; Singh *et al.*, 2021). The Atx2 Δ PAM2-GFP mutant is also capable of causing the formation of mRNP granules upon overexpression in S2 cells, but the protein components recruited to these are distinct from the mRNP granules produced by WT Atx2-GFP overexpression. While containing some of the same SG markers, such as Me31B and Rox8 (*Drosophila* homologs of DDX6 and Tia1), others, such as PABP, Caprin and dFMRP were excluded (**Figure 4.8, B**). A likely explanation for these observations is that, in the case of Atx2-overexpression induced granules, while condensation itself is dependent on the presence of the Atx2 cIDR (Bakthavachalu *et al.*, 2018), the array of protein components that phase separate with it is dependent on the well-folded LSM and PAM2 domains. It is well known that the PAM2 domain binds PABP (Xie *et al.*, 2014), meaning that it's probably directly recruited to the Atx2-overexpression granules by Atx2. Caprin and dFMR are then either dragged in through a direct interaction with Atx2, with PABP or, more likely, bound to their mRNA targets that are dragged in together with PABP. Hence, their absence from Atx2 Δ PAM2-overexpression induced granules. Me31B and Rox8 probably remain present in these granules either through binding of the Atx2-LSM domain, or through IDR-IDR mediated interactions.

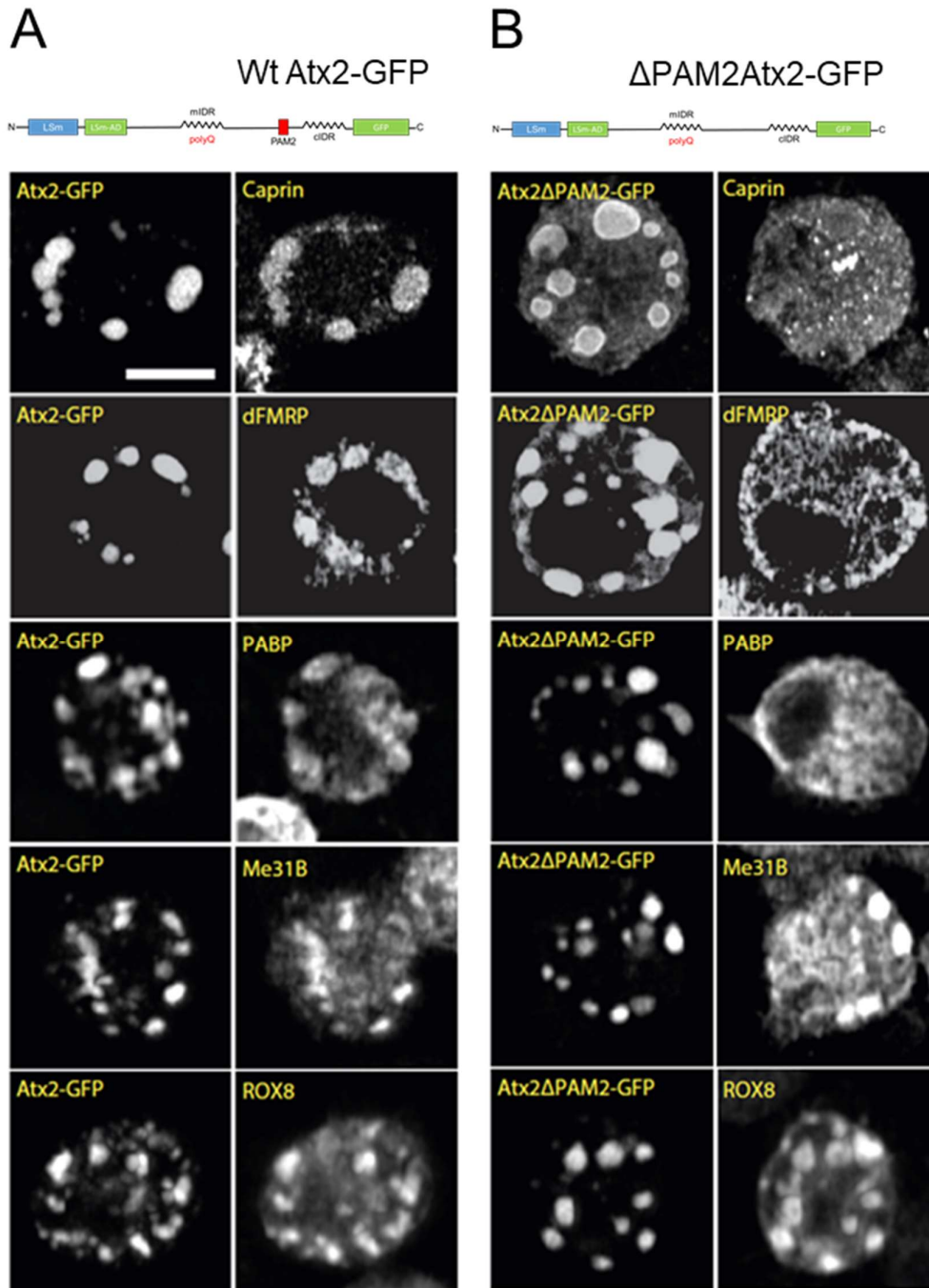


Figure 4.8: The PAM2 domain affects protein composition of Atx2-GFP overexpression induced granules in S2 cells. (A) Over-expression of Atx2-GFP in *Drosophila* S2 cells induces the formation of granules in which various SG markers co-localize. (B) Deletion of the PAM2 motif does not prevent granule formation but affects the protein composition of the granules. Some SG markers, in particular dFMR, Caprin and PABP fail to co-localize in these granules. Scale bar = 5 μm, applicable to both (A) and (B). *Staining and imaging carried out by Mr Arnas Petrauskas.*

4.2.2 PAM2 interactions with PABP are sufficient for Atx2 to associate with stress granules

In order to directly confirm Atx2 PAM2 motif interactions with PABP, and further understand the relevance of these to RNP granule formation, an approach previously shown for MeCP2 (Tillotson *et al*, 2017), consisting in designing radically truncated constructs and assessing their function, was applied. Constructs were generated encoding SNAP-epitope tagged Atx2/Atxn2 mutants. These consisted in forms of *Drosophila* and human Ataxin-2 containing only the LSm, LSm-AD and PAM2 elements, connected by short, flexible linkers. All the unstructured regions of the protein were deleted. These “minimized” Ataxin-2 constructs, together with their domain-deleted forms (**Figure 4.9**), allowed uncoupling of the functions of each structured region from those of the remaining long unstructured regions of Ataxin-2.

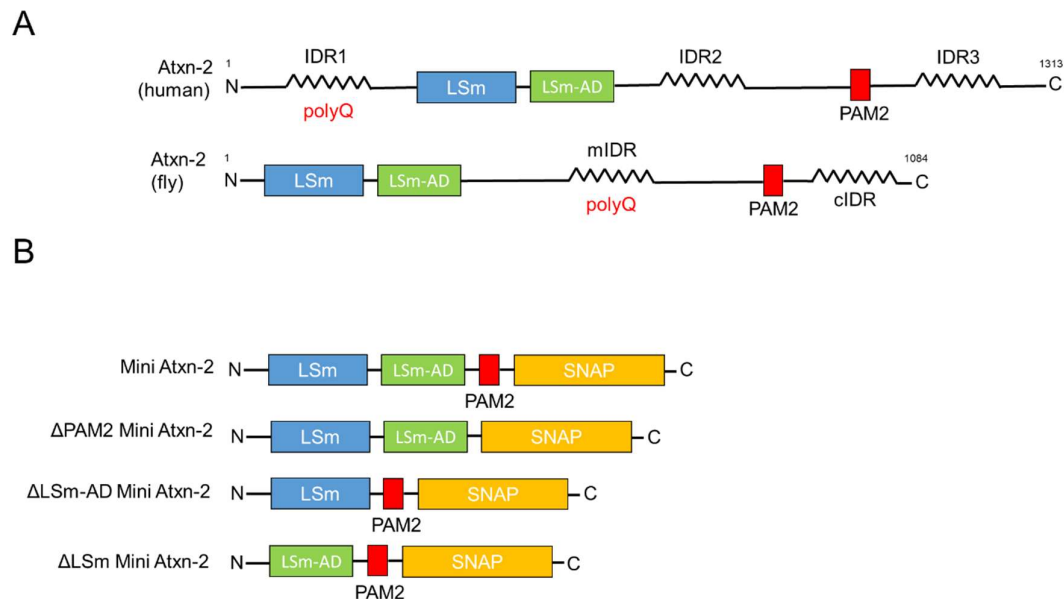


Figure 4.9: Schematic of minimized ataxin-2 constructs. (A) Domain structure of human and fly WT Ataxin-2. (B) Schematic of SNAP-tagged minimal and domain deleted constructs used to isolate the functions of the structured domains of the protein, disregarding the effects of the intrinsically disordered regions present. The structured domains of Ataxin-2 are highly conserved from fly to human, exhibiting high amino acid similarity (Clustal Ω) of 70% (LSm), 82% (LSm-AD) and 87% (PAM2). This suggests these domains maintain the same specific functions in both

organisms. Minimized constructs were made based both on human and fly Ataxin-2, each optimized for expression in fly/human cell lines respectively.

Furthermore, regarding the PAM2:PABP interaction, the key residues involved were identified, based on a previously solved crystal structure of a strongly conserved mammalian PAM2:PABPC1-MLLE domain complex (Kozlov *et al.*, 2010; Xie *et al.*, 2014) (**Figure 4.10, A**). Residues leucine 914 and phenylalanine 921 (L914 and F921) in the human Atxn2-PAM2 are predicted to interact hydrophobically with the PABPC-MLLE domain, and of these, F921 has previously been shown *via* mutagenesis followed by immunoprecipitation to be required for the PABPC1-Atxn2 interaction (Inagaki *et al.*, 2020). These residues (L859 and F866 respectively) are perfectly conserved in fly Atx2 (**Figure 4.10, B**). In order to demonstrate a more precise disruption of PAM2:PABP interactions, and to avoid potential unknown secondary effects of a full PAM2 motif deletion, mini Ataxin-2 constructs were generated where the predicted PABP interacting residues were singly or doubly altered to alanine. In the case of the human minimized Atxn2, a triple mutant was generated where glutamine 928 was, in addition to L914 and F921, altered to alanine, as it was predicted to add to the interaction *via* hydrogen bonding.

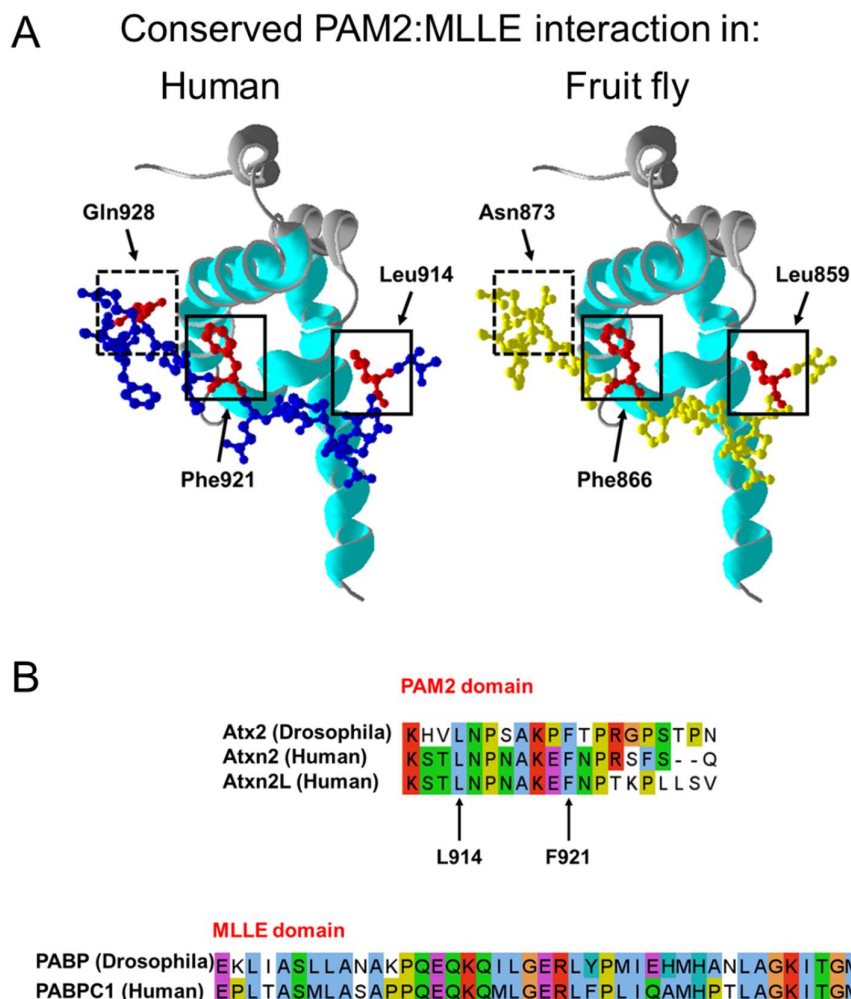


Figure 4.10: The Ataxin-2 PAM2 domain is conserved across human and fly, as is the PABP MLLE domain. (A) Structural model of the PABPC1 MLLE domain (ribbon) showing the near-perfect structural similarity of the human Atxn2 PAM2 domain (blue, uniprot ID: Q99700) with the *Drosophila* Atx2 PAM2 domain (yellow, uniprot ID: Q8SWR8). The *Drosophila* model for the Atx2 PAM2 domain was obtained by threading the *Drosophila* PAM2 sequence into a human Atxn2 PAM2:MLLE domain crystal structure (PDB ID: 3KTR). The key interacting residues that were chosen as point mutations to break this interaction are highlighted in red (Kozlov *et al.*, 2010). (B) Alignments of fly Atx2 with human Atxn2 and Atxn2L, as well as of the MLLE domains in fly and human PABP, showing the large degree of sequence similarity.

These constructs were then used for co-immunoprecipitation and co-localization analyses to examine how PAM2:PABP/PABPC1 affect RNP-granule formation.

SNAP-tagged WT and mutants of human and *Drosophila* mini Ataxin-2 were expressed in HEK293T and S2 cells respectively. Co-immunoprecipitations were carried out to test which Aaxin-2 domains are required for SNAP substrate beads to successfully

immunoprecipitate complexes containing PABPC1/PABP from cell lysates (**Figure 4.11**). Both the proteins LSM12 and PABPC1 proteins were co-immunoprecipitated with mammalian mini-Atxn2. However deletion of the PAM2 domain caused selective loss of PABPC1 co-immunoprecipitation, while deletion of the LSm domain caused selective loss of LSM12 co-immunoprecipitation. The LSm-AD domain did not have an effect on either PABPC1 or LSM12 pulldown.

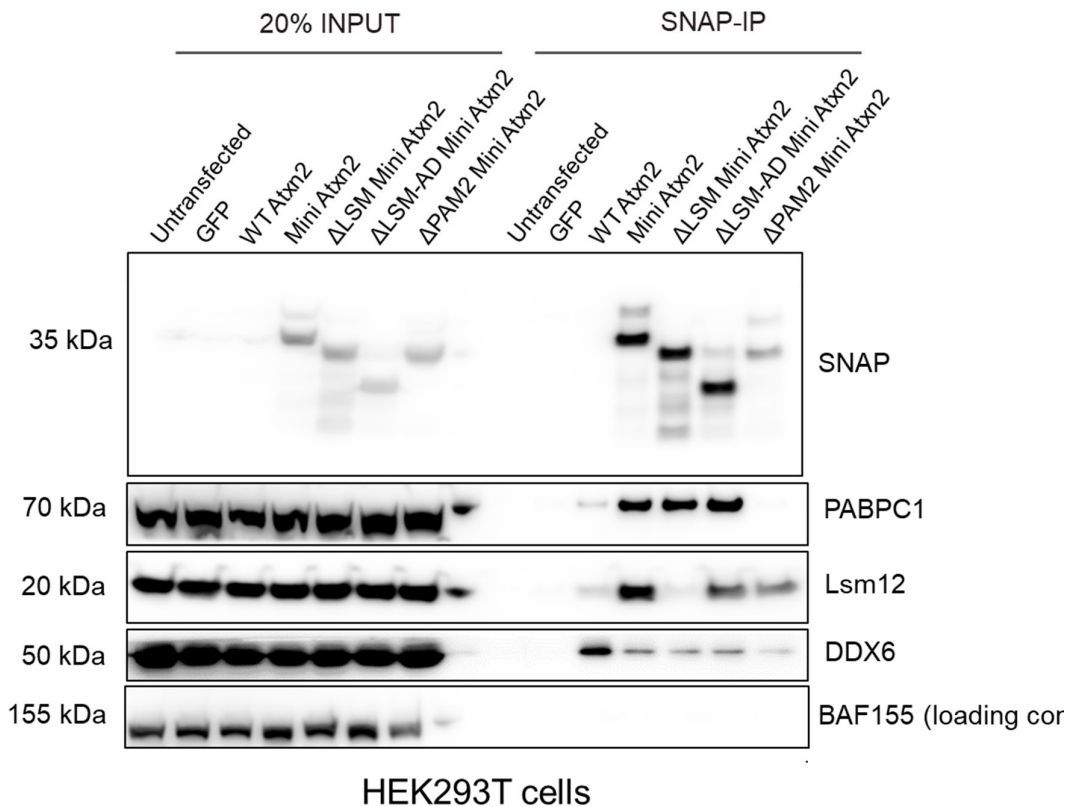


Figure 4.11: SNAP co-immunoprecipitation experiment carried out with wild-type and domain deleted minimized constructs. Human minimized Atxn2 SNAP IP-WB from HEK293T cells probing for PABPC1 and LSM12, showing the effects of different domain deletions. The PAM2 domain is necessary and sufficient for the Atxn2-PABP interaction, while the LSm domain is necessary and sufficient for the Atxn2-LSM12 interaction. Deletion of the LSm-AD domain did not appear to have a significant effect on pulling down either LSM12 or PABP. Regarding WT Atxn2, this result is not readily interpretable as the SNAP staining does not show a band, possibly due to problems transferring a large protein such as Atxn2 to the membrane. Interestingly, a faint band is observable in both the PABPC1 and the LSM12 stainings, and a strong DDX6 band is

present, suggesting that this helicase might be interacting with Atxn2 via regions or mechanisms other than its structured domains. *Experiment carried out by Mr Arnas Petrauskas.*

Following the co-immunoprecipitation results obtained with the full deletions, further co-immunoprecipitations with the PAM2 point mutants generated earlier (where the relevant PAM2 hydrophobic residues were switched to alanines) were carried out (**Figure 4.12**). All the PAM2 point mutations nearly perfectly replicated the effect of a full deletion of the PAM2 domain.

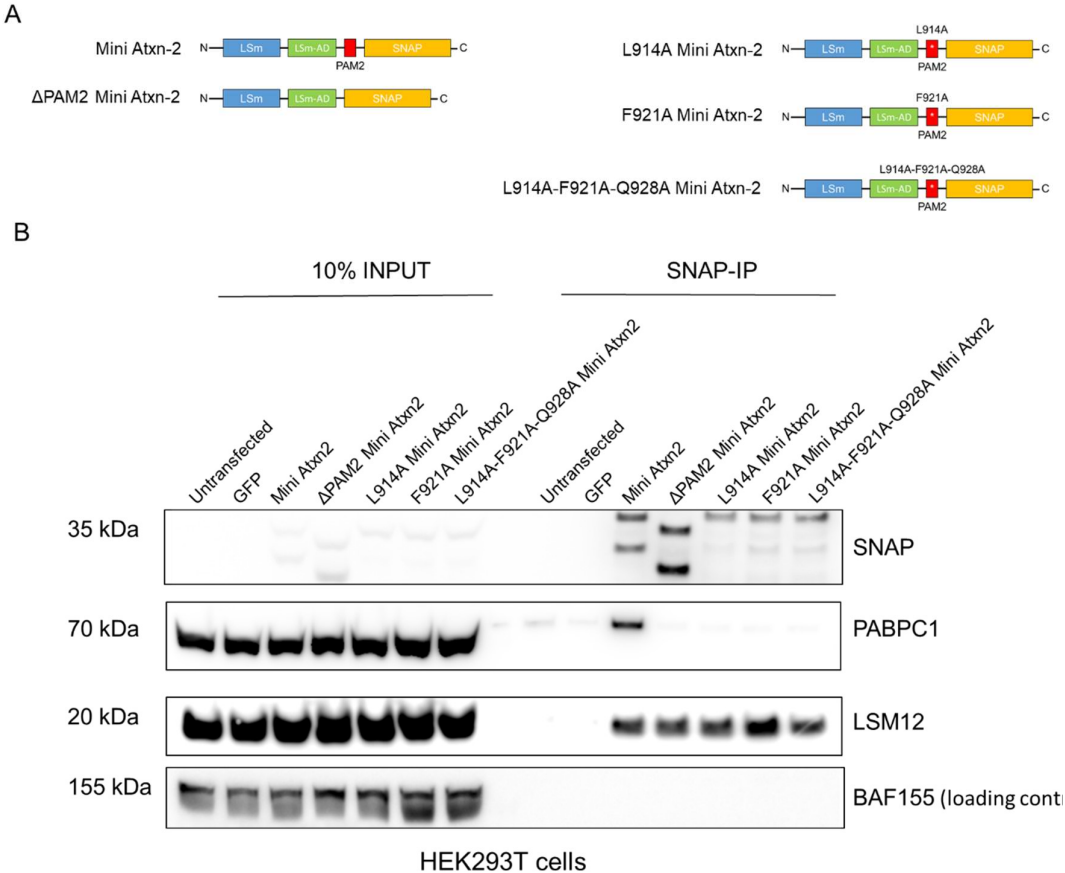


Figure 4.12: Single point mutations in key residues of the PAM2 domain are sufficient to replicate the effect of a full PAM2 deletion. (A) Domain schematics of the point mutants generated. (B) Point mutations targeting key interacting residues of the PAM2 domain were predicted to replicate the effect of a full PAM2 deletion in the minimized Atx2 construct. Human minimized Atxn2-SNAP IP WB from HEK293T cells showing exactly this: mutating either of the key hydrophobic residues L914 or F921 in the PAM2 domain is sufficient to prevent its interaction with PABPC1. The interaction with Lsm12 remains unaffected by any of the point mutations.

Together with the previous observation, the co-immunoprecipitation data seems to hint at the following sequence of molecular events: PABP/PABPC1 bound to mRNA poly-A tails

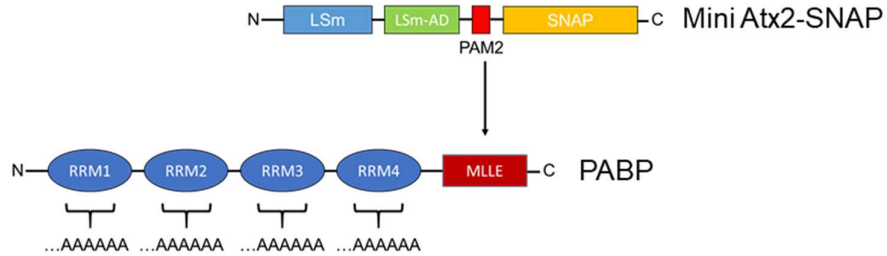
recruits Atx2/Atxn2 through the PAM2:MLLE domain interaction. If conditions are right, IDR-mediated condensation of these mRNPs into granules occurs.

4.2.3 Minimized Atxn2/Atx2 constructs are directed to stress granules via the PAM2:PABP interaction

Single mRNAs usually associate with multiple PABP molecules, as their poly-A tails are on average considerably longer than the ~24 bases required for PABP binding (Mangus *et al*, 2003). Henceforth, in cells expressing both endogenous Ataxin-2 and mini Ataxin-2, mRNAs could have both forms associated with their poly-A tails. While remaining diffuse in normal cellular function, in response to stress, mini Ataxin-2 would be expected to be recruited to SG *via* the same mechanism; formation of these SGs is facilitated by the IDRs on endogenous WT Ataxin-2 associated with the common target mRNAs. This hypothesis was tested in both *Drosophila* and human cells, before and after subjecting them to oxidative stress.

Drosophila and human mini Ataxin-2-SNAP, expressed in fly S2 or human U2OS cells respectively, remained diffuse in the cytoplasm and neither induced formation of Ataxin-2 granules. Upon exposure to sodium arsenite to induce SG formation, SNAP-tagged mini Atx2 (**Figure 4.13**) and mini-Atxn2 (**Figure 4.14**) were recruited to SGs. So, while Atx2 overexpression granules require the presence of its IDRs, the association of Ataxin-2 with mRNPs or mRNP granules only requires the presence of structured domain interactions.

A



B

S2 cells - Fly mini Atx2-SNAP 24 h transfection

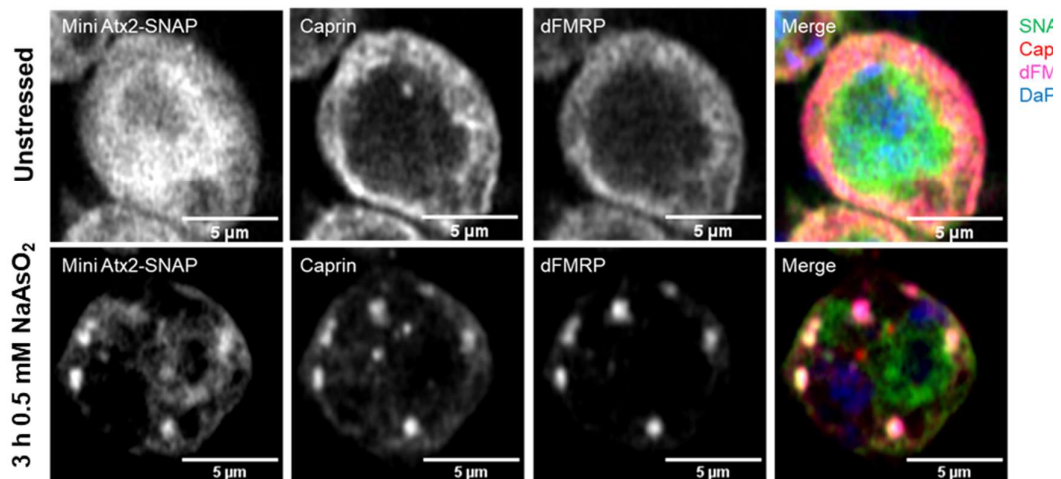


Figure 4.13: Mini Atx2 is recruited to SGs upon oxidative stress in S2 cells. (A) Proposed mechanism of mini Atx2 to SG: PABP binds the poly-A tails of target mRNAs through its four RRM domains (each recognizes ~ 4-6 adenosine bases). Mini Atx2 binds the PABP MLLE domain of some PABP molecules attached to the poly-A tail, and upon stress is recruited to SG granules together with wild type Atx2. (B) In unstressed S2 cells, mini atx2 remains soluble, as do SG markers Caprin and dFMRP. Upon oxidative stress, SG condensates (positive for SG markers Caprin and dFMRP) form, to which mini atx2 is recruited. As a small, soluble protein that is highly expressed, a significant percentage of mini atx2 remains soluble even under stress conditions.

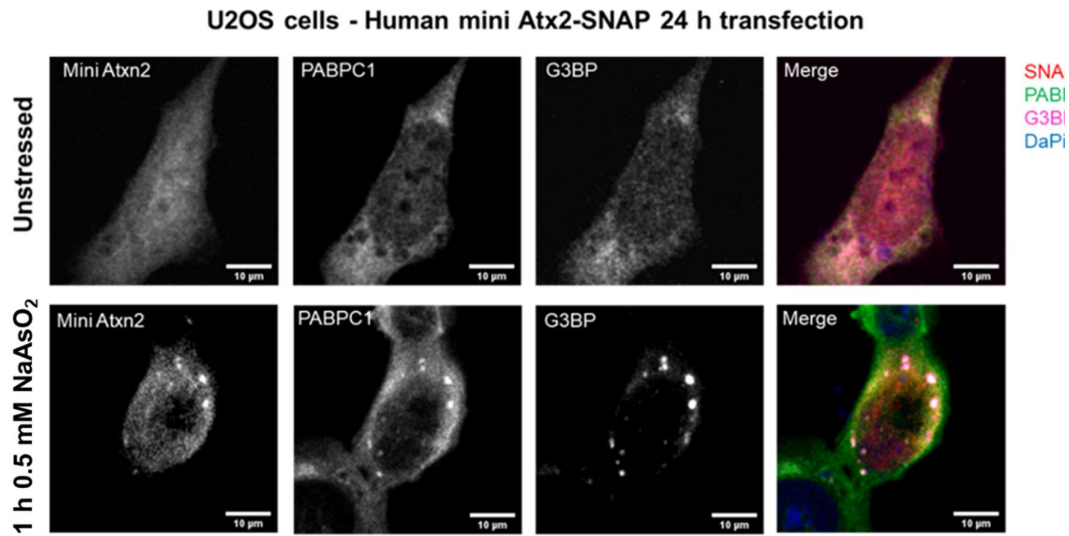


Figure 4.14: Mini Atxn2 is recruited to SGs upon oxidative stress in U2OS cells. As in S2 cells, in human U2OS cells mini Atxn2 remains soluble in normal conditions but is recruited to SGs upon oxidative stress.

Further experiments examined which of the structured domains was necessary for mini Ataxin-2 to associate with SGs, if the LSm, the LSm-AD or the PAM2 domain (**Figure 4.15**). *Drosophila* and human mini Ataxin-2 forms missing the PAM2 domain remained diffusely cytoplasmic even after stress in both S2 (**Figure 4.15, B**) and U2OS cells (**Figure 4.16, B**) while mutants lacking the LSm or LSm-AD domains were still capable of co-localization to SGs (**Figures 4.15, 4.16, C & D**).

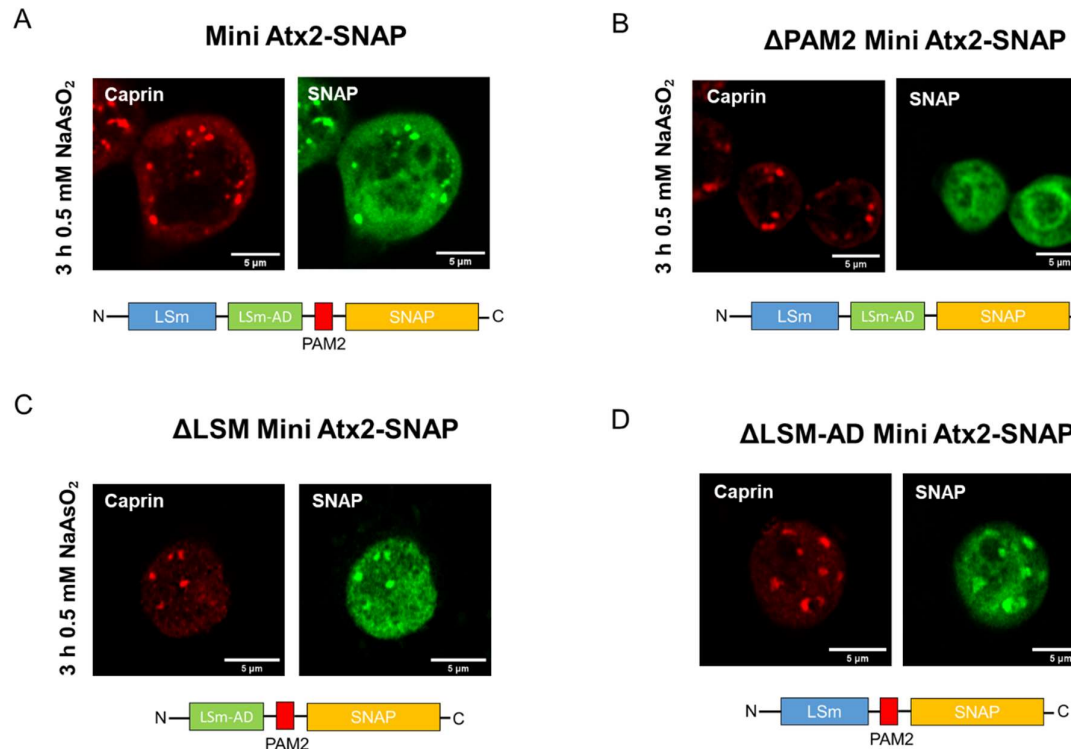


Figure 4.15: Deletion of the PAM2 domain, but not of the LSm and LSm-AD domains prevent mini Atx2 SG recruitment in S2 cells. (A) In *Drosophila* S2 cells mini Atx2-SNAP (green) is recruited to SGs induced by oxidative stress. (B) Deletion of the PAM2 domain prevents mini Atx2-SNAP from being recruited to SGs. (C) Deletion of the LSm domain does not have an effect on the ability of mini Atx2-SNAP to be recruited to SGs. (D) Likewise, the LSm-AD deletion does not affect the ability of mini Atx2-SNAP to be recruited to SGs. *Staining and imaging of ΔLSm and ΔLSm-AD constructs carried out by Mr Arnas Petrauskas.*

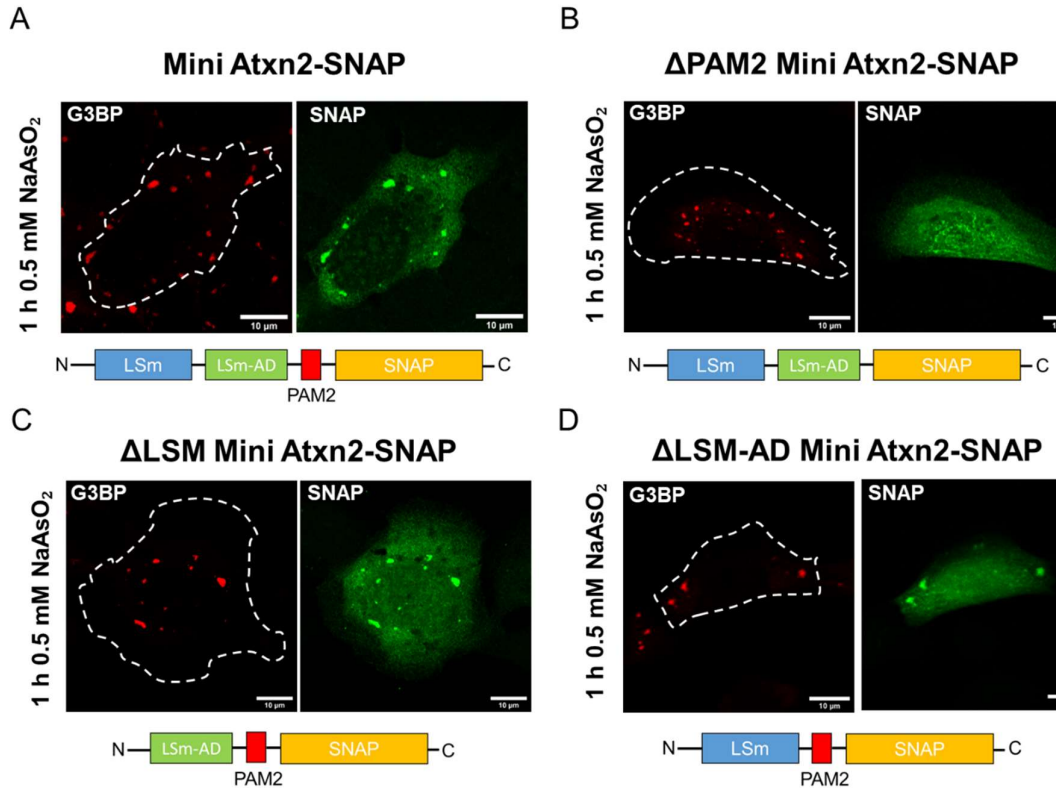


Figure 4.16: Deletion of the PAM2 domain, but not of the LSm and LSm-AD domains prevent mini Atx2 SG recruitment in U2OS cells. (A) In human U2OS cells mini Atxn2-SNAP (green) is recruited to SGs induced by oxidative stress. (B) Deletion of the PAM2 domain prevents mini Atxn2-SNAP from being recruited to SGs. (C) Deletion of the LSm domain does not have an effect on the ability of mini Atxn2-SNAP to be recruited to SGs. (D) Likewise, the LSm-AD deletion does not affect the ability of mini Atxn2-SNAP to be recruited to SGs. *Staining and imaging carried out by Mr Arnas Petrauskas.*

Having shown through co-localization experiments that the PAM2 domain is indeed responsible for targeting Ataxin-2 to SGs, the point mutants described previously were tested in both *Drosophila* S2 cells (**Figure 4.17**) and human U2OS cells (**Figure 4.18**). These exhibited the exact same behaviour as the Δ PAM2 mutants.

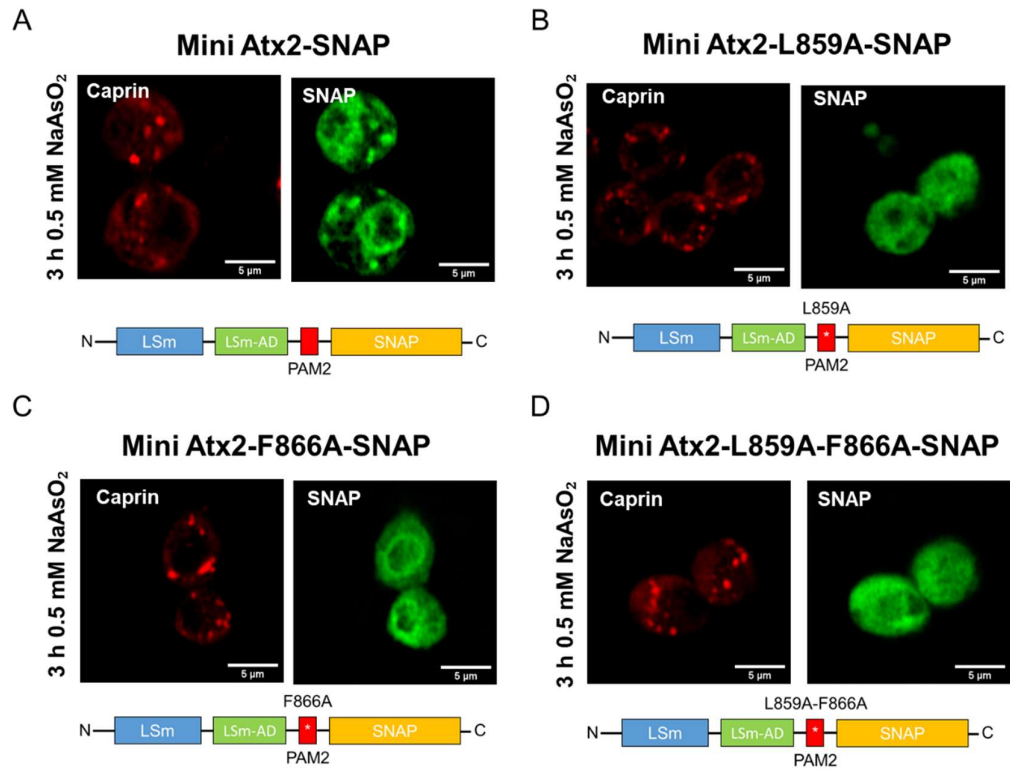


Figure 4.17: Point mutants in the PAM2 domain prevent fly mini Atx2 from being recruited to oxidative stress induced SGs in S2 cells. (A) Mini Atx2 co-localizes with SG components in *Drosophila* S2 cells upon being subjected to oxidative stress. (B-D) Point mutants targeting the hydrophobic interactions of the PAM2:MLLE domain, specifically leucine 859 to alanine, phenylalanine 866 to alanine and a double mutant with both residues turned to alanines cause mini Atx2 to remain diffuse in S2 cells, replicating the phenotype observed using the full PAM2 deletion.

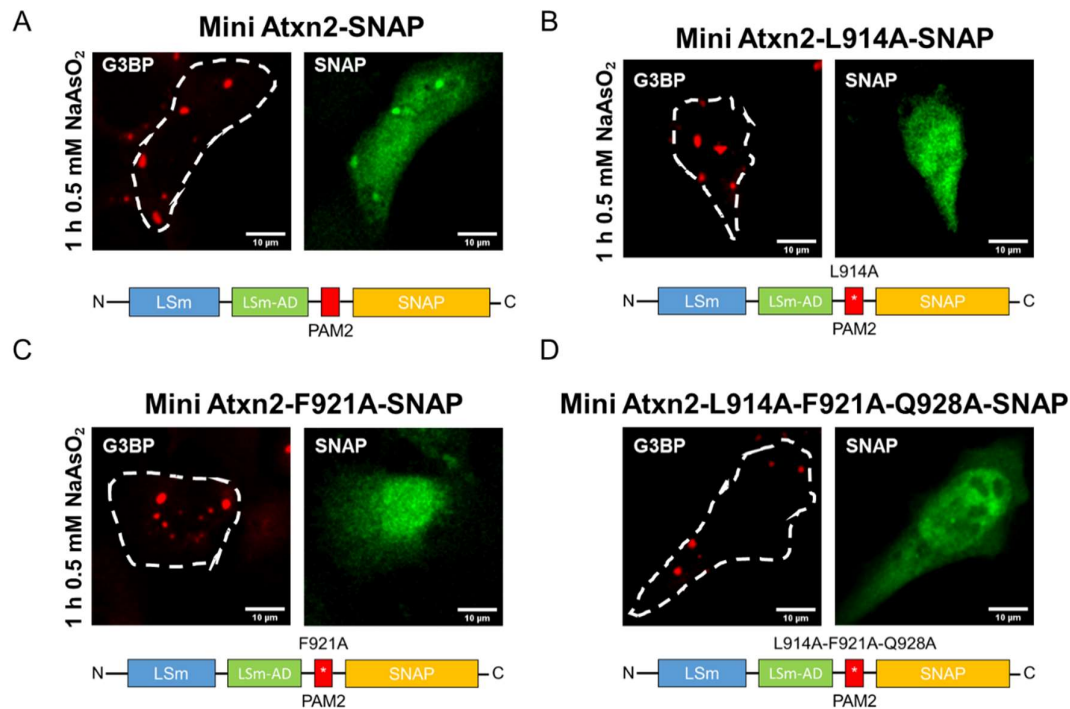


Figure 4.18: Point mutants in the PAM2 domain prevent human mini Atxn2 from being recruited to oxidative stress induced SGs in U2OS cells. (A) Mini Atxn2 co-localizes with SG components in human U2OS cells upon being subjected to oxidative stress. (B-D) Point mutants targeting the hydrophobic interactions of the PAM2:MLLE domain, specifically leucine 914 to alanine, phenylalanine 921 to alanine, as well as a triple mutant with both these residues and glutamine 928 (which appears hydrogen bound to the MLLE domain in the crystal structure (Kozlov *et al.*, 2010)) turned to alanines cause mini Atxn2 to remain diffuse in U2OS cells, replicating the phenotype observed using the full PAM2 deletion.

4.2.4 Atxn2 interacts with the protein LSM12 via a canonical LSm-LSm interaction

Having designed and generated point mutants in the PAM2 domain and shown that they could replicate the behaviour of minimized Atxn2 with the PAM2 domain fully deleted, the basis of the Atxn2 LSm:LSM12 interaction was investigated. The LSm domain itself was immediately found to be necessary for minimized Atxn2 to pull-down the protein LSM12 (**Figure 4.11**), which had been previously predicted and shown to interact with Atxn2 (Albrecht & Lengauer, 2004; Lee *et al.*, 2017). LSM proteins are known to form RNA-binding ring complexes. Examining existing crystal structures of these complexes (Zhou *et al.*, 2014) and running multiple sequence alignments (**Figure 4.19**) putative binding surfaces for an Atxn2 LSm:LSM12 interaction were identified.

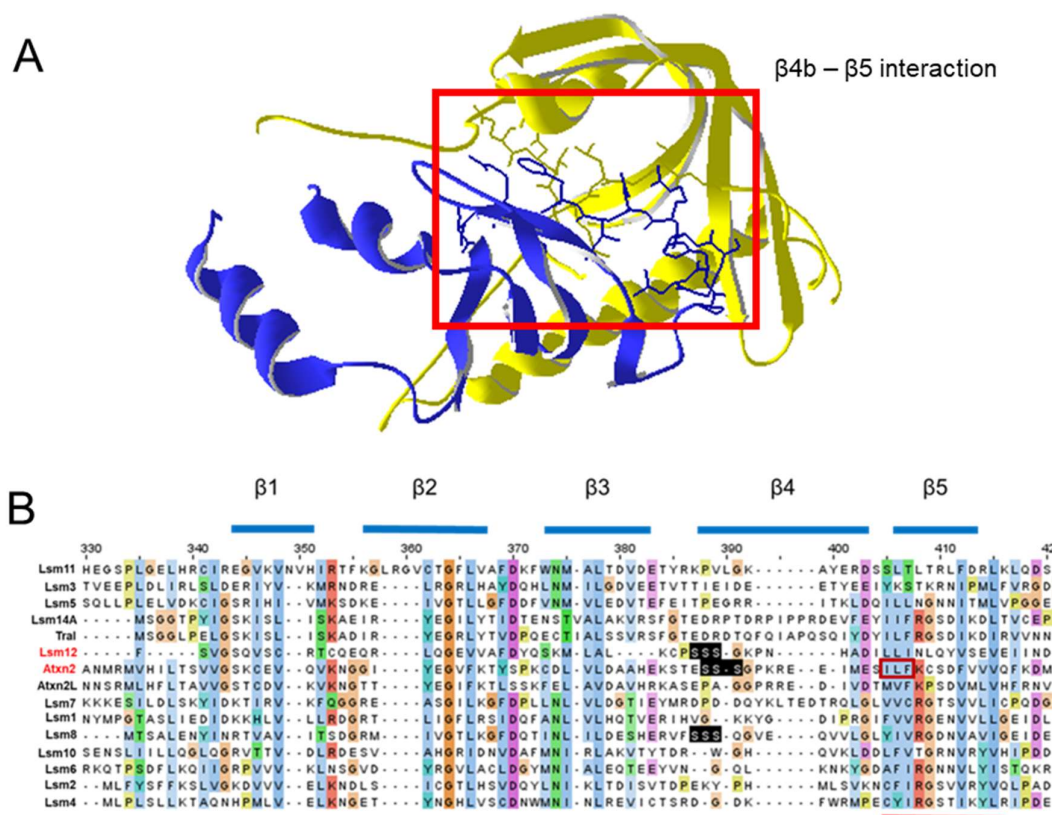


Figure 4.19: LSM protein alignment and mutant design. (A) Crystal structure (PDB ID: 4M7D) of canonical LSM:LSM binding (Zhou *et al.*, 2014). The $\beta 5$ strand of one Lsm protein binds the $\beta 4b$ strand of the next. (B) Multiple sequence alignment (Clustal Ω) across different Lsm proteins. Sequence and structural features are highly conserved. Lsm12 and the Atxn2 LSM domain exhibit higher sequence similarity with each other than with other LSM proteins. Highlighted in red are the mutations made for the block deletion and point mutants.

A canonical LSM:LSM interaction occurs between the $\beta 5$ strand of one of the Lsm proteins and the $\beta 4b$ strand of the other one. If the Atxn2 LSM:LSM12 interaction is canonical, deleting/mutating residues in this position should prevent mini Atxn2 from pulling down LSM12 in a co-immunoprecipitation experiment. Two mutants were generated, one block deletion of a portion of the $\beta 5$ strand (residues 328:339) and another, more ambitious mutant, where a triplet of highly conserved hydrophobic residues (isoleucine 328, leucine 329 and phenylalanine 330) in the putative binding surface were mutated to alanines. Indeed, both the block deletion and the point mutant failed to pull-down Lsm12 during co-immunoprecipitation experiments (Figure 4.20).

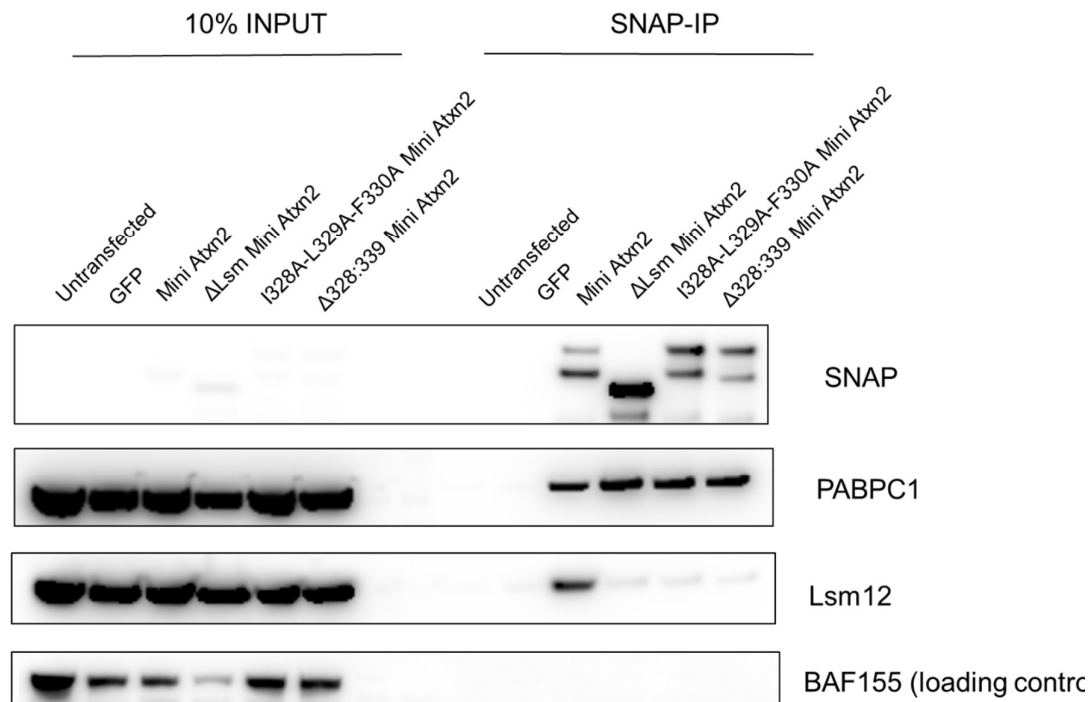


Figure 4.20: Point mutations in the Ataxin-2 LSm domain are sufficient to prevent the mini Ataxin-2:LSM12 interaction. A point mutant targeting a triplet of predicted key interacting residues of the Atxn2-LSm domain (I328-L329-F330), as well as a full deletion of the predicted interacting β -strand (Δ 328:339) were tested to see if they would replicate the effect of a full LSm domain deletion in the minimized Atx2 construct. Human minimized Atxn2-SNAP IP WB from HEK293T cells shows that either mutating the I328-L329-F330 triplet to alanine or deleting residues 328:339 in the LSm domain replicate the behaviour of a full LSm deletion. Lsm12 is no longer pulled down, while the mini Atxn2:PABPC1 interaction remains unaffected by the mutations.

As both mutations prevented mini Atxn2 from pulling down LSM12, it seems probable that the Atxn2 LSm:LSM12 interaction is a typical LSm:LSm interaction. Next, the possibility that mini Atxn2 might pull down other LSm-domain containing proteins was investigated. Specifically, a further human mini Atxn2-SNAP IP WB from HEK293T cells was carried out, this time probing also for Lsm14a, a PB marker known to bind DDX6 (Brandmann *et al*, 2018) and for Atxn2 itself, in case homo-dimerization/multimerization is occurring. The results showed (**Figure 4.21**) that mini Atxn2 does not pull down either Lsm14a or Atxn2, suggesting that the Atxn2 LSm:LSM12 interaction is specific. It must be noted though that WT Atxn2 is a very large protein with long intrinsically disordered stretches, so it's absence from the pull downs may not be readily taken as proof of lack of interaction with mini Atxn2.

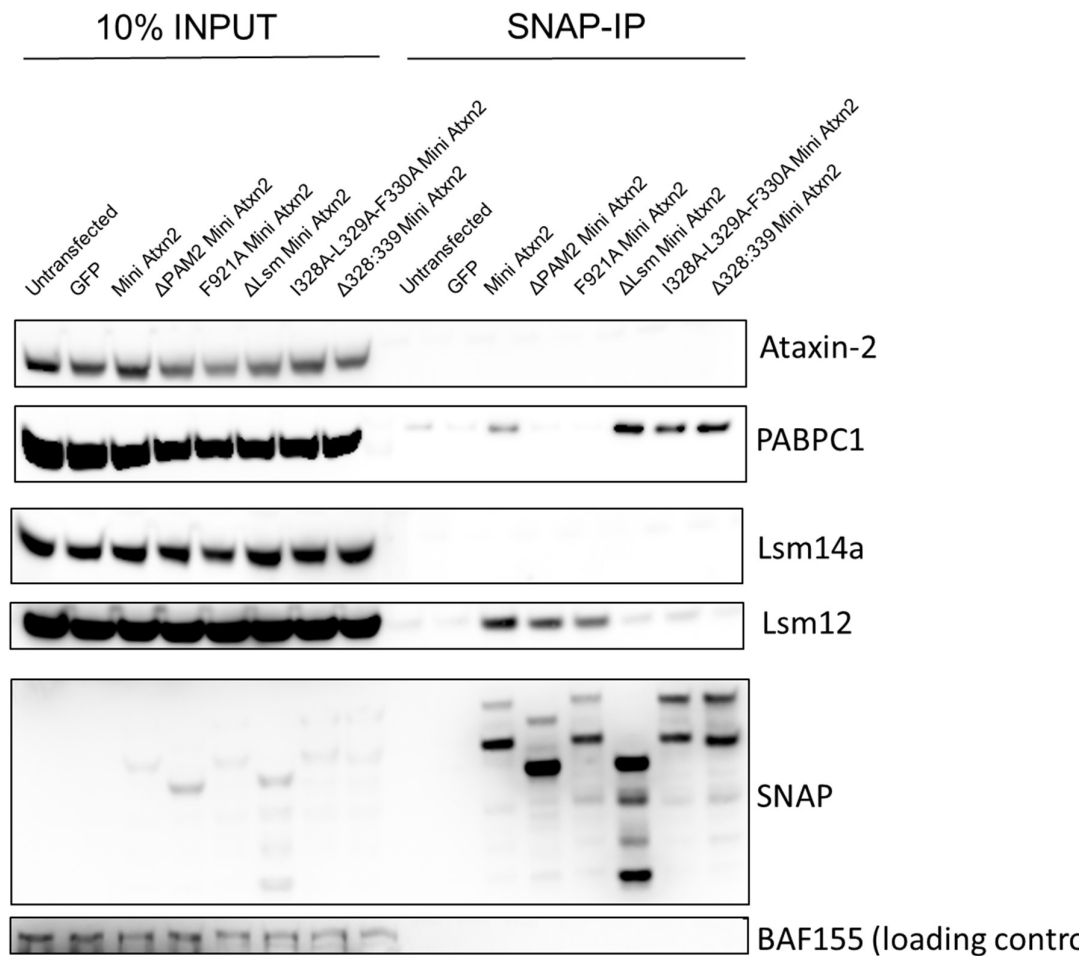


Figure 4.21: Lsm14a is not pulled down by mini Atxn2, regardless of the presence of the LSm domain. Lsm14a and Atxn2 are not pulled down by any of the mini Atxn2 constructs. Lsm12 is pulled down only if the Lsm domain is intact, and PABPC1 is pulled down only if the PAM2 domain is present/un-mutated.

4.2.5 The Atx2/Atxn2/Atxn2L IDRs are not functionally equivalent

In parallel to isolating and studying the roles of the different structured domains of Ataxin-2, the IDRs were also examined. In order to investigate if the Ataxin-2 IDRs ability to cause granule formation upon overexpression is sequence dependent, constructs were generated where the structured domains of human Atxn2 and human Atxn2L like were swapped with the structured domain of *Drosophila* Atx2 and vice-versa (**Figure 4.22**).

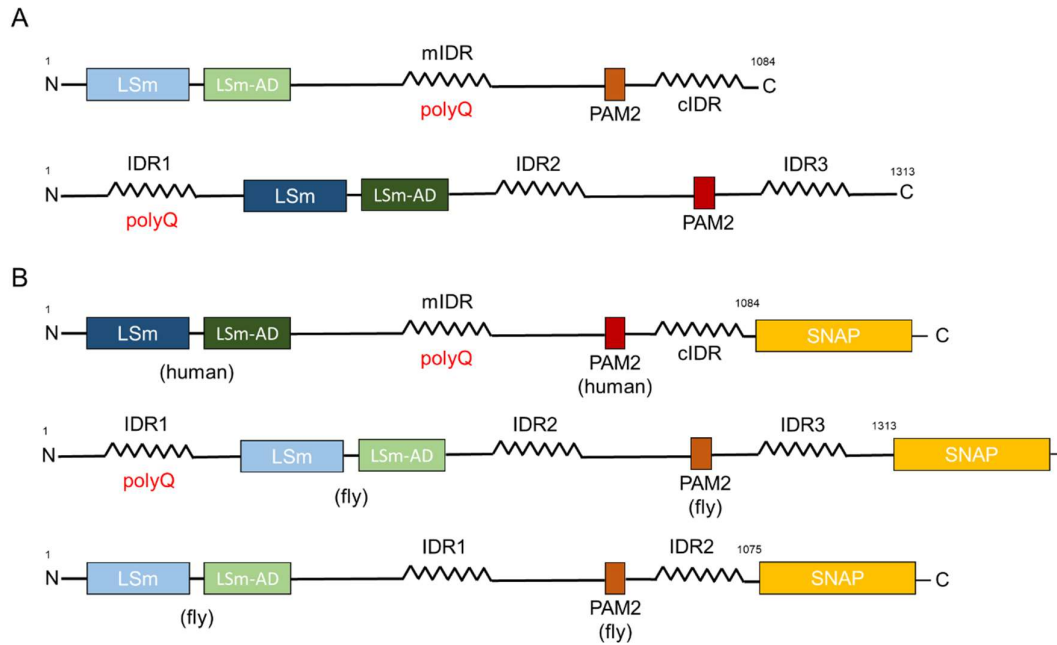


Figure 4.22: Human Atxn2/Atxn2L:Fly Atx2 structured domain swaps. (A) Domain structures of fly (top) and human (bottom) Atxin-2. (B) Domain schematics of the constructs generated to test differences in IDR behaviour. From top to bottom: human Atxn2 structured domains inserted in fly Atx2, fly Atx2 structured domains inserted in human Atxn2 and fly Atx2 structured domains inserted in human Atxn2L.

The construct containing the fly IDRs and human structured domains (HSDI Atx2) was transfected into U2OS cells to see what effect they would have. These formed condensates in both stressed and unstressed cells that appeared to co-localize upon oxidative stress – albeit not very strongly - with PABPC1, WT Atxn2 (**Figure 4.23**), LSM12 and G3BP (**Figure 4.24**).

U2OS-Atxn2-Human Structured domains-Drosophila Atx2 IDRs-SNAP

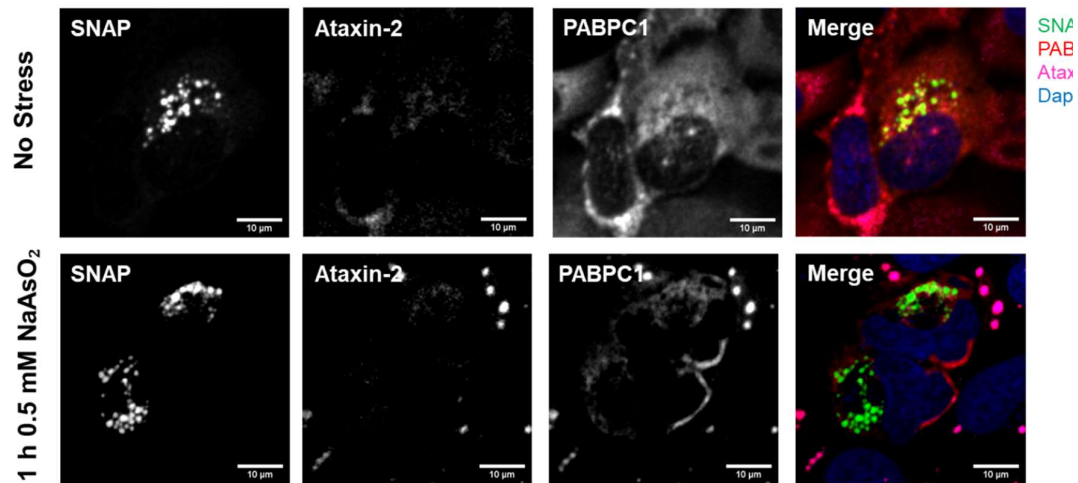


Figure 4.23: Construct with human Atxn2 structured domains and fly Atx2 IDRS (HSDI) expressed in U2OS cells, stained for PABPC1 and wild type Ataxin-2. The HSDI Atx2 construct formed condensates in U2OS cells in both stressed and unstressed conditions. Upon stress, WT Atxn2 and PABPC1 showed some evidence of low levels of co-localization with the HSDI construct. Interestingly, successfully transfected cells did not show evidence of normal stress granule formation after being subjected to oxidative stress.

U2OS-Atxn2-Human Structured domains-Drosophila Atx2 IDRs-SNAP

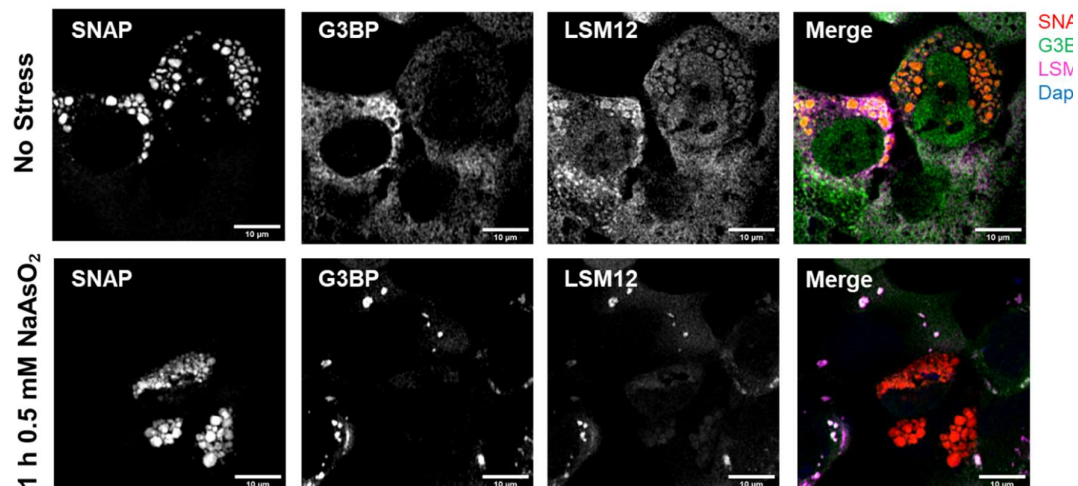


Figure 4.24: Construct with human Atxn2 structured domains and fly Atx2 IDRS (HSDI) expressed in U2OS cells, stained for G3BP and LSM12. The HSDI Atx2 construct formed

condensates in U2OS cells in both stressed and unstressed conditions. Upon stress, WT G3BP and LSM12 showed some evidence of low levels of co-localization with the HSDI construct. Interestingly, successfully transfected cells did not show evidence of normal stress granule formation after being subjected to oxidative stress.

The constructs containing the fly structured domain and the human Atxn2/Atxn2L IDRs (Atxn2-DSHI and Atxn2L-DSHI) were transfected into S2 cells to see if they would have an effect on condensation. The Atxn2-DSHI construct didn't form clear, localized condensates either in unstressed or stressed cells, but rather formed widespread areas of high concentration (**Figure 4.25**). Upon stress, SG markers Caprin and dFMRP seemed to be enriched in these areas. The Atxn2L-DSHI, somewhat unexpectedly, formed condensates in both unstressed and stressed S2 cells (**Figure 4.26**). Upon stress, Caprin and dFMRP seemed to co-localize to these granules to a certain extent. In both cases, as with the Atx2-HSDI, transfected cells did not form normal SGs upon oxidative stress.

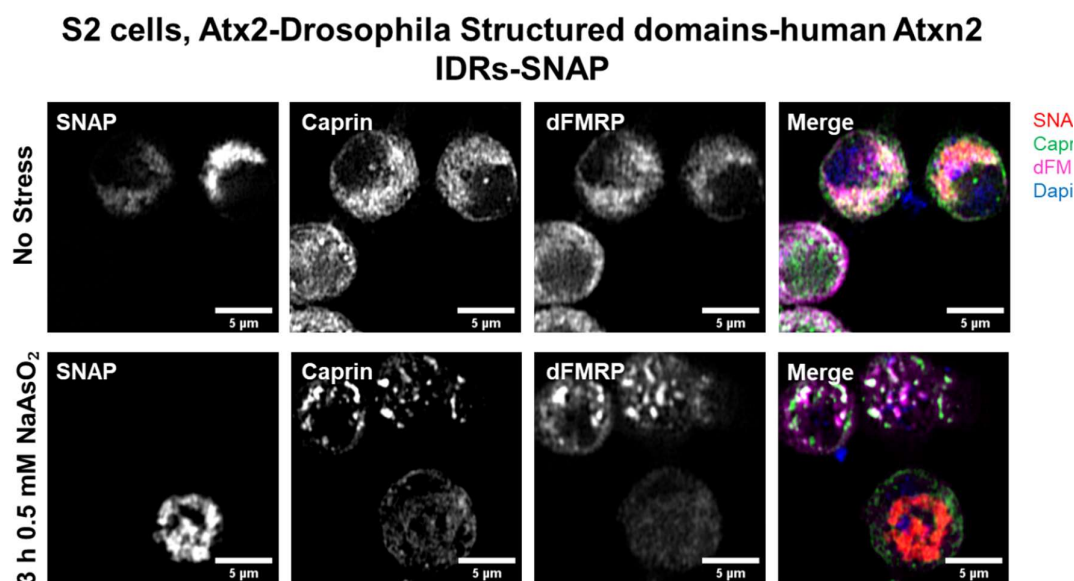


Figure 4.25: Construct with *Drosophila* Atxn2 structured domains and human Atxn2 IDRS (DSHI-Atxn2) expressed in S2 cells, stained for SG markers Caprin and dFMRP. Both before and after oxidative stress, the DSHI-Atxn2 construct does not form distinct puncta, but rather large spread out areas where it is at high concentration. Upon stress, Caprin and dFMRP appear to possibly co-localize to some extent with the construct. It was hypothesized that the lack of signal from these markers in transfected cells upon stress may reflect antibody penetration issues to condensates containing the DSHI-Atxn2 construct, that may have undergone a liquid-to-solid transition.

S2 cells, Atx2-Drosophila Structured domains-human Atxn2L IDRs-SNAP

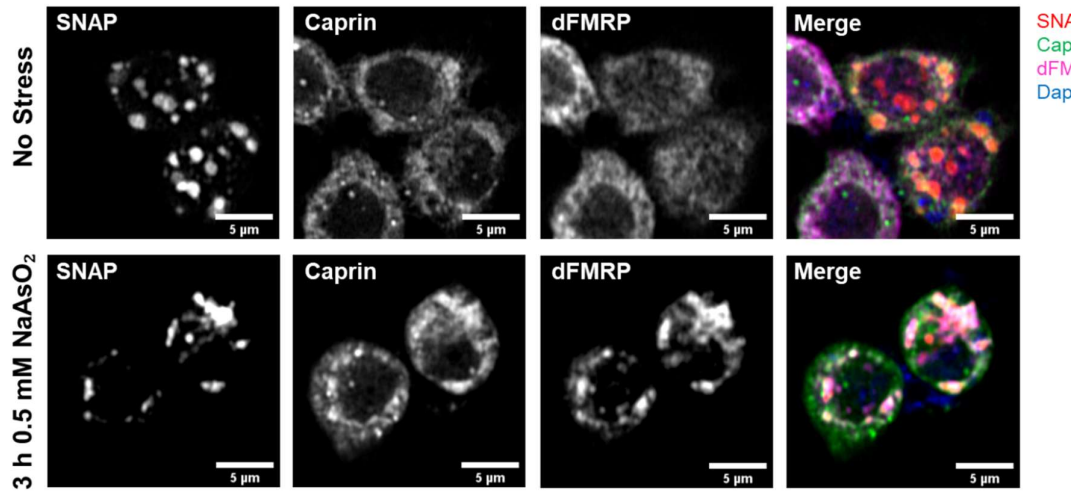


Figure 4.26: Construct with Drosophila Atxn2L structured domains and human Atxn2 IDRS (DSHI-Atxn2L) expressed in S2 cells, stained for SG markers Caprin and dFMRP. Both before and after stress, the DSHI-Atxn2L construct forms large granules, which upon stress appear to co-localize with SG markers Caprin and dFMRP.

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

The three different Atx2 domain deletions tested showed distinct effects on mRNP granule assembly in S2 cells. (i) IDR domain deletions prevented Atx2 granule formation, (ii) LSm-domain deletion enhanced the formation of Atx2 granules and (iii) PAM2 domain deletions resulted in the formation of unusual mRNP assemblies containing a different cohort of proteins (Bakthavachalu *et al.*, 2018; Singh *et al.*, 2021). Prior observations showing that Atx2 IDR deletions suppress cytotoxicity in *Drosophila* models for neurodegeneration indicate that mRNP granules are implicated in events that lead to degenerative disease (Bakthavachalu *et al.*, 2018; Becker *et al.*, 2017; Huelsmeier *et al.*, 2021; Scoles *et al.*, 2017). If there is a causative link between disease and granule condensation, then the expression of Atx2ΔLSm, which enhances granule assembly, should promote or potentially accelerate degeneration, while the expression of Atx2ΔcIDR should not. The expression of Atx2ΔPAM2 would be expected to support mRNP assemblies of different

composition from the ones containing WT or Δ LSm forms of Atx2. The effects on degeneration for this condition would be hard to predict.

In order to examine how the different Atx2 domain deletions affect nervous system integrity and function over time, a combination of a Gal4-responsive *UAS-Atx2* transgene with *elav-Gal4* and *TubGal80^{ts}* was used. This system allowed the use of a 18°C to 30°C temperature shift to induce *UAS-Atx2* transgene expression specifically in the brains of adult flies (**Figure 4.27, A**). It must be noted that the use of a temperature sensitive Gal4 system in *Drosophila* mutant may partially affect the results of the study, as (i) fly behaviour might be affected by the difference in temperature and (ii) temperature may play a role on the progression of molecular processes involved in neurodegeneration, such as protein aggregation and activation of heat-shock pathways. The rate at which flies climbed the walls of a glass cylinder was then analysed, which was interpreted as a proxy of motor ability, one-day and 15 days after transgene expression. All the genotypes tested exhibited robust and comparable levels of climbing ability on day 1. On day 15 variations in climbing ability across genotypes were instead observed. The 15-day old flies expressing Atx2WT or Atx2 Δ LSm showed a strong decline in climbing ability. On the other hand, in Atx2 Δ cIDR flies climbing ability deteriorated to a much lesser extent (**Figure 4.27, B**). These observations are consistent with the effects of these Atx2 types on granule formation. Strikingly, flies expressing the Atx2 Δ PAM variant, which formed compositionally distinct granules in S2 cells, showed no significant decline in climbing ability, suggesting that Atx2's ability to promote progressive decline of neural function doesn't depend as much on granule formation in itself, but rather the sequestration of critical factors such as PABP and its associated mRNAs (**Figure 4.27, B**). These observations support and extend prior work demonstrating that heterologous overexpression of full-length, but not PAM2-domain deleted forms of mammalian Atxn2 enhances mammalian TDP-43-induced degeneration of the *Drosophila* compound eye (Kim *et al.*, 2014). They are also consistent with work in mice showing that PABPC1 sequestration in inclusions correlates strongly with neurodegeneration progression (Damrath *et al.*, 2012).

The conclusion that PAM2 mediated interactions are required for progressive cytotoxicity is supported by a series of experiments done in parallel, in which *mef2-Gal4* was used instead of *elav-Gal4*, to *target* *UAS-Atx2* transgene expression to *Drosophila* adult muscles (**Figure 4.27, C**). Micro computed tomography (micro-CT) scanning was used to visualize the integrity of flight muscle fibres in whole-mount preparations revealing muscular degeneration in 20-day-old flies expressing WT Atx2. More severe degeneration

was observed in muscles of flies expressing Atx2 Δ LSm, while the muscles of similarly expressing Atx2 Δ clDR or Atx2 Δ PAM forms showed no morphological defects (**Figure 4.27, C**). *Experiments carried out by Khushboo Agrawal*. Western blots were then carried out on fly heads for the different transgenes (**Figure 4.28**), to ensure similar levels of expression across transgenes and that expression was indeed occurring in a temperature-dependent manner. This was found to be the case, although some larval expression was also observed.

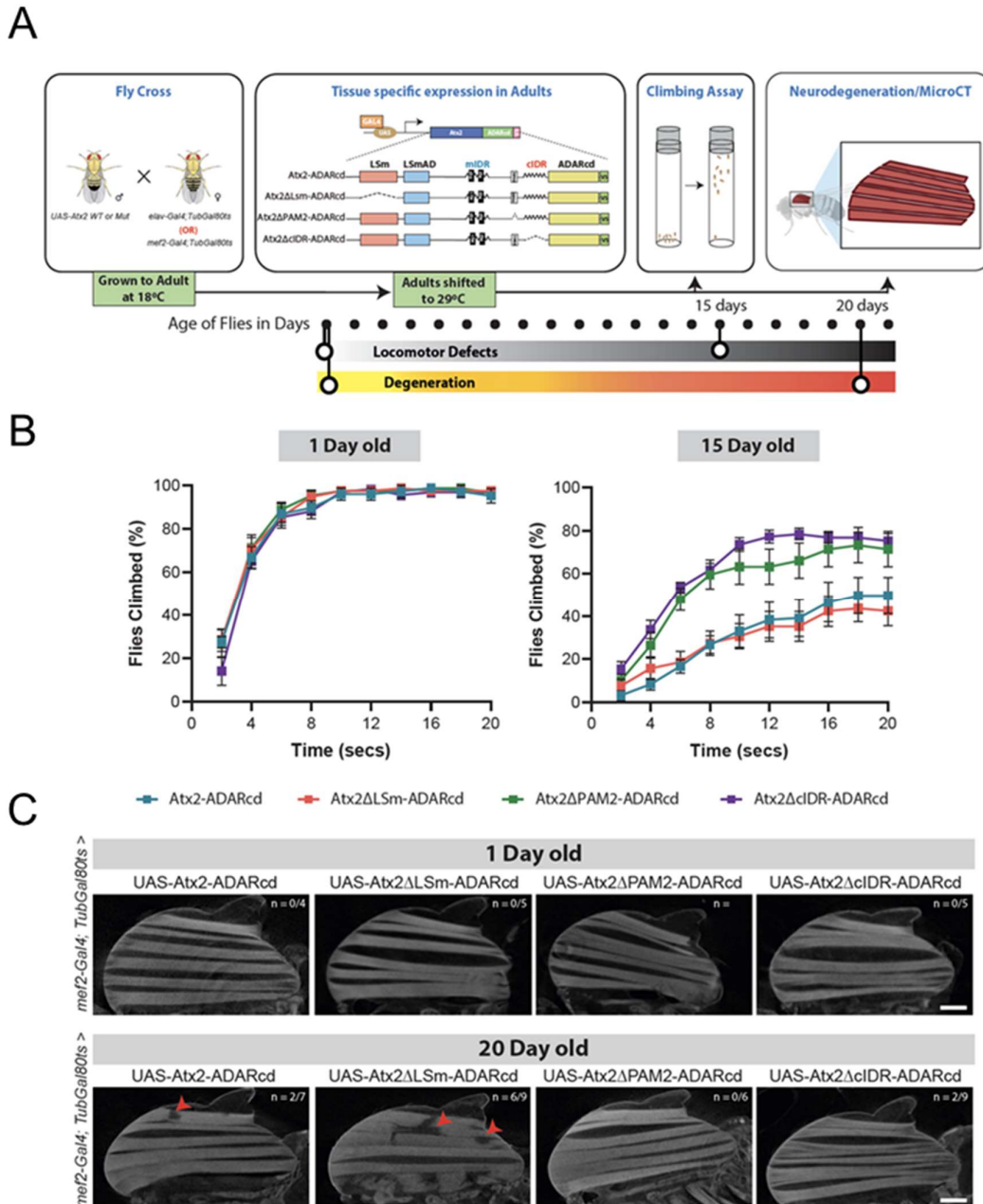


Figure 4.27: The IDR and the PAM2 domains promote while the LSm domain inhibits neurodegeneration in *Drosophila*. (A) Diagram depicting the experimental design for these experiments. The different UAS-transgenes were crossed with *elav-Gal4 ;TubGal80^{ts}* or *mef2-Gal4 ;TubGal80^{ts}* and kept at 18 °C until adult flies emerged. The flies were then shifted at 29 °C for 15/20 days depending on the following experimental steps. Either fly climbing or muscle cytotoxicity experiments were then carried out. (B) The *Drosophila* climbing behavior experiments were performed by driving the UAS-transgenes (Atx2WT, Atx2ΔcIDR, Atx2ΔPAM or Atx2ΔLSm) with *elav-Gal4*. A graph of the number of flies that managed to cross the 20 ml mark (shown on the Y-axis) against time (shown on the X-axis) was plotted. (C) Cellular toxicity

was measured by driving the UAS-transgenes (Atx2^{WT}, Atx2 Δ cIDR, Atx2 Δ PAM or Atx2 Δ LSm) with *mef2-Gal4*. Indirect flight muscles of the flies were imaged using micro-CT and the loss of muscle fibers is shown with solid red arrowheads. *Image courtesy of Baskar Bakthavachalu.*

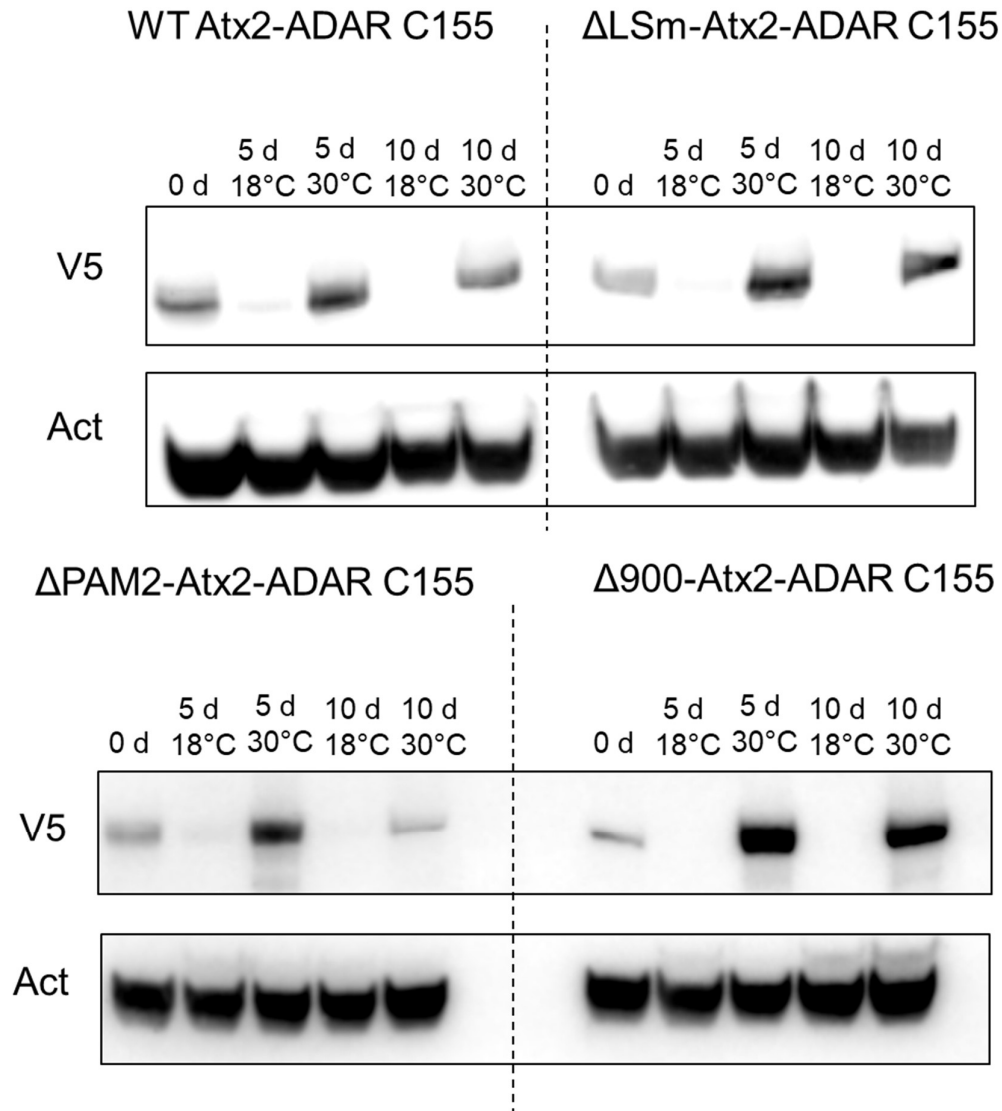


Figure 4.28: Fly head western blots to compare protein expression levels/ensure temperature-dependent expression. Western blots from *elav-Gal4 ;TubGal80^{ts}* fly heads were carried out to ensure expression levels of the different constructs were similar, and expression was indeed occurring in a temperature dependent fashion. Flies aged at 30 °C exhibited V5-tagged construct expression while flies grown at 18°C did not. Some larval expression was also observed

4.3 Discussion

The results presented in this chapter provide various insights regarding the function of Ataxin-2 in both healthy and diseased states. Some of these are demonstrated clearly in the aforementioned experiments, others are promising initial findings that require significant more work. Briefly: this study provides evidence for a model in which Atxn2 may modulate different translational states of its target mRNAs. It also shows that, in diseased states, different domains of Atxn2 exhibit different activities that can either promote or protect against NDs; this means that therapeutic targeting of specific protein-protein interactions mediated by Atxn2 may be more beneficial than simply reducing total levels of the protein. To this purpose, the point mutations identified that inhibit the interactions of the Atxn2 structured domains offer highly specific potential pharmacological targets. This work also presents a novel mechanism of mRNP granule assembly. Finally, information contained in the sequence of Atxn2 IDRs from different organisms is shown to be sufficient to alter the protein's condensation competency in cells.

4.3.1 Molecular mechanisms of Atxn2 function

Various RBPs are capable of remaining bound to mRNAs across different stages of their life cycle, namely RNA processing, transport, translation and translational control (Formicola *et al*, 2019; Gomes & Shorter, 2019; Hachet & Ephrussi, 2004; Harlen & Churchman, 2017; Lin *et al.*, 2015; Maniatis & Reed, 2002). Atxn2 has the ability to exert a variety of different effects on the mRNA transcripts it is found associated with: it is a translational activator of *period (per)* mRNA in *Drosophila*, can repress many miRNA reporters and plays a role in driving neuronal mRNP and SG condensation, while also having a generic stabilizing effect on most of its target mRNAs (Bakthavachalu *et al.*, 2018; Inagaki *et al.*, 2020; Lim & Allada, 2013; McCann *et al*, 2011; Nonhoff *et al.*, 2007; Sudhakaran *et al.*, 2014; Yokoshi *et al*, 2014; Zhang *et al*, 2013). This could imply either that Atxn2 can exert different functions depending on the specific mRNA transcript it interacts with, or, through specific interactions of its different domains, Atxn2 could act sequentially on the same mRNA transcript modulating mRNP remodelling at different stages of its life cycle.

It has been previously demonstrated that Atx2 can act as a translational activator of *per* mRNA by interacting with the proteins LSM12 and Twenty-four (TYF) through its LSM domain (Lim & Allada, 2013; Zhang *et al.*, 2013). Interpreting the current work, alongside the body of evidence presented here for Atxn2 binding LSM12 through a canonical LSM:LSM interaction, we suggest that this interaction occurs in translating poly-ribosomes and increase translation efficiency (Satterfield & Pallanck, 2006). This is consistent with

previous work showing that the Atxn2 LSM domain opposes the formation of mRNP granules, in which mRNAs are typically translationally repressed (Singh *et al.*, 2021).

It must be noted though, that the Atxn2 LSM domain might also have LSM12 independent functions as Atx2 LSM deletions are lethal in *Drosophila* while LSM12 null mutants are viable and fertile, although arrhythmic (Bakthavachalu *et al.*, 2018; Lee *et al.*, 2017). A possibility could be that the Atxn2 LSM domain may also interact with the DEAD-box helicase Me31B/DDX6, which forms a translational repressor complex that excludes LSM12 (Brandmann *et al.*, 2018; Lee *et al.*, 2017). In this framework, we suggest that Atxn2 acts as a translational activator when associated through a LSM:LSM interaction with LSM12. Association of Me31b/DDX6 in the complex occurs only if LSM12 is not present (or may itself somehow cause dissociation of LSM12) causing the Atxn2-bound mRNA to transition to a translationally repressed state. While there is evidence of direct/indirect binding of Me31b/DDX6 to Atxn2, the structural basis for this remains unclear, and may involve the Atxn2 IDRs.

A further point addressed is how Atxn2 PAM2 interactions with PABP favour translational activation. The co-IPs showing that PABP is pulled down by mini Atxn2 indicates that the Atxn2 PAM2:PABP association occurs before/independently of mRNP granule formation. Recent work showing that the PABP:PAM2 interaction is antagonistic to Atxn2 condensation (Boeynaems *et al.*, 2021) further supports the hypothesis that the PAM2 motif binds PABP in translating mRNAs, acting as an enhancer of translation. PABP is also known to associate with translational repressors that drive mRNA de-adenylation and/or storage (Machida *et al.*, 2018; Yoshida *et al.*, 2006). The data presented is consistent with this double role of Atxn2 in association with PABP. Initially when the mRNA is associated with me31B/DDX6, Atxn2 target mRNAs bound to PABP may recruit deadenylases and become translationally dormant or shuttled towards degradation pathways (Lee *et al.*, 2017; Machida *et al.*, 2018; Yi *et al.*, 2018; Yoshida *et al.*, 2006). Then Atxn2-bound mRNAs can transit to mRNP granules, which are formed also thanks to Atxn2 IDR condensation, the running hypothesis being that the mRNAs in these condensates is stored in a degradation-resistant form and may re-enter the pool of translating mRNAs.

4.3.2 Atxn2 as a therapeutic target

Treatments targeting Atxn2, in the form of antisense oligonucleotide (ASO) therapeutics, are currently being developed as treatments for ALS and SCA2 (Becker *et al.*, 2017; Biogen, 2021; Scoles *et al.*, 2017). The work presented in this chapter suggest that targeting certain domains (and indeed specific residues) of Atxn2 may be a more effective strategy than aiming to lower overall levels of the protein. Maintaining normal function of

the LSm domain while inhibiting IDR-mediated condensation and possibly PAM2:PABP interactions might be the ideal approach.

Previous work from the laboratory demonstrated that Atx2 mutants that lacked the cIDR required for granule formation were resistant to ND in *Drosophila* disease models (Bakthavachalu *et al.*, 2018; Huelsmeier *et al.*, 2021; Singh *et al.*, 2021). It was also shown that the LSm domain antagonizes granule formation (Singh *et al.*, 2021). In this chapter this observation was advanced by showing that Atx2 mutants in which the LSm domain was deleted exhibit stronger cytotoxic effects than WT or IDR deleted forms of Atx2 (**Figure 4.27**). The mechanisms through which this protective effect is exerted have not yet been fully elucidated but understanding the nature of the Atxn2 LSm:LSM12 binding (**Figures 4.20 – 4.21**) is a good starting point for forming plausible hypotheses. A lot more work is required to obtain a full picture of the upstream and downstream events around this association. This data supports a previously presented model where mRNP condensation efficiency is correlated with the speed and severity of ND in *Drosophila*. The relationship between SG formation and toxicity is nuanced and complicated. On one hand, SG formation is an essential step of the ISR, in which they play a central role in translational control, modulating levels of both housekeeping proteins and proteins involved in handling cellular stressors (Ivanov *et al.*, 2019). On the other hand, phase-separated condensates are by definition thought to be hubs capable of seeding protein aggregation for purely physical reasons, such as the slim energy threshold between LLPS and liquid-to-solid transition and the high local concentration of IDR-containing proteins with a tendency to self-associate (Li *et al.*, 2013; Wolozin & Ivanov, 2019). The results obtained in this work suggest that, at least in the case of Atxn2-positive RNP granules, toxicity is uncoupled from SG formation itself but rather related to the granule components recruited.

The Atxn2 PAM2 motif had been previously identified to promote degeneration through experiments showing that heterologous expression of human Atxn2 where the PAM2 domain had also been deleted suppressed cytotoxicity in *Drosophila* expressing human TDP-43, while full length mutant Atxn2 did not (Kim *et al.*, 2014). The observation shown here that expression of Atx2 Δ PAM2 is less toxic than WT Atx2 is consistent with this data. The protein TDP-43 can be a driver of pathology in both familial and sporadic ALS: its presence in hyper-phosphorylated, ubiquitinated cytoplasmic inclusion bodies is a hallmark of the disease (Kim *et al.*, 2020; Neumann *et al.*, 2006; Tziortzouda *et al.*, 2021). TDP-43 is primarily a nuclear protein, and mislocalization to the cytoplasm is potentially a central determinant of disease progression. While not directly explored in this work, this idea is consistent with the findings presented, which suggest a model for which:

- (i) a mutation-truncation in a localization sequence causes a disease related protein (for example TDP-43) to be recruited into regions of the cell not related to its normal function.
- (ii) This triggers the ISR, which causes SGs to form as well as activating the UPS.
- (iii) Due to IDRs interacting with RBPs or RNA transcripts themselves, the mislocalized protein could be recruited to SGs o.
- (iv) The physical environment of these LLPS structures favours the formation of insoluble aggregates of proteins that have this tendency. Disease associated mutations in Atxn2 further exacerbate this process either by interacting directly with mislocalized proteins such as TDP-43 or by perturbing the fluidity of the condensates themselves.

In fact, Atx2 Δ PAM2 was shown to form granules with a different protein composition than the ones formed by WT Atx2. This suggests that composition of the granules formed rather than the formation of granules in itself is the most relevant factor in NDs. So, while LLPS driven by IDRs could be necessary for the formation of different kinds of mRNP granules, granules with different protein components would have a different effect and cause different proteinopathies (De Graeve & Besse, 2018; Vogler *et al*, 2018).

4.3.2 Interactions between structured domains may be the main determinant of mRNP composition.

A large body of evidence exists suggesting that specific molecular interactions contribute to mRNP granule formation, notable examples being the PB-condensation-driving component EDC3 or the SG components G3BP and Caprin (Decker *et al*, 2007; Kedersha *et al*, 2016). *In vitro* evidence has also shown that LLPS is mediated both by specific interactions between structured domains and promiscuous associations between IDRs (Protter *et al*, 2018). The data presented in this chapter shows that the Atxn2 PAM2:PABP interaction is critical for the assembly of Atxn2-containing mRNP granules. This interaction may be key to selecting mRNA targets and protein components that will enter these condensates.

The suggested model entails that Atxn2, guided by PAM2:PABP interactions and LSM:LSM interactions, recruits target mRNAs and associating proteins into translating mRNPs (Satterfield & Pallanck, 2006). If translational arrest occurs (for example if cells are subjected to stress), the LSM:LSM interactions are altered (Lee *et al*, 2017), releasing

transcripts from stalled ribosomes. At this both mRNAs and Atxn2's IDRs contribute to the assembly of these mRNPs into granules. This sequence of events is consistent with:

- (i) the TRIBE data exhibiting a lower number of edits of WT Atx2 target mRNAs by Atx2 Δ PAM2-ADARcd.
- (ii) The fact that Δ PAM2 mini Atx2 or equivalent point mutants are not recruited to SGs.
- (iii) The fact that granules induced by Δ PAM2 Atx2 in S2 cells exhibit an aberrant protein composition and lack some key SG markers.

Furthermore, the observation that Atx2 Δ PAM2-ADARcd edits a large number of mRNAs not edited by WT Atx2-ADARcd shows that the PAM2:PABP interaction also prevents Atx2 from associating with non-target mRNAs.

A possible objection to this model is that mRNA components in SGs can be predicted quite accurately simply based on their size, arguing for a primary role of RNA-RNA interactions in SG assembly (Jain & Vale, 2017; Matheny *et al*, 2021; Van Treeck *et al*, 2018). It must be noted though, that in the experiments presented mechanisms through which mRNAs are selected in SGs are not directly investigated. Rather, the TRIBE data shows how Atx2 targets mRNAs to neuronal granules that exist *in vivo* without any stress being involved, while the microscopy experiments assess protein components formed in mRNP granules upon Atx2 over-expression. These mRNP granules may differ significantly by those observed upon stress, and could rely primarily on specific protein-protein and specific protein-mRNA interactions.

5. CONCLUSIONS & OUTLOOK

5.1 Conclusions

In the first chapter, a methodology to produce transient oligomeric aggregates of recombinant WT α SN at a physiological concentration was established. A variety of structural and analytical techniques were then used to characterize both their **size and the timeline of early aggregation progression**. The original goal of isolating α SN oligomers using SE-HPLC was unsuccessful, as at these initial stages of aggregation the oligomers were found to be too unstable, quickly reverting to an equilibrium state between the monomer and oligomers produced. The technique instead showed promise for the purpose of monitoring the early steps of aggregation in a time-resolved fashion, **the main advantage of SE-HPLC over other spectroscopy/light scattering based methodologies being its capability to monitor the growth of different protein species as individual components rather than as a population-averaged value**. The preliminary cytotoxicity and electrophysiology experiments that were carried out showed strong physiological effects, suggesting the oligomers **produced may indeed be relevant in disease progression**. This is in support of a growing body of evidence suggesting that – at the very least – **the threshold delimitating the minimal cytotoxic components of aggregating α SN appears very early in the aggregation process, ranging from low to medium M_w oligomers and never surpassing proto-fibrillar size** (Chen *et al.*, 2015; Danzer *et al.*, 2007; Fusco *et al.*, 2017; Rösener *et al.*, 2020). Instead the aggregates mostly observed post-mortem in LBs (hyper-phosphorylated, ubiquitinated amyloid fibrils) may be the end-result of protective proteostatic pathways rather than the central drivers of disease. The reversibility observed at these early stages of aggregation is in agreement with the standard model of amyloid formation kinetics known seeding-nucleation model (Chiti & Dobson, 2017) which implies an **initial, slow, possibly reversible rate determining step in which oligomeric seeds are formed**, followed by a cascade of monomeric/oligomeric units adding to this seed in a rapid, kinetically favourable manner. This idea forms part of the broader **permissive templating model** (Hardy, 2005), according to which the initial rate-determining step is concentration-dependant, but once seeding has occurred concentration is not important anymore for the conformational spread. The main conceptual advantage of this model is that it gives a solid theoretical foundation to the **prion hypothesis** of pathology spread in neurodegenerative diseases (Prusiner, 1998), according to which misfolded “seeds” (by definition soluble, typically thought to be oligomers) are capable of spreading from cell to cell. Due to permissive

templating, even low concentrations of misfolded seeds in the receive cell is enough to cause a conformational shift and initiate aggregation in the receiver cell.

While in many ways very preliminarily, the electrophysiology data hints at possible mechanisms of entry and toxicity of the α SN oligomers produced. The general consensus is that either α SN oligomers are capable of **spreading by punching holes in cellular membranes** (Danzon *et al.*, 2007; Fusco *et al.*, 2017), or they are capable of transmission in **a receptor-mediated manner** (Urrea *et al.*, 2017). Putatively, the increase in baseline synaptic transmission observed during the *in vivo* electrophysiology experiments, could be explained as **a spike in neurotransmitter release due to synaptic transmission activation** from the oligomer-receptor interaction, in line with reports tying α SN oligomers to NMDA receptor activation and Glutamate release (Diógenes *et al.*, 2012; Durante *et al.*, 2019; Trudler *et al.*, 2021). It also supports the idea of a phase-separation/synaptic vesicle cycling native function of α SN, where aggregation occurring at the synaptic terminal may impair synaptic transmission and lead to a dying-back neurodegeneration effect (Bridi & Hirth, 2018).

In the second chapter, significant progress was made towards the understanding of both the native function of Atxn2 and the role it may play in disease. Co-localization and co-IP experiments carried out with the minimized Atxn2 constructs showed that **the structured domains are necessary and sufficient for interaction with the Atxn2 binding partners PABP and LSM12**. Taken together with the differences in mRNA targets observed in the TRIBE experiments and the altered protein compositions observed in full length Atxn2 Δ PAM2 mRNP granules, this shows the central role that these structured domains play in the native functions of Atxn2. The effect of perturbing these interactions is probably not limited to the composition of mRNP granules, but is most likely generally true even in the Atxn2-mRNA interactions occurring outside of mRNP granules. The neurodegeneration experiments in *Drosophila* showed that **the presence of granules in itself is not necessarily a driver of disease, while the protein-mRNA composition of the granules themselves is**. This makes sense if aggregation is imagined as an outcome of LLPS: A liquid-to-solid transition in a liquid-liquid-phase-separated condensate containing a variety of different protein components driving phase separation could be inhibited by perturbing the composition of these condensates (*i.e.* by not allowing some LLPS driving components from being recruited to the condensate). The precise role of SGs in neurodegeneration itself remains unclear. While the Atxn2 IDRs are certainly part of the cohorts of proteins driving SG condensation, the central element to this protein interactome network has been shown to be the protein G3BP (Yang *et al.*, 2020), and SGs can form even in the absence of Atxn2. Atxn2-induce toxicity is nearly certainly partially decoupled from the formation of

SGs. The identification of specific point mutations capable of inhibiting these interactions show that (i) the associations are indeed occurring in the predicted manner, (ii) are almost certainly direct binding interactions and (iii) are potential pharmacological targets to inhibit the specific functions associated to the binding of these protein targets. The preliminary work carried on the IDRs shows that their function, although being mainly non-specific and physical in nature (*i.e.* Driving LLPS), is not fully sequence independent, **with different IDRs having different condensation-inducing potency**. Overall, the experiments are consistent with a model in which Atxn2 can play different roles during different stages of its target mRNAs life cycle.

5.2 Outlook

As average life expectancy grows longer and longer, so does the global incidence of neurodegenerative diseases. Onset of these diseases typically occurs after the average age of reproduction for our species, and are thus mostly unaffected by natural selection. These crippling diseases mostly result in a heavy reduction of quality of life, a massive economic burden and a premature death. Despite this ever-increasing chunk of the population of the World being affected, available treatments for most NDs are few, not very effective and often aimed merely at managing the symptoms rather than actually curing the disease. In this context, it is increasingly important to find potential therapeutic targets for these conditions.

An enormous body of evidence exists demonstrating the correlation between amyloid aggregation and disease. Amyloid aggregates have been shown to be present in diseases such as AD, PD, HT, ALS, SCAs and many more (Chen *et al.*, 2015; Cline *et al.*, 2018; Laffita-Mesa *et al.*, 2021; Neumann *et al.*, 2006; Recasens & Dehay, 2014; Taylor *et al.*, 2016; Tziortzouda *et al.*, 2021; Vogler *et al.*, 2018; Walsh *et al.*, 2002). A further impressive example helping establish a causal link between amyloid aggregation and disease progression is a series of experiments in which an artificial protein capable of forming amyloid aggregates was designed from first principles, based on the sequences of known proteins involved in pathology. Interestingly, not only did the artificial protein form cytotoxic aggregates, but was also shown to recruit a variety of other proteins, further evidence to the promiscuous nature of IDR-IDR interactions and the many ways they can malfunction in disease states (Olzscha *et al.*, 2011). From the study of the oligomers of α SN, preliminary cytotoxicity and electrophysiology experiments showed a very potent effect. Further cell-based and *in vivo* experiments are required to fully verify this effect and better understand their mechanism of action. But if these species are actually mediators (either directly or indirectly) of pathology, they could be promising pharmacological targets. In this scenario,

the possibility of isolating the pathogenic species would be of high-interest, because it would permit further *in vitro* structural studies and consequent drug design. An interesting future step would be to use the methods outlined (DOSY NMR and SE-HPLC) to investigate if their oligomer size or aggregation kinetics varies compared to wt α SN. Mutant E83Q in particular would be interesting as it has recently been shown to aggregate much more rapidly than wt α SN (Kumar *et al*, 2022), so a different aggregation profile might be detectable using either technique.

The study of Atxn2 is significant in a variety of different ways. Firstly, it provides insight into the complicated and poorly understood native function of Atxn2. That different domains of Atxn2 have the ability to mediate different mRNA processing functions is instead well established and accepted (Lee *et al.*, 2017). The data obtained in this work adds nuance to this statement, as it is also consistent with the idea that Atxn2 may have different binding partners at different times *while interacting with a single mRNA transcript*. This would imply a more central function in translational regulation of its target mRNAs than previously envisioned. This is consistent with the previous finding of the laboratory that deletion of WT Atx2 or of any of its structured domains is lethal in flies (Bakthavachalu *et al.*, 2018) even though deletion of the cIDR (necessary for Atx2 mRNP granule condensation) is not. The results of the IDR domain swaps, although very preliminary, already give some insight into the evolutionary history of Atxn2, as the function of insect Atx2 appears to have been partitioned in Atxn2 and Atxn2L in vertebrates, with Atxn2L potentially playing a stronger role in condensation. Interestingly, less work has been carried out so far on Atxn2L than on Atxn2, as it does not contain the poly-Q repeats implicated in ND progression. The point mutations in the PAM2 region could be an initial step in potential therapeutic intervention, as the PAM2:PABP interaction has been shown to be a possible disease driver, while further experiments using the LSm point mutations could help fully elucidate the role of the LSm domain in translational activation.

Overall this work advances our understanding of the molecular mechanisms directing Atxn2 function in both healthy and diseased states, and opens several novel, specific routes of enquiry that may lead to more targeted therapeutic approaches for devastating diseases like ALS and SCA2.

6. SUPPLEMENTARY INFORMATION

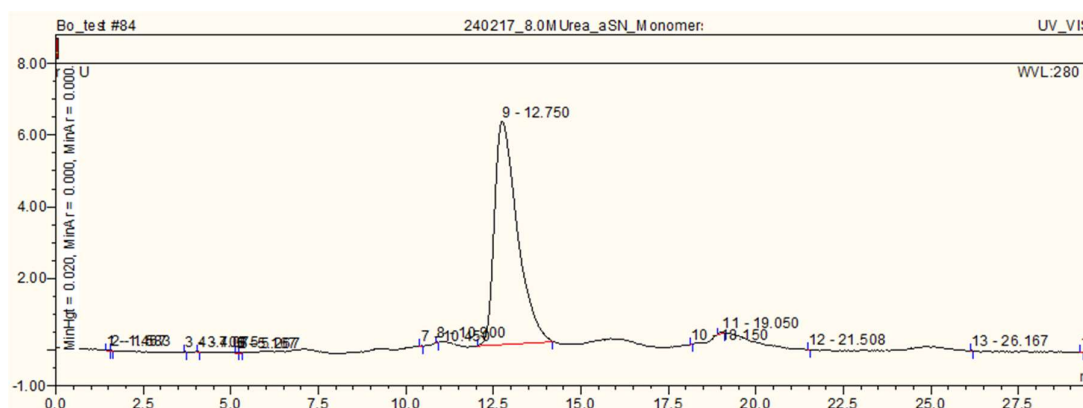


Figure 6.1: SE-HPLC chromatogram of α SN freshly dissolved in 8 M Urea. The α SN sample was injected using 8.0 M urea as a mobile phase. Retention time and shape appeared similar to what observed in phosphate buffer, although the estimated size was slightly larger (58.8 Å), suggesting some residual structure is conserved in native α SN.

Table 6.1: Hydrodynamic radi and elution volumes of standard curve proteins as obtain upon injection using 8 M urea as a mobile phase

Protein	Number of residues	R_H (Å)	Elution volume (mL)	Source of R_H value
Thyroglobulin	2769	262.8	2.18	(Kohn <i>et al.</i> , 2004)
γ -globulin	1364	78	2.47	(Nag <i>et al.</i> , 2015)
Human serum albumin	603	81.9	2.68	(White <i>et al.</i> , 2015)
β -Lactoglobulin	162	49.5	3.29	(Kohn <i>et al.</i> , 2004)
HEWL	129	34.6	4.07	(Wilkins <i>et al.</i> , 1999)
Cytochrome C	100	32.6	3.54	(Wilkins <i>et al.</i> , 1999)

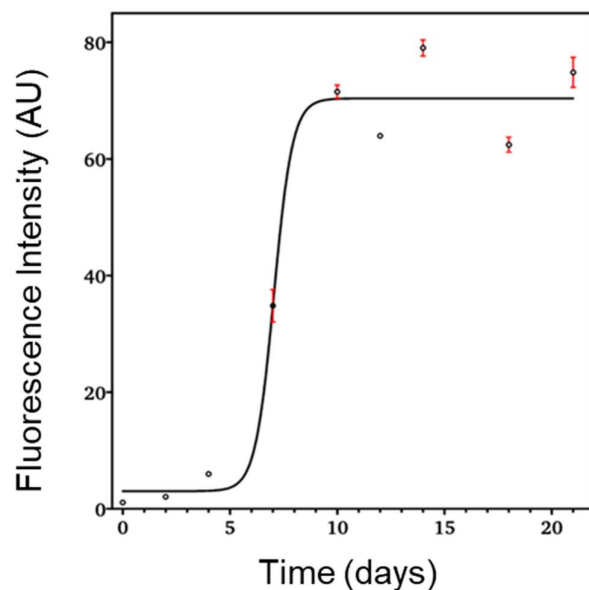


Figure 6.2: HEWL aggregation progression monitored using thioflavin T fluorescence spectrophotometry. Initial assessment of HEWL fibrillation monitored using thioflavin T spectrophotometry (n = 3). Fibrillation was induced by incubating HEWL at high concentration at low pH and high temperature.

Table 6.2: Mapping percentages of TRIBE replicates

S.NO	Sample Name	Million reads (Paired end)		Total Number of Million reads	% >= Q30	Mean Quality Score	Mapping Percentage Read 1	Mapping Percentage Read 2
		Read 1	Read 2					
1	Atx2ΔPAM2-ADARcd rep_1	23.75	23.75	47	97.04	38.26	92.20%	92.10%
2	Atx2ΔPAM2-ADARcd rep_2	22.66	22.66	45	97.03	38.26	91.30%	91.20%

Table 6.3: Targets obtained from TRIBE experiments

Chr	Start	End	Genes	Replicate 1 edit percentage	Replicate 2 edit percentage	Average edit percentage	Chr_coordinate
chr2L	1009080	1009081	IA-2	17.1	15.5	16.3	chr2L_1009081_IA-2
chr2L	1009440	1009441	IA-2	31.7	38.8	35.25	chr2L_1009441_IA-2
chr2L	1011482	1011483	IA-2	31.7	34.7	33.2	chr2L_1011483_IA-2
chr2L	13505104	13505105	B4	22.4	17.48	19.94	chr2L_13505105_B4
chr2L	17188307	17188308	beat-IIIc	18	21.7	19.85	chr2L_17188308_beat-IIIc
chr2L	19746161	19746162	CG10631	25.65	19.4	22.52	chr2L_19746162_CG10631
chr2L	20056018	20056019	sNPF	16	24.3	20.15	chr2L_20056019_sNPF
chr2L	7891375	7891376	Snoo	18.5	16.7	17.6	chr2L_7891376_Snoo
chr2L	8113160	8113161	Bsg	17.4	19.05	18.23	chr2L_8113161_Bsg
chr2L	8113894	8113895	Bsg	16.82	28.7	22.76	chr2L_8113895_Bsg
chr2L	8116005	8116006	Bsg	18.2	27.4	22.8	chr2L_8116006_Bsg
chr2L	9256901	9256902	Ggamma30A	19.5	32.6	26.05	chr2L_9256902_Ggamma30A
chr2L	9292530	9292531	Ggamma30A	42	54.4	48.2	chr2L_9292531_Ggamma30A
chr2L	9292829	9292830	Ggamma30A	42.6	53.55	48.08	chr2L_9292830_Ggamma30A
chr2L	9295031	9295032	Ggamma30A	23.5	26	24.75	chr2L_9295032_Ggamma30A
chr2R	12085463	12085464	jeb	16	28.4	22.2	chr2R_12085464_jeb
chr2R	13513577	13513578	Vmat	64.2	63.9	64.05	chr2R_13513578_Vmat
chr2R	13513654	13513655	Vmat	27.7	33.45	30.58	chr2R_13513655_Vmat
chr2R	13513732	13513733	Vmat	24	29.75	26.88	chr2R_13513733_Vmat
chr2R	13513762	13513763	Vmat	44.55	47.3	45.92	chr2R_13513763_Vmat
chr2R	13514350	13514351	Vmat	37.7	38.5	38.1	chr2R_13514351_Vmat
chr2R	13519080	13519081	Vmat	18.1	20.8	19.45	chr2R_13519081_Vmat
chr2R	23690701	23690702	Pal2	17.3	17.2	17.25	chr2R_23690702_Pal2
chr2R	24213943	24213944	CG30419	27.8	22.98	25.39	chr2R_24213944_CG30419
chr2R	24214234	24214235	CG30419	24.2	26.1	25.15	chr2R_24214235_CG30419
chr2R	24214648	24214649	CG30419	16.4	15.9	16.15	chr2R_24214649_CG30419
chr2R	24215193	24215194	CG30419	43.5	29.7	36.6	chr2R_24215194_CG30419
chr2R	24229341	24229342	CG30419	20.8	16.3	18.55	chr2R_24229342_CG30419
chr2R	6914804	6914805	CG30158	16.4	16	16.2	chr2R_6914805_CG30158
chr2R	6920815	6920816	CG30158	33.92	40.48	37.2	chr2R_6920816_CG30158
chr2R	6921435	6921436	CG30158	17.5	41.9	29.7	chr2R_6921436_CG30158
chr2R	6921976	6921977	CG30158	36	40.2	38.1	chr2R_6921977_CG30158
chr2R	7718070	7718071	CG18812	15.9	19.7	17.8	chr2R_7718071_CG18812
chr2R	9473772	9473773	Camta	21.75	30.18	25.96	chr2R_9473773_Camta
chr2R	9479421	9479422	Camta	23.1	26	24.55	chr2R_9479422_Camta
chr2R	9480019	9480020	Camta	32	41	36.5	chr2R_9480020_Camta

chr2R	9910328	9910329	FMRFa	42.5	25	33.75	chr2R_9910329_FMRFa
chr3L	11498209	11498210	chrb	15.82	20.5	18.16	chr3L_11498210_chrb
chr3L	12267895	12267896	CG32100	20	16.1	18.05	chr3L_12267896_CG32100
chr3L	1504295	1504296	Psa	18.6	15.9	17.25	chr3L_1504296_Psa
chr3L	1521650	1521651	Psa	18.5	34.6	26.55	chr3L_1521651_Psa
chr3L	1543675	1543676	CG7852	15.5	22	18.75	chr3L_1543676_CG7852
chr3L	17062355	17062356	Rbp6	15.2	18.5	16.85	chr3L_17062356_Rbp6
chr3L	17147382	17147383	Rbp6	18.23	26.62	22.42	chr3L_17147383_Rbp6
chr3L	17345219	17345220	Mip	29.9	27.2	28.55	chr3L_17345220_Mip
chr3L	17345290	17345291	Mip	17.9	16.9	17.4	chr3L_17345291_Mip
chr3L	19066983	19066984	Mkp3	18.9	15.7	17.3	chr3L_19066984_Mkp3
chr3L	21494821	21494822	Hr78	27.1	18.4	22.75	chr3L_21494822_Hr78
chr3L	21831417	21831418	CG7148	15.8	28.6	22.2	chr3L_21831418_CG7148
chr3L	21930851	21930852	mub	15	17.2	16.1	chr3L_21930852_mub
chr3L	21931110	21931111	mub	15.1	20.1	17.6	chr3L_21931111_mub
chr3L	22061206	22061207	Oct-TyrR	15.4	16.7	16.05	chr3L_22061207_Oct-TyrR
chr3L	22877661	22877662	Chro	35.3	45.7	40.5	chr3L_22877662_Chro
chr3L	23148124	23148125	CG32350	27.4	35.3	31.35	chr3L_23148125_CG32350
chr3L	23747549	23747550	CG17698	40.27	28.95	34.61	chr3L_23747550_CG17698
chr3L	23934990	23934991	CG40470	23.5	22.05	22.77	chr3L_23934991_CG40470
chr3L	3910071	3910072	Eip63F-1	20	24.5	22.25	chr3L_3910072_Eip63F-1
chr3L	3954338	3954339	CG12605	35.2	40.8	38	chr3L_3954339_CG12605
chr3L	3954933	3954934	CG12605	35.42	51.1	43.26	chr3L_3954934_CG12605
chr3L	3957068	3957069	CG12605	21.8	26.7	24.25	chr3L_3957069_CG12605
chr3L	3957671	3957672	CG12605	18.9	22.2	20.55	chr3L_3957672_CG12605
chr3L	3961590	3961591	CG12605	18.3	18.5	18.4	chr3L_3961591_CG12605
chr3L	3992789	3992790	scrt	21.08	27.28	24.18	chr3L_3992790_scrt
chr3L	4092142	4092143	CG14989	15.88	18.6	17.24	chr3L_4092143_CG14989
chr3L	4113123	4113124	Ack	18.4	16.7	17.55	chr3L_4113124_Ack
chr3L	4113297	4113298	Ack	18.4	17.9	18.15	chr3L_4113298_Ack
chr3L	572527	572528	hipk	29.8	31	30.4	chr3L_572528_hipk
chr3L	572530	572531	hipk	42.7	43	42.85	chr3L_572531_hipk
chr3L	575712	575713	hipk	36.12	52.92	44.52	chr3L_575713_hipk

chr3L	575753	575754	hipk	31.7	49.6	40.65	chr3L_575754_hipk
chr3L	576730	576731	hipk	17.1	15.2	16.15	chr3L_576731_hipk
chr3L	577020	577021	hipk	20	16.9	18.45	chr3L_577021_hipk
chr3L	577417	577418	hipk	51.42	62.12	56.77	chr3L_577418_hipk
chr3L	577970	577971	hipk	21.4	25.9	23.65	chr3L_577971_hipk
chr3L	578186	578187	hipk	18.2	24.9	21.55	chr3L_578187_hipk
chr3L	578453	578454	hipk	25.7	32.7	29.2	chr3L_578454_hipk
chr3L	579335	579336	hipk	18.1	17	17.55	chr3L_579336_hipk
chr3L	579634	579635	hipk	20.5	19.8	20.15	chr3L_579635_hipk
chr3L	580103	580104	hipk	63.1	67.7	65.4	chr3L_580104_hipk
chr3L	580500	580501	hipk	15.6	18.1	16.85	chr3L_580501_hipk
chr3L	580932	580933	hipk	17.9	24.1	21	chr3L_580933_hipk
chr3L	8970787	8970788	CG5026	17.9	16.7	17.3	chr3L_8970788_CG5026
chr3L	8993382	8993383	smg	25.9	15.3	20.6	chr3L_8993383_smg
chr3L	9074752	9074753	Tequila	29.8	19.5	24.65	chr3L_9074753_Tequila
chr3L	9103274	9103275	bol	23.9	26.3	25.1	chr3L_9103275_bol
chr3L	9136454	9136455	Use1	25	16	20.5	chr3L_9136455_Use1
chr3L	9669496	9669497	fry	18.2	18.2	18.2	chr3L_9669497_fry
chr3L	9945905	9945906	CG34356	19.4	32.1	25.75	chr3L_9945906_CG34356
chr3R	10158991	10158992	Invadolysin	17.9	18.2	18.05	chr3R_10158992_Invadolysin
chr3R	10862820	10862821	CG6574	20.8	33.3	27.05	chr3R_10862821_CG6574
chr3R	10877257	10877258	CR45195	32.7	19.5	26.1	chr3R_10877258_CR45195
chr3R	10889941	10889942	Leash	18.2	18.8	18.5	chr3R_10889942_Leash
chr3R	13224637	13224638	Ace	24.8	33.3	29.05	chr3R_13224638_Ace
chr3R	13227870	13227871	Ace	26.6	35.3	30.95	chr3R_13227871_Ace
chr3R	13227970	13227971	Ace	33.42	47.95	40.69	chr3R_13227971_Ace
chr3R	14330745	14330746	NK7.1	16.2	16.7	16.45	chr3R_14330746_NK7.1
chr3R	14660839	14660840	Hexim	17.1	20	18.55	chr3R_14660840_Hexim
chr3R	14669984	14669985	Meltrin	16.95	20.7	18.82	chr3R_14669985_Meltrin
chr3R	14746335	14746336	jvl	36.4	20.9	28.65	chr3R_14746336_jvl
chr3R	14746546	14746547	smp-30	45.1	24	34.55	chr3R_14746547_smp-30
chr3R	14804349	14804350	btsz	16.5	26.5	21.5	chr3R_14804350_btsz
chr3R	15255687	15255688	CG42404	18.2	17.6	17.9	chr3R_15255688_CG42404
chr3R	15356475	15356476	Atg4b	15.4	20.7	18.05	chr3R_15356476_At4b
chr3R	15414730	15414731	Atx2	22.82	26.25	24.54	chr3R_15414731_At2
chr3R	15414731	15414732	Atx2	19	22.6	20.8	chr3R_15414732_At2
chr3R	15849417	15849418	cv-d	18.8	22.2	20.5	chr3R_15849418_cv-d
chr3R	16611267	16611268	NPF	22.98	23.52	23.25	chr3R_16611268_NPF
chr3R	16645743	16645744	CG10324	26.1	35	30.55	chr3R_16645744_CG10324

chr3R	17090737	17090738	cal1	25	15.4	20.2	chr3R_17090738_cal1
chr3R	17731884	17731885	Lgr1	20	19	19.5	chr3R_17731885_Lgr1
chr3R	17802763	17802764	CG17806	20.7	18.8	19.75	chr3R_17802764_CG17806
chr3R	19155252	19155253	CG11779	15.6	19.1	17.35	chr3R_19155253_CG11779
chr3R	20783186	20783187	Syp	21.1	17	19.05	chr3R_20783187_Syp
chr3R	20797469	20797470	Syp	15.3	20	17.65	chr3R_20797470_Syp
chr3R	20820785	20820786	CG17271	31.2	19.75	25.48	chr3R_20820786_CG17271
chr3R	20862212	20862213	CG3822	17	26.3	21.65	chr3R_20862213_CG3822
chr3R	20992668	20992669	Calx	17.8	27.5	22.65	chr3R_20992669_Calx
chr3R	21213891	21213892	SNF4Agamm a	17.2	20	18.6	chr3R_21213892_SNF4Agamm a
chr3R	21354284	21354285	mod(mdg4)	15	19	17	chr3R_21354285_mod(mdg4)
chr3R	21527750	21527751	CG7956	30.3	26.5	28.4	chr3R_21527751_CG7956
chr3R	21625731	21625732	E2f	18	23.7	20.85	chr3R_21625732_E2f
chr3R	23261698	23261699	orb	15.2	16.1	15.65	chr3R_23261699_orb
chr3R	23681818	23681819	eIF-3p66	20.4	21.5	20.95	chr3R_23681819_eIF-3p66
chr3R	23698883	23698884	prr	15.4	23.4	19.4	chr3R_23698884_prr
chr3R	23723418	23723419	CG10365	16.1	21.1	18.6	chr3R_23723419_CG10365
chr3R	23732353	23732354	Rpn9	20.4	16.7	18.55	chr3R_23732354_Rpn9
chr3R	24664388	24664389	slo	23.1	30	26.55	chr3R_24664389_slo
chr3R	24802164	24802165	polybromo	29.2	20.98	25.09	chr3R_24802165_polybromo
chr3R	24820485	24820486	Saf-B	15	22	18.5	chr3R_24820486_Saf-B
chr3R	25234647	25234648	CG10420	22.2	34.5	28.35	chr3R_25234648_CG10420
chr3R	26233920	26233921	CG12290	25.6	18.4	22	chr3R_26233921_CG12290
chr3R	28050111	28050112	CG34362	15.6	25.6	20.6	chr3R_28050112_CG34362
chr3R	28838531	28838532	Apc	20.4	18.8	19.6	chr3R_28838532_Apc
chr3R	29659085	29659086	Dop1R2	15.8	39.5	27.65	chr3R_29659086_Dop1R2
chr3R	29674515	29674516	Bub3	27.6	17.8	22.7	chr3R_29674516_Bub3
chr3R	31457698	31457699	Gprk2	23.5	24.5	24	chr3R_31457699_Gprk2
chr3R	31841367	31841368	RhoGAP100F	17.73	19	18.37	chr3R_31841368_RhoGAP100F
chr3R	5811503	5811504	CG11000	21.3	32.4	26.85	chr3R_5811504_CG11000
chr3R	5811505	5811506	CG11000	20.8	25.7	23.25	chr3R_5811506_CG11000
chr3R	7126383	7126384	CG10098	18.6	19.1	18.85	chr3R_7126384_CG10098
chr3R	8244435	8244436	CG18749	15.4	16.1	15.75	chr3R_8244436_CG18749
chr3R	8244435	8244436	CG33722	15.4	16.1	15.75	chr3R_8244436_CG33722
chr3R	9416762	9416763	alpha-Man-II	30.8	32.3	31.55	chr3R_9416763_alpha-Man-II
chr3R	9441505	9441506	ps	37.92	33.52	35.72	chr3R_9441506_ps
chr3R	9471135	9471136	CG16779	43.05	34.85	38.95	chr3R_9471136_CG16779
chr3R	9525979	9525980	CG8176	16.2	24.4	20.3	chr3R_9525980_CG8176
chr3R	9539464	9539465	mura	21.75	20.05	20.9	chr3R_9539465_mura
chr3R	9794186	9794187	CG8516	19.4	30.8	25.1	chr3R_9794187_CG8516
chr4	478956	478957	Asator	15.1	20	17.55	chr4_478957_Asator

chr4	532906	532907	zfh2	20	19.7	19.85	chr4_532907_zfh2
chr4	92946	92947	pan	22.5	20.8	21.65	chr4_92947_pan
chrX	10309093	10309094	alpha-Man-I	24.5	29	26.75	chrX_10309094_alpha-Man-I
chrX	12331302	12331303	Ten-a	34.8	34	34.4	chrX_12331303_Ten-a
chrX	16075472	16075473	Tob	23.68	29.32	26.5	chrX_16075473_Tob
chrX	16075887	16075888	Tob	22.2	22	22.1	chrX_16075888_Tob
chrX	16075888	16075889	Tob	21.6	24	22.8	chrX_16075889_Tob
chrX	16076959	16076960	Tob	20	23.3	21.65	chrX_16076960_Tob
chrX	16077193	16077194	Tob	17.85	19.15	18.5	chrX_16077194_Tob
chrX	16089317	16089318	Tob	15.2	16.1	15.65	chrX_16089318_Tob
chrX	3321068	3321069	dnc	22.7	28.9	25.8	chrX_3321069_dnc
chrX	3342369	3342370	dnc	17.4	22.2	19.8	chrX_3342370_dnc
chrX	6325433	6325434	CG15894	32	20.8	26.4	chrX_6325434_CG15894
chrX	9172798	9172799	mei-P26	17.3	23.1	20.2	chrX_9172799_mei-P26
chrX	9179452	9179453	mei-P26	17.35	22.23	19.79	chrX_9179453_mei-P26
chrX	9188891	9188892	mei-P26	15.8	18.4	17.1	chrX_9188892_mei-P26

7. BIBLIOGRAPHY

Alberti S, Dormann D (2019) Liquid–Liquid Phase Separation in Disease. *Annual Review of Genetics* 53: 171-194

Alberti S, Gladfelter A, Mittag T (2019) Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* 176: 419-434

Alberti S, Mateju D, Mediani L, Carra S (2017) Granulostasis: Protein Quality Control of RNP Granules. *Frontiers in Molecular Neuroscience* 10

Albrecht M, Lengauer T (2004) Novel Sm-like proteins with long C-terminal tails and associated methyltransferases. *FEBS Lett* 569: 18-26

Ali MM, Roe SM, Vaughan CK, Meyer P, Panaretou B, Piper PW, Prodromou C, Pearl LH (2006) Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex. *Nature* 440: 1013-1017

Anders S, Pyl PT, Huber W (2015) HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31: 166-169

Anderson P (2010) Post-transcriptional regulons coordinate the initiation and resolution of inflammation. *Nat Rev Immunol* 10: 24-35

Anderson P, Kedersha N (2009) RNA granules: post-transcriptional and epigenetic modulators of gene expression. *Nat Rev Mol Cell Biol* 10: 430-436

Araki K, Yagi N, Nakatani R, Sekiguchi H, So M, Yagi H, Ohta N, Nagai Y, Goto Y, Mochizuki H (2016) A small-angle X-ray scattering study of alpha-synuclein from human red blood cells. *Sci Rep* 6: 30473

Armstrong MJ, Okun MS (2020) Diagnosis and Treatment of Parkinson Disease: A Review. *JAMA* 323: 548-560

Ascherio A, Schwarzschild MA (2016) The epidemiology of Parkinson's disease: risk factors and prevention. *Lancet Neurol* 15: 1257-1272

Aulić S, Masperone L, Narkiewicz J, Isopi E, Bistaffa E, Ambrosetti E, Pastore B, De Cecco E, Scaini D, Zago P *et al* (2017) α -Synuclein Amyloids Hijack Prion Protein to Gain Cell Entry, Facilitate Cell-to-Cell Spreading and Block Prion Replication. *Sci Rep* 7: 10050

Azkanaz M, Rodríguez López A, De Boer B, Huiting W, Angrand P-O, Vellenga E, Kampinga HH, Bergink S, Martens JH, Schuringa JJ *et al* (2019) Protein quality control in the nucleolus safeguards recovery of epigenetic regulators after heat shock. *eLife* 8

Bakthavachalu B, Huelsmeier J, Sudhakaran IP, Hillebrand J, Singh A, Petrauskas A, Thiagarajan D, Sankaranarayanan M, Mizoue L, Anderson EN *et al* (2018) RNP-Granule Assembly via Ataxin-2 Disordered Domains Is Required for Long-Term Memory and Neurodegeneration. *Neuron* 98: 754-766.e754

Banani SF, Lee HO, Hyman AA, Rosen MK (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology* 18: 285-298

Bartels T, Choi JG, Selkoe DJ (2011) α -Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature* 477: 107-110

Bassell GJ, Warren ST (2008) Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron* 60: 201-214

Becker LA, Huang B, Bieri G, Ma R, Knowles DA, Jafar-Nejad P, Messing J, Kim HJ, Soriano A, Auburger G *et al* (2017) Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice. *Nature* 544: 367-371

Bellani S, Sousa VL, Ronzitti G, Valtorta F, Meldolesi J, Chieregatti E (2010) The regulation of synaptic function by alpha-synuclein. *Commun Integr Biol* 3: 106-109

Benbrook DM, Long A (2012) Integration of autophagy, proteasomal degradation, unfolded protein response and apoptosis. *Exp Oncol* 34: 286-297

Biancalana M, Koide S (2010) Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochim Biophys Acta* 1804: 1405-1412

Biogen, 2021. <https://clinicaltrials.gov/ct2/show/NCT04494256>.

Bley N, Lederer M, Pfalz B, Reinke C, Fuchs T, Glaß M, Möller B, Hüttelmaier S (2015) Stress granules are dispensable for mRNA stabilization during cellular stress. *Nucleic Acids Res* 43: e26

Bliss TV, Lomo T (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J Physiol* 232: 331-356

- Boeynaems S, Dorone Y, Marian A, Shabardina V, Huang G, Kim G, Sanyal A, Şen N-E, Docampo R, Ruiz-Trillo I *et al* (2021) Poly(A)-binding protein is an ataxin-2 chaperone that emulsifies biomolecular condensates. *bioRxiv*: 2021.2008.2023.457426
- Boisvert FM, van Koningsbruggen S, Navascués J, Lamond AI (2007) The multifunctional nucleolus. *Nat Rev Mol Cell Biol* 8: 574-585
- Bouhamdani N, Comeau D, Turcotte S (2021) A Compendium of Information on the Lysosome. *Frontiers in Cell and Developmental Biology* 9
- Braak H, Del Tredici K, Rüb U, de Vos RA, Jansen Steur EN, Braak E (2003) Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol Aging* 24: 197-211
- Brandmann T, Fakim H, Padamsi Z, Youn JY, Gingras AC, Fabian MR, Jinek M (2018) Molecular architecture of LSM14 interactions involved in the assembly of mRNA silencing complexes. *Embo j* 37
- Brengues M, Teixeira D, Parker R (2005) Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. *Science* 310: 486-489
- Bridi JC, Hirth F (2018) Mechanisms of α -Synuclein Induced Synaptopathy in Parkinson's Disease. *Frontiers in Neuroscience* 12
- Brodin L, Milovanovic D, Rizzoli SO, Shupliakov O (2022) α -Synuclein in the Synaptic Vesicle Liquid Phase: Active Player or Passive Bystander? *Frontiers in Molecular Biosciences* 9
- Buddika K, Ariyapala IS, Hazuga MA, Riffert D, Sokol NS (2020) Canonical nucleators are dispensable for stress granule assembly in intestinal progenitors. *Journal of Cell Science* 133: jcs243451
- Burgess RR (2018) A brief practical review of size exclusion chromatography: Rules of thumb, limitations, and troubleshooting. *Protein Expr Purif* 150: 81-85
- Burré J (2015) The Synaptic Function of α -Synuclein. *J Parkinsons Dis* 5: 699-713
- Chaari A, Fahy C, Chevillot-Biraud A, Rholam M (2015) Insights into Kinetics of Agitation-Induced Aggregation of Hen Lysozyme under Heat and Acidic Conditions from Various Spectroscopic Methods. *PLoS One* 10: e0142095

- Chan CS, Gertler TS, Surmeier DJ (2010) A molecular basis for the increased vulnerability of substantia nigra dopamine neurons in aging and Parkinson's disease. *Mov Disord* 25 Suppl 1: S63-70
- Chen SW, Drakulic S, Deas E, Ouberaï M, Aprile FA, Arranz R, Ness S, Roodveldt C, Guillems T, De-Genst EJ *et al* (2015) Structural characterization of toxic oligomers that are kinetically trapped during α -synuclein fibril formation. *Proc Natl Acad Sci U S A* 112: E1994-2003
- Chen Y, Ding F, Nie H, Serohijos AW, Sharma S, Wilcox KC, Yin S, Dokholyan NV (2008) Protein folding: then and now. *Arch Biochem Biophys* 469: 4-19
- Chipman David M, Sharon N (1969) Mechanism of Lysozyme Action. *Science* 165: 454-465
- Chiti F, Dobson CM (2017) Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu Rev Biochem* 86: 27-68
- Chou A, Krukowski K, Jopson T, Zhu PJ, Costa-Mattioli M, Walter P, Rosi S (2017) Inhibition of the integrated stress response reverses cognitive deficits after traumatic brain injury. *Proc Natl Acad Sci U S A* 114: E6420-e6426
- Cline EN, Bicca MA, Viola KL, Klein WL (2018) The Amyloid- β Oligomer Hypothesis: Beginning of the Third Decade. *J Alzheimers Dis* 64: S567-s610
- Cougot N, Babajko S, Séraphin B (2004) Cytoplasmic foci are sites of mRNA decay in human cells. *J Cell Biol* 165: 31-40
- Cui X, Wei Z, Zhang L, Liu H, Sun L, Zhang S-W, Huang Y, Meng J (2016) Guitar: An R/Bioconductor Package for Gene Annotation Guided Transcriptomic Analysis of RNA-Related Genomic Features. *Biomed Res Int* 2016: 8367534-8367534
- Damrath E, Heck MV, Gispert S, Azizov M, Nowock J, Seifried C, Rüb U, Walter M, Auburger G (2012) ATXN2-CAG42 Sequesters PABPC1 into Insolubility and Induces FBXW8 in Cerebellum of Old Ataxic Knock-In Mice. *PLoS Genetics* 8: e1002920
- Danzer KM, Haasen D, Karow AR, Moussaud S, Habeck M, Giese A, Kretzschmar H, Hengerer B, Kostka M (2007) Different species of alpha-synuclein oligomers induce calcium influx and seeding. *J Neurosci* 27: 9220-9232

Davey NE, Van Roey K, Weatheritt RJ, Toedt G, Uyar B, Altenberg B, Budd A, Diella F, Dinkel H, Gibson TJ (2012) Attributes of short linear motifs. *Mol Biosyst* 8: 268-281

De Graeve F, Besse F (2018) Neuronal RNP granules: from physiological to pathological assemblies. *Biological Chemistry* 399: 623-635

Decker CJ, Parker R (2012) P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. *Cold Spring Harbor perspectives in biology* 4: a012286-a012286

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

Deo RC, Bonanno JB, Sonenberg N, Burley SK (1999) Recognition of Polyadenylate RNA by the Poly(A)-Binding Protein. *Cell* 98: 835-845

Dettmer U, Newman AJ, Soldner F, Luth ES, Kim NC, von Saucken VE, Sanderson JB, Jaenisch R, Bartels T, Selkoe D (2015) Parkinson-causing α -synuclein missense mutations shift native tetramers to monomers as a mechanism for disease initiation. *Nature Communications* 6: 7314

Dice JF (1990) Peptide sequences that target cytosolic proteins for lysosomal proteolysis. *Trends Biochem Sci* 15: 305-309

Diógenes MJ, Dias RB, Rombo DM, Vicente Miranda H, Maiolino F, Guerreiro P, Näsström T, Franquelim HG, Oliveira LMA, Castanho MARB *et al* (2012) Extracellular Alpha-Synuclein Oligomers Modulate Synaptic Transmission and Impair LTP Via NMDA-Receptor Activation. *The Journal of Neuroscience* 32: 11750

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15-21

Dobson CM (2004) Principles of protein folding, misfolding and aggregation. *Semin Cell Dev Biol* 15: 3-16

Dollins DE, Warren JJ, Immormino RM, Gewirth DT (2007) Structures of GRP94-nucleotide complexes reveal mechanistic differences between the hsp90 chaperones. *Mol Cell* 28: 41-56

Durante V, de Iure A, Loffredo V, Vaikath N, De Risi M, Paciotti S, Quiroga-Varela A, Chiasserini D, Mellone M, Mazzocchetti P *et al* (2019) Alpha-synuclein targets GluN2A NMDA receptor subunit causing striatal synaptic dysfunction and visuospatial memory alteration. *Brain* 142: 1365-1385

Durr A (2010) Autosomal dominant cerebellar ataxias: polyglutamine expansions and beyond. *Lancet Neurol* 9: 885-894

Duvick L, Barnes J, Ebner B, Agrawal S, Andresen M, Lim J, Giesler GJ, Zoghbi HY, Orr HT (2010) SCA1-like disease in mice expressing wild-type ataxin-1 with a serine to aspartic acid replacement at residue 776. *Neuron* 67: 929-935

Edelhoch H (1960) The properties of thyroglobulin. I. The effects of alkali. *J Biol Chem* 235: 1326-1334

Ehrnhoefer DE, Bieschke J, Boeddrich A, Herbst M, Masino L, Lurz R, Engemann S, Pastore A, Wanker EE (2008) EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nat Struct Mol Biol* 15: 558-566

Eisele YS, Obermüller U, Heilbronner G, Baumann F, Kaeser SA, Wolburg H, Walker LC, Staufenbiel M, Heikenwalder M, Jucker M (2010) Peripherally applied Abeta-containing inoculates induce cerebral beta-amyloidosis. *Science* 330: 980-982

El-Naggar AM, Sorensen PH (2018) Translational control of aberrant stress responses as a hallmark of cancer. *J Pathol* 244: 650-666

Elbaum-Garfinkle S, Kim Y, Szczepaniak K, Chen Carlos C-H, Eckmann Christian R, Myong S, Brangwynne Clifford P (2015) The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proceedings of the National Academy of Sciences* 112: 7189-7194

Elden AC, Kim HJ, Hart MP, Chen-Plotkin AS, Johnson BS, Fang X, Armarkola M, Geser F, Greene R, Lu MM *et al* (2010) Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* 466: 1069-1075

Emery P (2007) Protein extraction from Drosophila heads. *Methods Mol Biol* 362: 375-377

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

Faustini G, Bono F, Valerio A, Pizzi M, Spano P, Bellucci A (2017) Mitochondria and α -Synuclein: Friends or Foes in the Pathogenesis of Parkinson's Disease? *Genes (Basel)* 8: 377

Fekete S, Beck A, Veuthey JL, Guilleme D (2014) Theory and practice of size exclusion chromatography for the analysis of protein aggregates. *J Pharm Biomed Anal* 101: 161-173

Fink AL (2006) The aggregation and fibrillation of alpha-synuclein. *Acc Chem Res* 39: 628-634

Finkbeiner S (2020) The Autophagy Lysosomal Pathway and Neurodegeneration. *Cold Spring Harb Perspect Biol* 12

Formicola N, Vijayakumar J, Besse F (2019) Neuronal ribonucleoprotein granules: Dynamic sensors of localized signals. *Traffic* 20: 639-649

Fu H, Hardy J, Duff KE (2018) Selective vulnerability in neurodegenerative diseases. *Nature Neuroscience* 21: 1350-1358

Fusco G, Chen SW, Williamson PTF, Cascella R, Perni M, Jarvis JA, Cecchi C, Vendruscolo M, Chiti F, Cremades N *et al* (2017) Structural basis of membrane disruption and cellular toxicity by α -synuclein oligomers. *Science* 358: 1440-1443

Ganowiak K, Hultman P, Engström U, Gustavsson A, Westermarck P (1994) Fibrils from synthetic amyloid-related peptides enhance development of experimental AA-amyloidosis in mice. *Biochem Biophys Res Commun* 199: 306-312

Giguère N, Burke Nanni S, Trudeau L-E (2018) On Cell Loss and Selective Vulnerability of Neuronal Populations in Parkinson's Disease. *Frontiers in Neurology* 9

Glasscock E, Singhanian A, Tanouye MA (2005) The mei-P26 gene encodes a RING finger B-box coiled-coil-NHL protein that regulates seizure susceptibility in *Drosophila*. *Genetics* 170: 1677-1689

Goedert M (2015) Neurodegeneration. Alzheimer's and Parkinson's diseases: The prion concept in relation to assembled A β , tau, and α -synuclein. *Science* 349: 1255555

Gomes E, Shorter J (2019) The molecular language of membraneless organelles. *J Biol Chem* 294: 7115-7127

Grabbe C, Husnjak K, Dikic I (2011) The spatial and temporal organization of ubiquitin networks. *Nat Rev Mol Cell Biol* 12: 295-307

Hachet O, Ephrussi A (2004) Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization. *Nature* 428: 959-963

Halliday M, Radford H, Zents KAM, Molloy C, Moreno JA, Verity NC, Smith E, Ortori CA, Barrett DA, Bushell M *et al* (2017) Repurposed drugs targeting eIF2 α -P-mediated translational repression prevent neurodegeneration in mice. *Brain* 140: 1768-1783

Harding HP, Zhang Y, Bertolotti A, Zeng H, Ron D (2000) Perk is essential for translational regulation and cell survival during the unfolded protein response. *Mol Cell* 5: 897-904

Hardy J (2005) Expression of normal sequence pathogenic proteins for neurodegenerative disease contributes to disease risk: 'permissive templating' as a general mechanism underlying neurodegeneration. *Biochem Soc Trans* 33: 578-581

Harlen KM, Churchman LS (2017) The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nature Reviews Molecular Cell Biology* 18: 263-273

Harte NP, Klyubin I, McCarthy EK, Min S, Garrahy SA, Xie Y, Davey GP, Boland JJ, Rowan MJ, Mok KH (2015a) Amyloid Oligomers and Mature Fibrils Prepared from an Innocuous Protein Cause Diverging Cellular Death Mechanisms. *J Biol Chem* 290: 28343-28352

Harte NP, Klyubin I, McCarthy EK, Min S, Garrahy SA, Xie Y, Davey GP, Boland JJ, Rowan MJ, Mok KH (2015b) Amyloid Oligomers and Mature Fibrils Prepared from an Innocuous Protein Cause Diverging Cellular Death Mechanisms. *J Biol Chem* 290: 28343-28352

He L, Zhang J, Zhao J, Ma N, Kim SW, Qiao S, Ma X (2018) Autophagy: The Last Defense against Cellular Nutritional Stress. *Advances in Nutrition* 9: 493-504

Heras-Garvin A, Weckbecker D, Ryazanov S, Leonov A, Griesinger C, Giese A, Wenning GK, Stefanova N (2018) Anle138b modulates α -synuclein oligomerization and prevents motor decline and neurodegeneration in a mouse model of multiple system atrophy. *Mov Disord*

Hershko A, Ciechanover A (1998) The ubiquitin system. *Annu Rev Biochem* 67: 425-479

Hetz C (2012) The unfolded protein response: controlling cell fate decisions under ER stress and beyond. *Nature Reviews Molecular Cell Biology* 13: 89-102

Hetz C, Zhang K, Kaufman RJ (2020) Mechanisms, regulation and functions of the unfolded protein response. *Nature Reviews Molecular Cell Biology* 21: 421-438

Hirsch EC, Vyas S, Hunot S (2012) Neuroinflammation in Parkinson's disease. *Parkinsonism Relat Disord* 18 Suppl 1: S210-212

Hochberg-Laufer H, Schwed-Gross A, Neugebauer KM, Shav-Tal Y (2019) Uncoupling of nucleo-cytoplasmic RNA export and localization during stress. *Nucleic Acids Research* 47: 4778-4797

Hollien J (2013) Evolution of the unfolded protein response. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1833: 2458-2463

Hong P, Koza S, Bouvier ES (2012) Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates. *J Liq Chromatogr Relat Technol* 35: 2923-2950

Huang S, Mo X, Wang J, Ye X, Yu H, Liu Y (2022) α -Synuclein phase separation and amyloid aggregation are modulated by C-terminal truncations. *FEBS Letters* 596: 1388-1400

Hubstenberger A, Courel M, Bénard M, Souquere S, Ernoult-Lange M, Chouaib R, Yi Z, Morlot JB, Munier A, Fradet M *et al* (2017) P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. *Mol Cell* 68: 144-157.e145

Huelsmeier J, Walker E, Bakthavachalu B, Ramaswami M (2021) A C-terminal ataxin-2 disordered region promotes Huntingtin protein aggregation and neurodegeneration in *Drosophila* models of Huntington's disease. *G3 (Bethesda)* 11

Hyman AA, Weber CA, Jülicher F (2014) Liquid-Liquid Phase Separation in Biology. *Annual Review of Cell and Developmental Biology* 30: 39-58

Ichimiya T, Yamakawa T, Hirano T, Yokoyama Y, Hayashi Y, Hirayama D, Wagatsuma K, Itoi T, Nakase H (2020) Autophagy and Autophagy-Related Diseases: A Review. *Int J Mol Sci* 21: 8974

- Iliff JJ, Lee H, Yu M, Feng T, Logan J, Nedergaard M, Benveniste H (2013) Brain-wide pathway for waste clearance captured by contrast-enhanced MRI. *J Clin Invest* 123: 1299-1309
- Ilijina M, Garcia GA, Horrocks MH, Tosatto L, Choi ML, Ganzinger KA, Abramov AY, Gandhi S, Wood NW, Cremades N *et al* (2016) Kinetic model of the aggregation of alpha-synuclein provides insights into prion-like spreading. *Proc Natl Acad Sci U S A* 113: E1206-1215
- Inagaki H, Hosoda N, Tsuiji H, Hoshino S-I (2020) Direct evidence that Ataxin-2 is a translational activator mediating cytoplasmic polyadenylation. *Journal of Biological Chemistry* 295: 15810-15825
- Ivanov P, Kedersha N, Anderson P (2019) Stress Granules and Processing Bodies in Translational Control. *Cold Spring Harb Perspect Biol* 11
- Jain A, Vale RD (2017) RNA phase transitions in repeat expansion disorders. *Nature* 546: 243-247
- Jarrett JT, Lansbury PT, Jr. (1993) Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* 73: 1055-1058
- Jessen NA, Munk AS, Lundgaard I, Nedergaard M (2015) The Glymphatic System: A Beginner's Guide. *Neurochem Res* 40: 2583-2599
- Jiang S, Fagman JB, Chen C, Alberti S, Liu B (2020) Protein phase separation and its role in tumorigenesis. *eLife* 9
- Jones JA, Wilkins DK, Smith LJ, Dobson MC (1997) Characterisation of protein unfolding by NMR diffusion measurements. *Journal of Biomolecular NMR*
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A *et al* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596: 583-589
- Jung T, Grune T (2012) Structure of the proteasome. *Prog Mol Biol Transl Sci* 109: 1-39
- Kahvejian A, Svitkin YV, Sukarieh R, M'Boutchou MN, Sonenberg N (2005) Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. *Genes Dev* 19: 104-113

Kandiyal PS, Kim JY, Fortunati DL, Mok KH (2019) Size Determination of Protein Oligomers/Aggregates Using Diffusion NMR Spectroscopy. *Methods Mol Biol* 2039: 173-183

Kaushik S, Cuervo AM (2012) Chaperone-mediated autophagy: a unique way to enter the lysosome world. *Trends Cell Biol* 22: 407-417

Kedersha N, Anderson P (2007) Mammalian stress granules and processing bodies. *Methods Enzymol* 431: 61-81

Kedersha N, Chen S, Gilks N, Li W, Miller IJ, Stahl J, Anderson P (2002) Evidence that ternary complex (eIF2-GTP-tRNA(i)(Met))-deficient preinitiation complexes are core constituents of mammalian stress granules. *Mol Biol Cell* 13: 195-210

Kedersha N, Cho MR, Li W, Yacono PW, Chen S, Gilks N, Golan DE, Anderson P (2000) Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. *J Cell Biol* 151: 1257-1268

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

Kedersha N, Stoecklin G, Ayodele M, Yacono P, Lykke-Andersen J, Fritzler MJ, Scheuner D, Kaufman RJ, Golan DE, Anderson P (2005) Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J Cell Biol* 169: 871-884

Kedersha NL, Gupta M, Li W, Miller I, Anderson P (1999) RNA-Binding Proteins Tia-1 and Tiar Link the Phosphorylation of Eif-2 α to the Assembly of Mammalian Stress Granules. *Journal of Cell Biology* 147: 1431-1442

Kiebler MA, Bassell GJ (2006) Neuronal RNA granules: movers and makers. *Neuron* 51: 685-690

Kim DH, Sarbassov DD, Ali SM, King JE, Latek RR, Erdjument-Bromage H, Tempst P, Sabatini DM (2002) mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. *Cell* 110: 163-175

Kim G, Gautier O, Tassoni-Tsuchida E, Ma XR, Gitler AD (2020) ALS Genetics: Gains, Losses, and Implications for Future Therapies. *Neuron* 108: 822-842

Kim H-J, Raphael AR, Ladow ES, McGurk L, Weber RA, Trojanowski JQ, Lee VM-Y, Finkbeiner S, Gitler AD, Bonini NM (2014) Therapeutic modulation of eIF2 α phosphorylation rescues TDP-43 toxicity in amyotrophic lateral sclerosis disease models. *Nature Genetics* 46: 152-160

Kim HJ, Kim NC, Wang YD, Scarborough EA, Moore J, Diaz Z, MacLea KS, Freibaum B, Li S, Molliex A *et al* (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495: 467-473

Kimball SR, Horetsky RL, Ron D, Jefferson LS, Harding HP (2003) Mammalian stress granules represent sites of accumulation of stalled translation initiation complexes. *Am J Physiol Cell Physiol* 284: C273-284

Klockgether T, Mariotti C, Paulson HL (2019) Spinocerebellar ataxia. *Nat Rev Dis Primers* 5: 24

Kohn JE, Millett IS, Jacob J, Zagrovic B, Dillon TM, Cingel N, Dothager RS, Seifert S, Thiyagarajan P, Sosnick TR *et al* (2004) Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc Natl Acad Sci U S A* 101: 12491-12496

Kouli A TK, Kuan WL. (2018) *Parkinson's Disease: Etiology, Neuropathology, and Pathogenesis*. . Codon Publications, Brisbane (AU)

Kozlov G, Safaee N, Rosenauer A, Gehring K (2010) Structural basis of binding of P-body-associated proteins GW182 and ataxin-2 by the Mlle domain of poly(A)-binding protein. *J Biol Chem* 285: 13599-13606

Kozlov G, Trempe J-F, Khaleghpour K, Kahvejian A, Ekiel I, Gehring K (2001) Structure and function of the C-terminal PABC domain of human poly(A)-binding protein. *Proceedings of the National Academy of Sciences* 98: 4409-4413

Kumar ST, Donzelli S, Chiki A, Syed MMK, Lashuel HA (2020) A simple, versatile and robust centrifugation-based filtration protocol for the isolation and quantification of α -synuclein monomers, oligomers and fibrils: Towards improving experimental reproducibility in α -synuclein research. *J Neurochem*: e14955

Kumar ST, Mahul-Mellier AL, Hegde RN, Rivière G, Moons R, Ibáñez de Opakua A, Magalhães P, Rostami I, Donzelli S, Sobott F *et al* (2022) A NAC domain mutation (E83Q) unlocks the pathogenicity of human alpha-synuclein and recapitulates its pathological diversity. *Sci Adv* 8: eabn0044

Laffita-Mesa JM, Paucar M, Svenningsson P (2021) Ataxin-2 gene: a powerful modulator of neurological disorders. *Current Opinion in Neurology* 34

Latonen L (2019) Phase-to-Phase With Nucleoli – Stress Responses, Protein Aggregation and Novel Roles of RNA. *Frontiers in Cellular Neuroscience* 13: 151

Lee J, Kim M, Itoh TQ, Lim C (2018) Ataxin-2: A versatile posttranscriptional regulator and its implication in neural function. *Wiley Interdiscip Rev RNA* 9: e1488

Lee J, Yoo E, Lee H, Park K, Hur JH, Lim C (2017) LSM12 and ME31B/DDX6 Define Distinct Modes of Posttranscriptional Regulation by ATAXIN-2 Protein Complex in *Drosophila* Circadian Pacemaker Neurons. *Mol Cell* 66: 129-140.e127

Lee KH, Zhang P, Kim HJ, Mitrea DM, Sarkar M, Freibaum BD, Cika J, Coughlin M, Messing J, Molliex A *et al* (2016) C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. *Cell* 167: 774-788.e717

Lefebvre S, Bürglen L, Reboullet S, Clermont O, Burlet P, Viollet L, Benichou B, Cruaud C, Millasseau P, Zeviani M *et al* (1995) Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 80: 155-165

Levin J, Schmidt F, Boehm C, Prix C, Bötzel K, Ryazanov S, Leonov A, Griesinger C, Giese A (2014) The oligomer modulator anle138b inhibits disease progression in a Parkinson mouse model even with treatment started after disease onset. *Acta Neuropathol* 127: 779-780

Li J-Y, Englund E, Holton JL, Soulet D, Hagell P, Lees AJ, Lashley T, Quinn NP, Rehnacrona S, Björklund A *et al* (2008) Lewy bodies in grafted neurons in subjects with Parkinson's disease suggest host-to-graft disease propagation. *Nature Medicine* 14: 501-503

Li P, Banjade S, Cheng H-C, Kim S, Chen B, Guo L, Llaguno M, Hollingsworth JV, King DS, Banani SF *et al* (2012a) Phase transitions in the assembly of multivalent signalling proteins. *Nature* 483: 336-340

Li WW, Li J, Bao JK (2012b) Microautophagy: lesser-known self-eating. *Cell Mol Life Sci* 69: 1125-1136

Li Y, Zhao C, Luo F, Liu Z, Gui X, Luo Z, Zhang X, Li D, Liu C, Li X (2018) Amyloid fibril structure of α -synuclein determined by cryo-electron microscopy. *Cell Res* 28: 897-903

- Li YR, King OD, Shorter J, Gitler AD (2013) Stress granules as crucibles of ALS pathogenesis. *J Cell Biol* 201: 361-372
- Lim C, Allada R (2013) ATAXIN-2 activates PERIOD translation to sustain circadian rhythms in *Drosophila*. *Science* 340: 875-879
- Limousin P, Foltynie T (2019) Long-term outcomes of deep brain stimulation in Parkinson disease. *Nature Reviews Neurology* 15: 234-242
- Lin Y, Currie SL, Rosen MK (2017) Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J Biol Chem* 292: 19110-19120
- Lin Y, Protter SW, David, Rosen K, Michael, Parker R (2015) Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Molecular Cell* 60: 208-219
- Lømo T (2003) The discovery of long-term potentiation. *Philos Trans R Soc Lond B Biol Sci* 358: 617-620
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15: 550
- Luk KC, Kehm V, Carroll J, Zhang B, O'Brien P, Trojanowski JQ, Lee VM (2012a) Pathological α -synuclein transmission initiates Parkinson-like neurodegeneration in nontransgenic mice. *Science* 338: 949-953
- Luk KC, Kehm VM, Zhang B, O'Brien P, Trojanowski JQ, Lee VM (2012b) Intracerebral inoculation of pathological α -synuclein initiates a rapidly progressive neurodegenerative α -synucleinopathy in mice. *J Exp Med* 209: 975-986
- Machida K, Shigeta T, Yamamoto Y, Ito T, Svitkin Y, Sonenberg N, Imataka H (2018) Dynamic interaction of poly(A)-binding protein with the ribosome. *Scientific Reports* 8: 17435
- Mahul-Mellier AL, Bartscher J, Maharjan N, Weerens L, Croisier M, Kuttler F, Leleu M, Knott GW, Lashuel HA (2020) The process of Lewy body formation, rather than simply α -synuclein fibrillization, is one of the major drivers of neurodegeneration. *Proc Natl Acad Sci U S A* 117: 4971-4982
- Malik BR, Maddison DC, Smith GA, Peters OM (2019) Autophagic and endo-lysosomal dysfunction in neurodegenerative disease. *Molecular Brain* 12: 100

- Malinowska L, Kroschwald S, Alberti S (2013) Protein disorder, prion propensities, and self-organizing macromolecular collectives. *Biochim Biophys Acta* 1834: 918-931
- Man WK, Tahirbegi B, Vrettas MD, Preet S, Ying L, Vendruscolo M, De Simone A, Fusco G (2021) The docking of synaptic vesicles on the presynaptic membrane induced by α -synuclein is modulated by lipid composition. *Nature Communications* 12: 927
- Mandrioli J, Mediani L, Alberti S, Carra S (2020) ALS and FTD: Where RNA metabolism meets protein quality control. *Semin Cell Dev Biol* 99: 183-192
- Mangus DA, Evans MC, Jacobson A (2003) Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol* 4: 223
- Maniatis T, Reed R (2002) An extensive network of coupling among gene expression machines. *Nature* 416: 499-506
- Martin EW, Holehouse AS (2020) Intrinsically disordered protein regions and phase separation: sequence determinants of assembly or lack thereof. *Emerg Top Life Sci* 4: 307-329
- Mateju D, Franzmann TM, Patel A, Kopach A, Boczek EE, Maharana S, Lee HO, Carra S, Hyman AA, Alberti S (2017) An aberrant phase transition of stress granules triggered by misfolded protein and prevented by chaperone function. *The EMBO journal* 36: 1669-1687
- Matheny T, Van Treeck B, Huynh TN, Parker R (2021) RNA partitioning into stress granules is based on the summation of multiple interactions. *RNA* 27: 174-189
- McCann C, Holohan EE, Das S, Dervan A, Larkin A, Lee JA, Rodrigues V, Parker R, Ramaswami M (2011) The Ataxin-2 protein is required for microRNA function and synapse-specific long-term olfactory habituation. *Proc Natl Acad Sci U S A* 108: E655-662
- McCormick C, Khaperskyy DA (2017) Translation inhibition and stress granules in the antiviral immune response. *Nat Rev Immunol* 17: 647-660
- McEwen E, Kedersha N, Song B, Scheuner D, Gilks N, Han A, Chen JJ, Anderson P, Kaufman RJ (2005) Heme-regulated inhibitor kinase-mediated phosphorylation of eukaryotic translation initiation factor 2 inhibits translation, induces stress granule formation, and mediates survival upon arsenite exposure. *J Biol Chem* 280: 16925-16933

- McMahon AC, Rahman R, Jin H, Shen JL, Fieldsend A, Luo W, Rosbash M (2016) TRIBE: Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding Proteins. *Cell* 165: 742-753
- Meisl G, Kirkegaard JB, Arosio P, Michaels TC, Vendruscolo M, Dobson CM, Linse S, Knowles TP (2016) Molecular mechanisms of protein aggregation from global fitting of kinetic models. *Nat Protoc* 11: 252-272
- Meyer-Luehmann M, Coomaraswamy J, Bolmont T, Kaeser S, Schaefer C, Kilger E, Neuenschwander A, Abramowski D, Frey P, Jaton AL *et al* (2006) Exogenous induction of cerebral beta-amyloidogenesis is governed by agent and host. *Science* 313: 1781-1784
- Middleton ER, Rhoades E (2010) Effects of curvature and composition on α -synuclein binding to lipid vesicles. *Biophys J* 99: 2279-2288
- Milovanovic D, De Camilli P (2017) Synaptic Vesicle Clusters at Synapses: A Distinct Liquid Phase? *Neuron* 93: 995-1002
- Mittag T, Parker R (2018) Multiple Modes of Protein-Protein Interactions Promote RNP Granule Assembly. *J Mol Biol* 430: 4636-4649
- Moon SL, Morisaki T, Khong A, Lyon K, Parker R, Stasevich TJ (2019) Multicolour single-molecule tracking of mRNA interactions with RNP granules. *Nat Cell Biol* 21: 162-168
- Morgan SA, Lavenir I, Fan J, Masuda-Suzukake M, Passarella D, DeTure MA, Dickson DW, Ghetti B, Goedert M (2020) α -Synuclein filaments from transgenic mouse and human synucleinopathy-containing brains are major seed-competent species. *J Biol Chem* 295: 6652-6664
- Morris GA (2009) *Diffusion-Ordered Spectroscopy*. John Wiley & Sons, Ltd
- Munishkina LA, Fink AL, Uversky VN (2008) Concerted action of metals and macromolecular crowding on the fibrillation of alpha-synuclein. *Protein Pept Lett* 15: 1079-1085
- Munishkina LA, Phelan C, Uversky VN, Fink AL (2003) Conformational behavior and aggregation of alpha-synuclein in organic solvents: modeling the effects of membranes. *Biochemistry* 42: 2720-2730

Murray DT, Kato M, Lin Y, Thurber KR, Hung I, McKnight SL, Tycko R (2017) Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell* 171: 615-627.e616

Nag M, Das D, Bandyopadhyay D, Basak S (2015) Unusual denaturation trajectory of bovine gamma globulin studied by fluorescence correlation spectroscopy. *Phys Chem Chem Phys* 17: 19139-19148

Nath S, Meuvlis J, Hendrix J, Carl SA, Engelborghs Y (2010) Early aggregation steps in alpha-synuclein as measured by FCS and FRET: evidence for a contagious conformational change. *Biophys J* 98: 1302-1311

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM *et al* (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* 314: 130-133

Nonhoff U, Ralser M, Welzel F, Piccini I, Balzereit D, Yaspo ML, Lehrach H, Krobitsch S (2007) Ataxin-2 interacts with the DEAD/H-box RNA helicase DDX6 and interferes with P-bodies and stress granules. *Mol Biol Cell* 18: 1385-1396

Nover L, Scharf K, Neumann D (1983) Formation of cytoplasmic heat shock granules in tomato cells. *Molecular and cellular biology* 3: 1648-1655

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

Olzscha H, Schermann SM, Woerner AC, Pinkert S, Hecht MH, Tartaglia GG, Vendruscolo M, Hayer-Hartl M, Hartl FU, Vabulas RM (2011) Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell* 144: 67-78

Ostrowski LA, Hall AC, Mekhail K (2017) Ataxin-2: From RNA Control to Human Health and Disease. *Genes (Basel)* 8: 157

Pakos-Zebrucka K, Koryga I, Mnich K, Lujic M, Samali A, Gorman AM (2016) The integrated stress response. *EMBO Rep* 17: 1374-1395

Papoulas O, Monzo KF, Cantin GT, Ruse C, Yates JR, Ryu YH, Sisson JC (2010) dFMRP and Caprin, translational regulators of synaptic plasticity, control the cell cycle at the *Drosophila* mid-blastula transition. *Development* 137: 4201-4209

- Park D, Wu Y, Lee S-E, Kim G, Jeong S, Milovanovic D, De Camilli P, Chang S (2021) Cooperative function of synaptophysin and synapsin in the generation of synaptic vesicle-like clusters in non-neuronal cells. *Nature Communications* 12: 263
- Parker R, Noel TR, Brownsey GJ, Laos K, Ring SG (2005) The nonequilibrium phase and glass transition behavior of beta-lactoglobulin. *Biophys J* 89: 1227-1236
- Patel A, Lee HO, Jawerth L, Maharana S, Jahnelt M, Hein MY, Stoyanov S, Mahamid J, Saha S, Franzmann TM *et al* (2015) A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* 162: 1066-1077
- Paulson HL, Shakkottai VG, Clark HB, Orr HT (2017) Polyglutamine spinocerebellar ataxias - from genes to potential treatments. *Nat Rev Neurosci* 18: 613-626
- Pérez Ortiz JM, Mollema N, Toker N, Adamski CJ, O'Callaghan B, Duvick L, Friedrich J, Walters MA, Strasser J, Hawkinson JE *et al* (2018) Reduction of protein kinase A-mediated phosphorylation of ATXN1-S776 in Purkinje cells delays onset of Ataxia in a SCA1 mouse model. *Neurobiol Dis* 116: 93-105
- Poewe W, Seppi K, Tanner CM, Halliday GM, Brundin P, Volkman J, Schrag A-E, Lang AE (2017) Parkinson disease. *Nature Reviews Disease Primers* 3: 17013
- Powers ET, Morimoto RI, Dillin A, Kelly JW, Balch WE (2009) Biological and chemical approaches to diseases of proteostasis deficiency. *Annu Rev Biochem* 78: 959-991
- Preissler S, Ron D (2019) Early Events in the Endoplasmic Reticulum Unfolded Protein Response. *Cold Spring Harb Perspect Biol* 11
- Price WS (2002) *Diffusion-Based Studies of Aggregation, Binding and Conformation of Biomolecules: Theory and Practice*. John Wiley & Sons, Chichester
- Protter DSW, Parker R (2016) Principles and Properties of Stress Granules. *Trends Cell Biol* 26: 668-679
- Protter DSW, Rao BS, Van Treeck B, Lin Y, Mizoue L, Rosen MK, Parker R (2018) Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell Rep* 22: 1401-1412
- Prusiner SB (1998) Prions. *Proc Natl Acad Sci U S A* 95: 13363-13383

Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841-842

Ragland SA, Criss AK (2017) From bacterial killing to immune modulation: Recent insights into the functions of lysozyme. *PLoS Pathog* 13: e1006512

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

Rao G, Croft B, Teng C, Awasthi V (2015) Ubiquitin-Proteasome System in Neurodegenerative Disorders. *J Drug Metab Toxicol* 6: 187

Ray S, Singh N, Kumar R, Patel K, Pandey S, Datta D, Mahato J, Panigrahi R, Navalkar A, Mehra S *et al* (2020) α -Synuclein aggregation nucleates through liquid–liquid phase separation. *Nature Chemistry* 12: 705-716

Recasens A, Dehay B (2014) Alpha-synuclein spreading in Parkinson's disease. *Frontiers in Neuroanatomy* 8

Reineke LC, Kedersha N, Langereis MA, van Kuppeveld FJ, Lloyd RE (2015) Stress granules regulate double-stranded RNA-dependent protein kinase activation through a complex containing G3BP1 and Caprin1. *mBio* 6: e02486

Rekas A, Ahn KJ, Kim J, Carver JA (2012) The chaperone activity of α -synuclein: Utilizing deletion mutants to map its interaction with target proteins. *Proteins* 80: 1316-1325

Rekas A, Knott RB, Sokolova A, Barnham KJ, Perez KA, Masters CL, Drew SC, Cappai R, Curtain CC, Pham CL (2010) The structure of dopamine induced alpha-synuclein oligomers. *Eur Biophys J* 39: 1407-1419

Rösener NS, Gremer L, Wördehoff MM, Kupreichyk T, Etzkorn M, Neudecker P, Hoyer W (2020) Clustering of human prion protein and α -synuclein oligomers requires the prion protein N-terminus. *Communications Biology* 3: 365

Ross CA, Poirier MA (2004) Protein aggregation and neurodegenerative disease. *Nat Med* 10 Suppl: S10-17

Rousseau A, Bertolotti A (2016) An evolutionarily conserved pathway controls proteasome homeostasis. *Nature* 536: 184-189

Rousseaux MWC, Tschumperlin T, Lu HC, Lackey EP, Bondar VV, Wan YW, Tan Q, Adamski CJ, Friedrich J, Twaroski K *et al* (2018) ATXN1-CIC Complex Is the Primary Driver of Cerebellar Pathology in Spinocerebellar Ataxia Type 1 through a Gain-of-Function Mechanism. *Neuron* 97: 1235-1243.e1235

Roy G, De Crescenzo G, Khaleghpour K, Kahvejian A, O'Connor-McCourt M, Sonenberg N (2002) Paip1 interacts with poly(A) binding protein through two independent binding motifs. *Mol Cell Biol* 22: 3769-3782

Ruggeri FS, Charmet J, Kartanas T, Peter Q, Chia S, Habchi J, Dobson CM, Vendruscolo M, Knowles TPJ (2018) Microfluidic deposition for resolving single-molecule protein architecture and heterogeneity. *Nat Commun* 9: 3890

Safaei N, Kozlov G, Noronha AM, Xie J, Wilds CJ, Gehring K (2012) Interdomain allostery promotes assembly of the poly(A) mRNA complex with PABP and eIF4G. *Mol Cell* 48: 375-386

Saibil H (2013) Chaperone machines for protein folding, unfolding and disaggregation. *Nature Reviews Molecular Cell Biology* 14: 630-642

Sangwan S, Sahay S, Murray KA, Morgan S, Guenther EL, Jiang L, Williams CK, Vinters HV, Goedert M, Eisenberg DS (2020) Inhibition of synucleinopathic seeding by rationally designed inhibitors. *Elife* 9

Satterfield TF, Pallanck LJ (2006) Ataxin-2 and its Drosophila homolog, ATX2, physically assemble with polyribosomes. *Hum Mol Genet* 15: 2523-2532

Savica R, Grossardt BR, Bower JH, Ahlskog JE, Rocca WA (2013) Incidence and pathology of synucleinopathies and tauopathies related to parkinsonism. *JAMA Neurol* 70: 859-866

Sawner AS, Ray S, Yadav P, Mukherjee S, Panigrahi R, Poudyal M, Patel K, Ghosh D, Kummerant E, Kumar A *et al* (2021) Modulating α -Synuclein Liquid-Liquid Phase Separation. *Biochemistry* 60: 3676-3696

Schöls L, Bauer P, Schmidt T, Schulte T, Riess O (2004) Autosomal dominant cerebellar ataxias: clinical features, genetics, and pathogenesis. *Lancet Neurol* 3: 291-304

Schuller AP, Wojtynek M, Mankus D, Tatli M, Kronenberg-Tenga R, Regmi SG, Dip PV, Lytton-Jean AKR, Brignole EJ, Dasso M *et al* (2021) The cellular environment shapes the nuclear pore complex architecture. *Nature* 598: 667-671

Schwarz DS, Blower MD (2016) The endoplasmic reticulum: structure, function and response to cellular signaling. *Cellular and molecular life sciences : CMLS* 73: 79-94

Scoles DR, Meera P, Schneider MD, Paul S, Dansithong W, Figueroa KP, Hung G, Rigo F, Bennett CF, Otis TS *et al* (2017) Antisense oligonucleotide therapy for spinocerebellar ataxia type 2. *Nature* 544: 362-366

Selkoe DJ (2003) Folding proteins in fatal ways. *Nature* 426: 900-904

Sharma P, Verma N, Singh PK, Korpole S, Ashish (2016) Characterization of heat induced spherulites of lysozyme reveals new insight on amyloid initiation. *Sci Rep* 6: 22475

Shiau AK, Harris SF, Southworth DR, Agard DA (2006) Structural Analysis of *E. coli* hsp90 Reveals Dramatic Nucleotide-Dependent Conformational Rearrangements. *Cell* 127: 329-340

Shimizu A, Ikeguchi M, Sugai S (1994) Appropriateness of DSS and TSP as internal references for (1)H NMR studies of molten globule proteins in aqueous media. *J Biomol NMR* 4: 859-862

Shin Y, Brangwynne CP (2017) Liquid phase condensation in cell physiology and disease. *Science* 357

Shulman JM, Feany MB (2003) Genetic Modifiers of Tauopathy in Drosophila. *Genetics* 165: 1233-1242

Sidrauski C, Mcgeachy AM, Ingolia NT, Walter P (2015) The small molecule ISRIB reverses the effects of eIF2 α phosphorylation on translation and stress granule assembly. *eLife* 4

Singh A, Hulsmeier J, Kandi AR, Pothapragada SS, Hillebrand J, Petrauskas A, Agrawal K, Rt K, Thiagarajan D, Jayaprakashappa D *et al* (2021) Antagonistic roles for Ataxin-2 structured and disordered domains in RNP condensation. *eLife* 10

Singh A, Kandi AR, Jayaprakashappa D, Thuery G, Purohit DJ, Huelsmeier J, Singh R, Pothapragada SS, Ramaswami M, Bakthavachalu B (2022) The transcriptional response

to oxidative stress is independent of stress-granule formation. *Molecular Biology of the Cell* 33: ar25

Soto C (2003) Unfolding the role of protein misfolding in neurodegenerative diseases. *Nat Rev Neurosci* 4: 49-60

Soto C, Pritzkow S (2018) Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases. *Nat Neurosci* 21: 1332-1340

Spasic M, Friedel CC, Schott J, Kreth J, Leppek K, Hofmann S, Ozgur S, Stoecklin G (2012) Genome-Wide Assessment of AU-Rich Elements by the AREScore Algorithm. *PLOS Genetics* 8: e1002433

Spillantini MG, Schmidt ML, Lee VMY, Trojanowski JQ, Jakes R, Goedert M (1997) α -Synuclein in Lewy bodies. *Nature* 388: 839-840

Srivastava SP, Kumar KU, Kaufman RJ (1998) Phosphorylation of eukaryotic translation initiation factor 2 mediates apoptosis in response to activation of the double-stranded RNA-dependent protein kinase. *J Biol Chem* 273: 2416-2423

Standart N, Weil D (2018) P-Bodies: Cytosolic Droplets for Coordinated mRNA Storage. *Trends Genet* 34: 612-626

Stejskal EO, Tanner JE (1965) Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *J Chem Phys*

Stoecklin G, Kedersha N (2013) Relationship of GW/P-bodies with stress granules. *Adv Exp Med Biol* 768: 197-211

Sudhakaran IP, Hillebrand J, Dervan A, Das S, Holohan EE, Hülsmeier J, Sarov M, Parker R, VijayRaghavan K, Ramaswami M (2014) FMRP and Ataxin-2 function together in long-term olfactory habituation and neuronal translational control. *Proc Natl Acad Sci U S A* 111: E99-E108

Surmeier DJ, Obeso JA, Halliday GM (2017) Selective neuronal vulnerability in Parkinson disease. *Nature Reviews Neuroscience* 18: 101-113

Takayama S, Reed JC, Homma S (2003) Heat-shock proteins as regulators of apoptosis. *Oncogene* 22: 9041-9047

Taylor JP, Brown RH, Cleveland DW (2016) Decoding ALS: from genes to mechanism. *Nature* 539: 197-206

Thibautaud TA, Anderson RT, Smith DM (2018) A common mechanism of proteasome impairment by neurodegenerative disease-associated oligomers. *Nature Communications* 9: 1097

Tillotson R, Selfridge J, Koerner MV, Gadalla KKE, Guy J, De Sousa D, Hector RD, Cobb SR, Bird A (2017) Radically truncated MeCP2 rescues Rett syndrome-like neurological defects. *Nature* 550: 398-401

Toste Rêgo A, da Fonseca PCA (2019) Characterization of Fully Recombinant Human 20S and 20S-PA200 Proteasome Complexes. *Molecular Cell* 76: 138-147.e135

Tran J, Anastacio H, Bardy C (2020) Genetic predispositions of Parkinson's disease revealed in patient-derived brain cells. *npj Parkinson's Disease* 6: 8

Troncoso-Escudero P, Parra A, Nassif M, Vidal RL (2018) Outside in: Unraveling the Role of Neuroinflammation in the Progression of Parkinson's Disease. *Frontiers in Neurology* 9

Trudler D, Sanz-Blasco S, Eisele YS, Ghatak S, Bodhinathan K, Akhtar MW, Lynch WP, Piña-Crespo JC, Talantova M, Kelly JW *et al* (2021) α -Synuclein Oligomers Induce Glutamate Release from Astrocytes and Excessive Extrasynaptic NMDAR Activity in Neurons, Thus Contributing to Synapse Loss. *J Neurosci* 41: 2264-2273

Tsunemi T, Ishiguro Y, Yoroisaka A, Hattori N (2021) Analysis of α -Synuclein in Exosomes. *Methods Mol Biol* 2322: 41-45

Twelves D, Perkins KS, Counsell C (2003) Systematic review of incidence studies of Parkinson's disease. *Mov Disord* 18: 19-31

Tziortzouda P, Van Den Bosch L, Hirth F (2021) Triad of TDP43 control in neurodegeneration: autoregulation, localization and aggregation. *Nat Rev Neurosci* 22: 197-208

Ulmer TS, Bax A, Cole NB, Nussbaum RL (2005) Structure and Dynamics of Micelle-bound Human α -Synuclein*. *Journal of Biological Chemistry* 280: 9595-9603

Urrea L, Ferrer I, Gavín R, Del Río JA (2017) The cellular prion protein (PrP(C)) as neuronal receptor for α -synuclein. *Prion* 11: 226-233

Uversky VN (2003) A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders. *J Biomol Struct Dyn* 21: 211-234

van Blitterswijk M, Mullen B, Heckman MG, Baker MC, DeJesus-Hernandez M, Brown PH, Murray ME, Hsiung GY, Stewart H, Karydas AM *et al* (2014) Ataxin-2 as potential disease modifier in C9ORF72 expansion carriers. *Neurobiol Aging* 35: 2421.e2413-2427

Van Treeck B, Parker R (2018) Emerging Roles for Intermolecular RNA-RNA Interactions in RNP Assemblies. *Cell* 174: 791-802

Van Treeck B, Protter DSW, Matheny T, Khong A, Link CD, Parker R (2018) RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. *Proc Natl Acad Sci U S A* 115: 2734-2739

Vance C, Rogelj B, Hortobágyi T, De Vos KJ, Nishimura AL, Sreedharan J, Hu X, Smith B, Ruddy D, Wright P *et al* (2009) Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 323: 1208-1211

Vandelli A, Cid Samper F, Torrent Burgas M, Sanchez de Groot N, Tartaglia GG (2022) The Interplay Between Disordered Regions in RNAs and Proteins Modulates Interactions Within Stress Granules and Processing Bodies. *Journal of Molecular Biology* 434: 167159

Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A *et al* (2022) AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research* 50: D439-D444

Vlad C, Lindner K, Karreman C, Schildknecht S, Leist M, Tomczyk N, Rontree J, Langridge J, Danzer K, Ciossek T *et al* (2011) Autoproteolytic fragments are intermediates in the oligomerization/aggregation of the Parkinson's disease protein alpha-synuclein as revealed by ion mobility mass spectrometry. *Chembiochem* 12: 2740-2744

Vogler TO, Wheeler JR, Nguyen ED, Hughes MP, Britson KA, Lester E, Rao B, Betta ND, Whitney ON, Ewachiw TE *et al* (2018) TDP-43 and RNA form amyloid-like myo-granules in regenerating muscle. *Nature* 563: 508-513

Volpicelli-Daley LA, Luk KC, Patel TP, Tanik SA, Riddle DM, Stieber A, Meaney DF, Trojanowski JQ, Lee VM (2011) Exogenous α -synuclein fibrils induce Lewy body pathology leading to synaptic dysfunction and neuron death. *Neuron* 72: 57-71

- Voronina E, Seydoux G, Sassone-Corsi P, Nagamori I (2011) RNA granules in germ cells. *Cold Spring Harb Perspect Biol* 3
- Walsh DM, Klyubin I, Fadeeva JV, Cullen WK, Anwyl R, Wolfe MS, Rowan MJ, Selkoe DJ (2002) Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo. *Nature* 416: 535-539
- Wang W, Perovic I, Chittuluru J, Kaganovich A, Nguyen LT, Liao J, Auclair JR, Johnson D, Landeru A, Simorellis AK *et al* (2011) A soluble α -synuclein construct forms a dynamic tetramer. *Proc Natl Acad Sci U S A* 108: 17797-17802
- Wegmann S, Eftekharzadeh B, Tepper K, Zoltowska KM, Bennett RE, Dujardin S, Laskowski PR, MacKenzie D, Kamath T, Commins C *et al* (2018) Tau protein liquid-liquid phase separation can initiate tau aggregation. *Embo j* 37
- Weiss MS, Palm GJ, Hilgenfeld R (2000) Crystallization, structure solution and refinement of hen egg-white lysozyme at pH 8.0 in the presence of MPD. *Acta Crystallogr D Biol Crystallogr* 56: 952-958
- Wek SA, Zhu S, Wek RC (1995) The histidyl-tRNA synthetase-related sequence in the eIF-2 alpha protein kinase GCN2 interacts with tRNA and is required for activation in response to starvation for different amino acids. *Mol Cell Biol* 15: 4497-4506
- White J, Heß D, Raynes J, Laux V, Haertlein M, Forsyth T, Jeyasingham A (2015) The aggregation of "native" human serum albumin. *Eur Biophys J* 44: 367-371
- Wilkins DK, Grimshaw SB, Receveur V, Dobson CM, Jones JA, Smith LJ (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* 38: 16424-16431
- Wolozin B, Ivanov P (2019) Stress granules and neurodegeneration. *Nature Reviews Neuroscience* 20: 649-666
- Wong YC, Krainc D (2017) α -synuclein toxicity in neurodegeneration: mechanism and therapeutic strategies. *Nature Medicine* 23: 1-13
- Wong YL, Lebon L, Edalji R, Lim HB, Sun C, Sidrauski C (2018) The small molecule ISRIB rescues the stability and activity of Vanishing White Matter Disease eIF2B mutant complexes. *eLife* 7

Wright PE, Dyson HJ (2015) Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology* 16: 18-29

Xiao Y, Ma B, McElheny D, Parthasarathy S, Long F, Hoshi M, Nussinov R, Ishii Y (2015) A β (1–42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. *Nature Structural & Molecular Biology* 22: 499-505

Xie J, Kozlov G, Gehring K (2014) The "tale" of poly(A) binding protein: the MLLE domain and PAM2-containing proteins. *Biochim Biophys Acta* 1839: 1062-1068

Yang P, Mathieu C, Kolaitis RM, Zhang P, Messing J, Yurtsever U, Yang Z, Wu J, Li Y, Pan Q *et al* (2020) G3BP1 Is a Tunable Switch that Triggers Phase Separation to Assemble Stress Granules. *Cell* 181: 325-345.e328

Yaremchuk A, Tukalo M, Grøtli M, Cusack S (2001) A succession of substrate induced conformational changes ensures the amino acid specificity of *Thermus thermophilus* prolyl-tRNA synthetase: comparison with histidyl-tRNA synthetase. *J Mol Biol* 309: 989-1002

Yi H, Park J, Ha M, Lim J, Chang H, Kim VN (2018) PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay. *Mol Cell* 70: 1081-1088.e1085

Yokoshi M, Li Q, Yamamoto M, Okada H, Suzuki Y, Kawahara Y (2014) Direct binding of Ataxin-2 to distinct elements in 3' UTRs promotes mRNA stability and protein expression. *Mol Cell* 55: 186-198

Yoshida M, Yoshida K, Kozlov G, Lim NS, De Crescenzo G, Pang Z, Berlanga JJ, Kahvejian A, Gehring K, Wing SS *et al* (2006) Poly(A) binding protein (PABP) homeostasis is mediated by the stability of its inhibitor, Paip2. *The EMBO Journal* 25: 1934-1944

Youn JY, Dyakov BJA, Zhang J, Knight JDR, Vernon RM, Forman-Kay JD, Gingras AC (2019) Properties of Stress Granule and P-Body Proteomes. *Mol Cell* 76: 286-294

Zhang J, Guan X, Shah K, Yan J (2021) Lsm12 is an NAADP receptor and a two-pore channel regulatory protein required for calcium mobilization from acidic organelles. *Nature Communications* 12: 4739

Zhang K, Daigle JG, Cunningham KM, Coyne AN, Ruan K, Grima JC, Bowen KE, Wadhwa H, Yang P, Rigo F *et al* (2018) Stress Granule Assembly Disrupts Nucleocytoplasmic Transport. *Cell* 173: 958-971.e917

Zhang Y, Ling J, Yuan C, Dubruille R, Emery P (2013) A role for *Drosophila* ATX2 in activation of PER translation and circadian behavior. *Science* 340: 879-882

Zhou L, Hang J, Zhou Y, Wan R, Lu G, Yin P, Yan C, Shi Y (2014) Crystal structures of the Lsm complex bound to the 3' end sequence of U6 small nuclear RNA. *Nature* 506: 116-120

Zhu M, Rajamani S, Kaylor J, Han S, Zhou F, Fink AL (2004) The flavonoid baicalein inhibits fibrillation of alpha-synuclein and disaggregates existing fibrils. *J Biol Chem* 279: 26846-26857

Zyryanova AF, Kashiwagi K, Rato C, Harding HP, Crespillo-Casado A, Perera LA, Sakamoto A, Nishimoto M, Yonemochi M, Shirouzu M *et al* (2021) ISRIB Blunts the Integrated Stress Response by Allosterically Antagonising the Inhibitory Effect of Phosphorylated eIF2 on eIF2B. *Mol Cell* 81: 88-103.e106

8. PUBLICATIONS & PRESENTATIONS

Arnas Petrauskas* & Daniel L. Fortunati* *et al.* **Structured and disordered regions of Ataxin-2 contribute differently to the specificity and efficiency of mRNP granule formation.** bioRxiv (2022).

Antonín Brisuda, James C. S. Ho, Pancham S. Kandiyal, Justin T-Y. Ng, Ines Ambite, Daniel S. C. Butler, Jaromir Háček, Murphy Lam Yim Wan, Thi Hien Tran, Aftab Nadeem, Tuan Hiep Tran, Anna Hastings, Petter Storm, Daniel L. Fortunati *et al.* **Bladder cancer therapy using a conformationally fluid tumoricidal peptide complex.** Nat Commun 12, 3427 (2021).

Pancham S. Kandiyal, JiYoon Kim, Daniel L. Fortunati, Mok KH. **Size Determination of Protein Oligomers/Aggregates Using Diffusion NMR Spectroscopy.** Methods Mol Biol. (2019).

Poster presentation at Molecular and Computational Biology Symposium, 2018, in Dublin (Ireland), titled: Structural & kinetic studies of oligomeric aggregates of α -synuclein.

Poster presentation at 28th ICMRBS Congress, 2018, in Dublin (Ireland), titled: Structural & kinetic studies of oligomeric aggregates of α -synuclein.

Symposium presentation at the 24th IUBMB Congress and 15th FAOBMB Congress, 2018, in Seoul (Korea), titled: Structural & Kinetic studies of oligomeric aggregates of α -synuclein.

9. APPENDIX

Pancham S. Kandiyal, JiYoon Kim, Daniel L. Fortunati, Mok KH. **Size Determination of Protein Oligomers/Aggregates Using Diffusion NMR Spectroscopy**. Methods Mol Biol. (2019).



Chapter 13

Size Determination of Protein Oligomers/Aggregates Using Diffusion NMR Spectroscopy

Pancham S. Kandiyal, Ji Yoon Kim, Daniel L. Fortunati,
and K. H. Mok

Abstract

Diffusion-ordered spectroscopy (DOSY) is a widely used NMR technique for the identification of different chemical moieties/compounds contained in mixtures and has been successfully employed for the separation of small molecules based on hydrodynamic radii. Herein we show that DOSY can also be applied for the size determination of larger biomolecules such as proteins and protein oligomers/aggregates. Proof-of-principle is first shown with a cross-linked oligomeric protein mixture where the hydrodynamic volumes of each component are estimated and subsequently verified with size-exclusion HPLC and SDS polyacrylamide gel electrophoresis. We then determine the sizes of protein oligomers contained in a protein solution subjected under amyloid fibrillogenesis conditions. These studies aim to provide insight into the kinetics behind protein aggregation involved in amyloidosis as well as to determine the hydrodynamic radii of proteins within the mixture.

Key words DOSY, Hydrodynamic radii, Protein, Oligomers, Aggregation, Diffusion coefficient, Pulsed-field gradient (PFG), Stokes–Einstein equation

1 Introduction

Nuclear magnetic resonance (NMR) spectroscopy is used for characterization of conformational and atomic level details of small molecules, peptides and proteins in solution. Availability of various multidimensional, multinuclear experiments and isotopically labeled techniques have made life easier to study the 3D structure of macromolecules using NMR spectroscopy [1]. Diffusion is an intrinsic property of all molecules and varies on the basis of a particle's size, shape, solubility, charge distribution, etc. [2]. The theoretical and experimental demonstration for self-diffusion property of molecules was first given by Stejskal and Tanner in 1965 [3]. Advancement in hardware and software technology in NMR have made the performing of diffusion experiments straightforward for small molecules, leading to (1) the determination of the

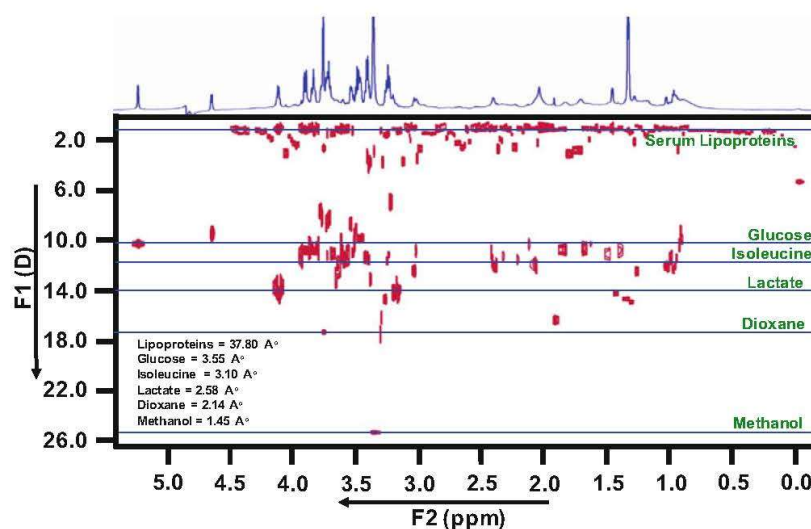


Fig. 1 Separation of metabolites and lipoproteins in a serum sample using DOSY NMR

mobility of a compound, (2) binding of small molecules to protein, (3) aggregation of molecules, and (4) identification of individual compounds in a mixture (Fig. 1) [4].

One of the key technical features in a diffusion NMR experiment pulse sequence is the utilization of pulsed-field gradients (PFGs) that spatially encode the solutes, the magnetizations of which are then refocused through either a spin echo or stimulated-echo of the spins (PFGSE or PFGSTE) [5]. For the past two decades, DOSY has been implemented for the characterization of mixtures containing small molecules [6, 7]. In contrast, its application for macromolecules has been less due to the shorter relaxation times and the overlapping of chemical shifts, despite the fact that in principle, DOSY NMR can be very helpful to understand the rate of protein aggregation and oligomer formation. While previous work with DOSY NMR has been able to show the separation of proteins and lipids, samples containing a mixture of different size proteins have not been shown with any frequency [8]. Here we have implemented 2D DOSY NMR to determine the size of an intentionally prepared, chemically cross-linked oligomeric mixture of a protein by addition of a cross-linker [9]. For our experiment, we used DOSY bipolar pulse pair stimulated echo (Dbppste) pulse sequence [10] and optimized the experiment by varying different parameters such as the diffusion gradient length, gradient diffusion delay time, relaxation delay and gradient recovery time.

When samples contain a mixture of different proteins, it is hard to readily distinguish all of the different sized-proteins according as a function of their diffusion coefficient value (D value) due to the large overlapping of their chemical shifts. For our cross-linked sample contained a mixture of overlapping signals originating from monomers, dimers and trimers of a protein, extracting the diffusion coefficients from requires deconvolution of the data. By plotting the intensities/peak areas obtained against the square of the gradient strength (G^2) according to the Stejskal–Tanner equation and fitting the data points with multiple Gaussian curves, it was possible to capture more subtle differences in diffusion coefficients (D) which would not be accessible using a less involved analysis. The D value used further for calculating the hydrodynamic radii using the Stokes–Einstein equation [11].

From these data and the exponential/Gaussian fitting of them, we were able to extract three different sizes of component protein, that is, monomer, dimer, and trimer for lysozyme, despite the fact that the overlapping NMR signals of the oligomers made it impossible to consider these proteins as separate components from the 2D DOSY experiment alone, and as a result making it difficult to calculate their different diffusion coefficient values.

2 Materials

All reagents should be prepared using ultrapure water.

1. Proteins; Hen egg-white lysozyme (HEWL; lyophilized powder, protein $\geq 90\%$, $\geq 40,000$ units/mg protein), bovine thyroglobulin, bovine γ -globulins, human serum albumin, β -lactoglobulin, horse cytochrome c , L-tryptophan.
2. Deuterium oxide (D_2O).
3. 1,4-Dioxane.
4. The water-soluble, homobifunctional cross-linker bis(sulfosuccinimidyl) substrate (BS^3). BS^3 contains sulfo-NHS (N -hydroxysuccinimide) esters on both carboxylate ends of the molecule, providing greater solubility due to the negatively charged sulfonate groups, making the molecule very hydrophilic.
5. Chromatography was performed on a Dionex Ultimate HPLC 3000 Standard System running Chromeleon 6 software (Dionex, Thermo Scientific), and the size exclusion-HPLC (SE-HPLC) was a Superdex 200 10/300 GL column (10 mm $\times$ 30 cm, Particle size 13 μm , GE Healthcare).
6. SDS electrophoresis gels, tank, and power supply.
7. NMR experiments were carried out on an Agilent Technologies 18.8T (800 MHz) DD2 Premium Compact spectrometer

with a triple-resonance, 5 mm enhanced cold probe; however, any high-field magnet with pulsed-field gradient capabilities will be sufficient. (The interest is more in monitoring the decay of signal as a function of applied gradient than the resolution of the spectra.)

8. HPLC elution buffer; 0.05 M sodium phosphate at pH 7.0 containing 0.1 M Na_2SO_4 , filtered through a 0.45 μm (or 0.2 μm) pore-size membrane and degassed.
9. Conjugation buffer and; 0.1 M sodium phosphate buffer at pH 7.0 containing 0.15 M NaCl, filtered through a 0.45 μm filter.
10. Quenching buffer for cross-linking with BS^3 ; 1 M Tris-HCl, pH 7.5 filtered through a 0.45 μm (or 0.2 μm) pore-size membrane.
11. Solutions for NMR; 0.05 M sodium phosphate buffer at pH 7.0 filtered through a 0.45 μm (or 0.2 μm) pore-size membrane. Other buffers are also acceptable provided that the buffer does not contain components (usually of aliphatic nature) that have strong, unexchangeable ^1H spectral signals. All solutions for NMR should contain at least 10% of D_2O that will be used as a deuterium lock signal.

3 Methods

3.1 Initial Diffusion Coefficient Value Estimation from SE-HPLC Data

An initial, rough estimate of the expected diffusion coefficients of the cross-linked HEWL species can be obtained from analysis of size exclusion-HPLC (SE-HPLC) data.

1. Dissolve the calibration standards at a concentration of 0.1 mM with the elution buffer and inject onto the SE-HPLC column eluting with 0.05 M sodium phosphate buffer pH 7.0 containing 0.1 M Na_2SO_4 at a flow rate of 0.50 mL/min with detection at 280 nm.
2. Generate a standard calibration curve using a suitable range of protein molecular weight standards (for example, bovine thyroglobulin, bovine γ -globulins, human serum albumin, β -lactoglobulin, horse cytochrome c , L-tryptophan) including human serum albumin (HSA), which has a hydrodynamic radius (R_H) of 40 Å (Figs. 2 and 3). The calibration curve is a plot of R_H values of these proteins vs. the elution volume obtained from the SE-HPLC.
3. Prepare cross-linked forms of HEWL using BS^3 . For a HEWL concentration of 5 mg/mL, use a tenfold molar excess of BS^3 cross-linker and react at room temperature for 3 h. Quench the reaction using the quenching buffer to a final concentration of

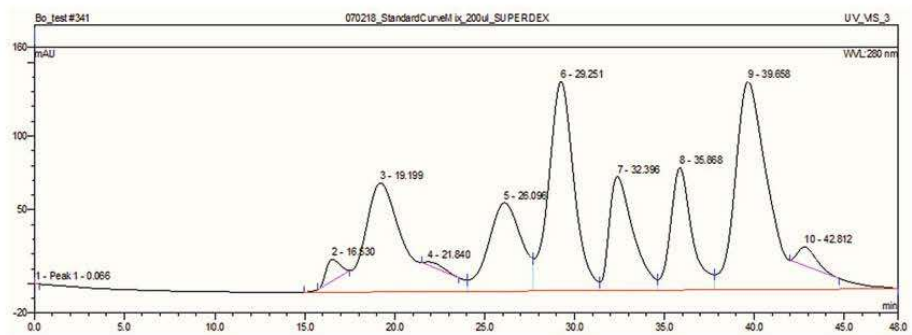


Fig. 2 Protein calibration standard chromatogram using size-exclusion HPLC

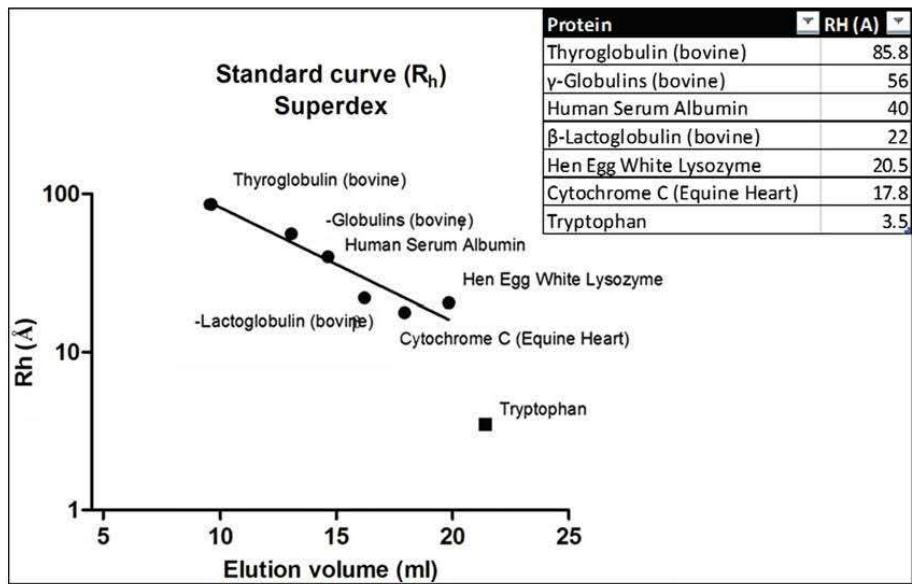


Fig. 3 SE-HPLC standard curve

- 20–50 mM Tris. Dialyze the protein solution with 0.05 M sodium phosphate buffer pH 7.0 (or any buffer that will keep the protein well solubilized).
4. Run a SDS-PAGE to observe the extent of cross-linking.
 5. Inject an approximate 0.1 mM concentration sample of cross-linked HEWL onto the SE-HPLC and measure the retention times (*see Note 1*).

Table 1**Rough estimates of hydrodynamic radii/diffusion coefficients obtained from SE-HPLC results**

HEWL peak	R_t (min)	R_H (Å)	D (10^{10} m ² /s)
Trimer	36.5	28	0.87
Dimer	39.0	23	1.10
Monomer	41.0	20	1.21
Compacted monomer	42.9	17	1.42

6. Obtain the R_H values by using the linear relationship found on the standard calibration curve (Fig. 3). From the Stokes-Einstein equation, a rough estimate of the diffusion coefficients (D) for each species is as follows:

$$D = \frac{k_b T}{6\pi\eta R_H}$$

where k_b is Boltzmann's constant, T is the absolute temperature, η is the dynamic viscosity, and R_H is the radius of the spherical particle.

Typical retention times and corresponding hydrodynamic radii/diffusion coefficients obtained shown in Table 1.

3.2 DOSY-NMR and the Identification of Protein Oligomer/Aggregate Components in a Mixture

3.2.1 Background

The diffusion coefficient (D) of a single species of molecule can be extracted from DOSY-NMR data by plotting the Intensity of the peaks (I) obtained against the square of the gradient strength (g^2), according to the Stejskal-Tanner Eq. 3:

$$I = I_0 e^{-D\gamma^2 g^2 \delta^2 (\Delta - \delta/3)} \quad (1)$$

which takes the general form

$$y = A e^{-Qx^2} \quad (2)$$

where Q corresponds to the diffusion coefficient multiplied by $(\gamma^2 \delta^2 (\Delta - \delta/3))$. (Here, γ is the gyromagnetic ratio, Δ is diffusion time given, and δ is the time for the gradient pulse.)

For a mixture of different components, the graph will represent a summation of different Gaussian curves (poly-Gaussian), so the equation will become:

$$\sum i = A_1 e^{-Q_1 x^2} + A_2 e^{-Q_2 x^2} + \dots + A_i e^{-Q_i x^2} \quad (3)$$

3.2.2 Practical

To fit the DOSY-NMR data of a mixture of components with an equation of this type, a sequential approach is taken.

1. Dissolve the HEWL samples in 50 mM sodium phosphate buffer (pH 7.4, 90% H₂O, 10% D₂O), but the pH and buffer of the sample should be optimized to the conditions necessary for the oligomers/aggregates (*see Note 2*).
2. Perform diffusion-ordered spectroscopy (DOSY) measurements at 293 K. The DgscSL_dpfgsc DOSY pulse program (Agilent VNMR) is used, which consists of gradient compensated stimulated echo with spin lock using the excitation sculpting solvent suppression method [12]. A spectral window of 13,020 Hz was used, with an acquisition time of 2.46 s with a relaxation delay of 3 s. The FIDs were collected with 32,000 complex data points with 64 scans. Logarithmically the gradient pulse strength was increased from 3% to 86% of the maximum strength of 32,767 G/cm in 60 steps. A diffusion time (Δ) of 100 ms and bipolar half-sine-shaped gradient pulses (δ) of 5 ms were applied. 1,4-Dioxane, which is known to behave independently of protein concentration and the folded state of the protein, was used as an internal chemical shift reference and hydrodynamic radius calibration reference (3.75 ppm; $R_H = 2.12$ Å) [11, 13]. Three replicate acquisitions were given for each sample, and the resulting diffusion coefficient (D) values calculated.
3. Upon acquisition of the NMR spectra, fit first a sample of neat dioxane with a single Gaussian to obtain the diffusion value for dioxane (Fig. 4). By referencing to dioxane it is then possible to obtain relative diffusion values/hydrodynamic radii for other species without having to correct for all the other factors contained in Q (which is instead necessary to obtain absolute measurements) (*see Note 3*).
4. Running the identical experiment and observing dioxane in a protein mixture sample clearly results in a sum of different poly-

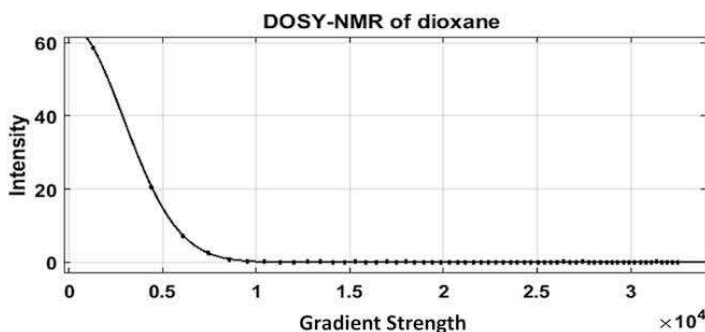


Fig. 4 DOSY-NMR fitting of pure 1,4-dioxane. Pure dioxane fit with a single curve (Eq. 2). The diffusion coefficient (un-adjusted) obtained from this fit was 5.868×10^{-8}

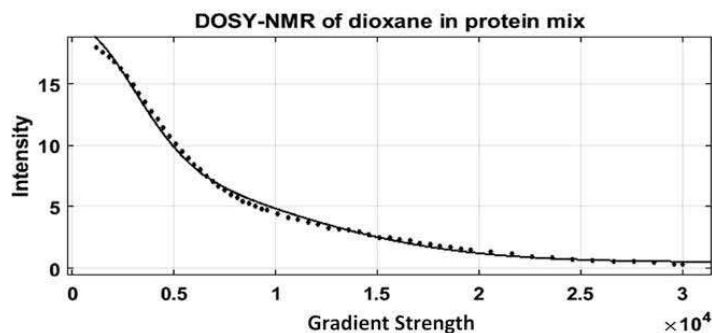


Fig. 5 DOSY-NMR fitting of 1,4-dioxane in a protein solution. Dioxane from the protein mix, fit with a poly-Gaussian (Eq. 3). In this case, it is clear that the fit could not be a single Gaussian like in the case of the pure dioxane sample, probably due to protein peaks overlapping in the dioxane region. By fixing the diffusion value of one of the components of the poly-Gaussian, it was possible to obtain a diffusion values for a second component, equal to 6.137×10^{-9}

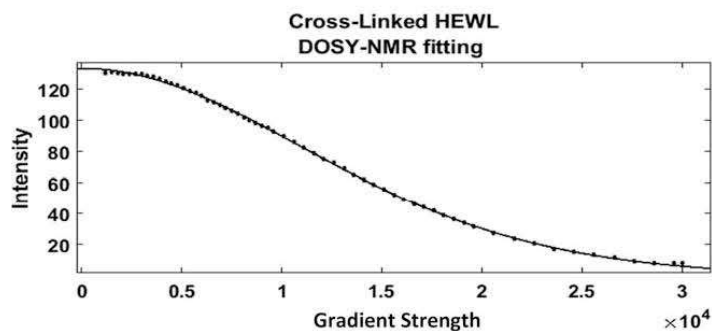


Fig. 6 DOSY-NMR fitting of cross-linked hen egg white lysozyme. Poly-Gaussian fitting of cross-linked HEWL DOSY-NMR data. A four component poly-Gaussian was chosen to approximate the population of cross-linked species present in the sample (although probably more different species than this were present, the SE-HPLC and SDS PAGE results suggested four major species to be present in the sample). The initial Gaussian was fixed at the value obtained for the second component of the dioxane with HEWL sample, and other components were added sequentially, increasing the quality of the fit while narrowing the range of values of the possible diffusion coefficients used. Diffusion coefficients obtained are presented in Table 2

Gaussian curves (Fig. 5). By fixing the diffusion coefficient of one of the values, it is possible to extract the diffusion coefficient for the second component, which (proportionally to the pure dioxane value) is comparable to that of lysozyme. The mixture of cross-linked HEWL was fit by sequentially adding curves to an initial (imperfect) fit, fixing values as fits got better and narrowing the range toward the values obtained through SE-HPLC (Fig. 6).

Table 2
Diffusion coefficients/hydrodynamic radii obtained from DOSY-NMR

Species	$D \times 10^9$ (a.u.)	R_h (Å)	I_0 (a.u.)
Dioxane	58.7	2.12	64.8
HEWL1 (trimer)	3.0	41.4	82.5
HEWL2 (dimer)	4.8	25.9	16.19
HEWL3 (mono)	6.1	20.3	34.97
HEWL4 (compact monomer)	6.5	19.1	0.4

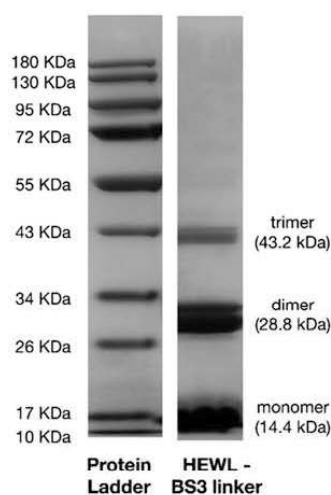


Fig. 7 SDS-PAGE gel of the selected HEWL-BS³ cross-linked mixture sample

4 Notes

1. Compared to the SDS-PAGE results, which reflect disulfide bridge-reduced and fully denatured proteins (Fig. 7), the separation of monomer, dimer, and trimer of HEWL showed that the hydrodynamic radii of these species (in their nondenatured/reduced state) were not as widely distributed (Figs. 8 and 9). Nevertheless, not only was it possible to distinguish different species in the shoulders of the peak obtained by deconvolution (Figs. 9 and 10), but one is able to identify a further-compacted monomer which appeared to have resulted from intramolecular covalent cross-linkages.

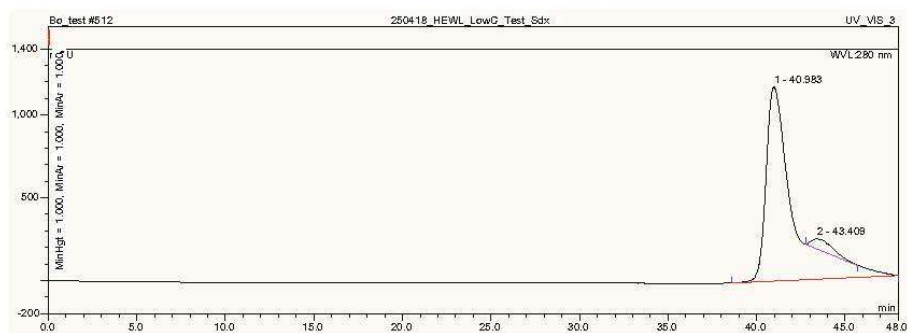


Fig. 8 Size-exclusion chromatogram of monomeric HEWL

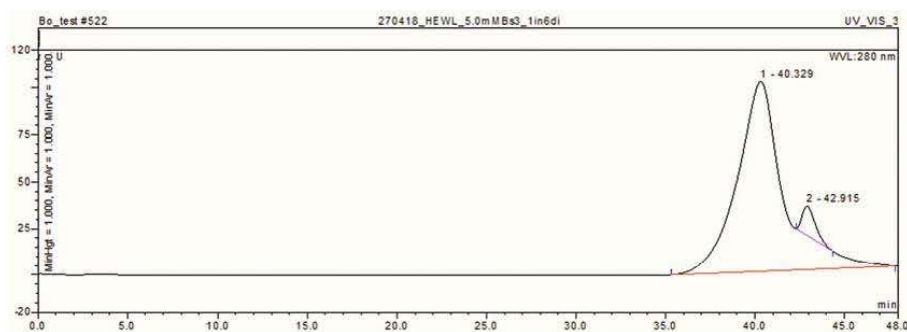


Fig. 9 Size-exclusion chromatogram of a mixture of cross-linked HEWL (trimer, dimer, and monomer)

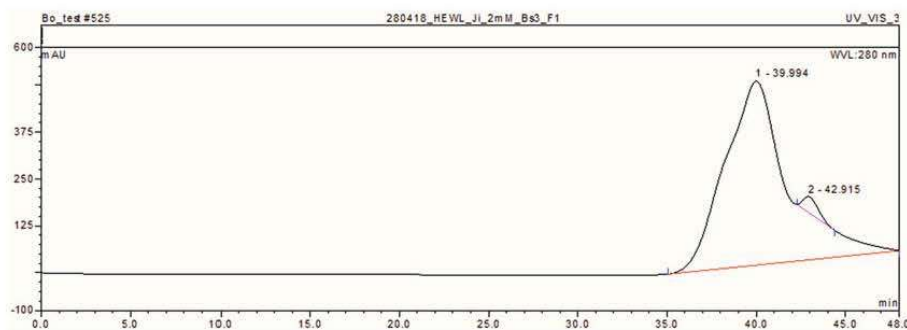


Fig. 10 Size-exclusion chromatogram of mixture of cross-linked HEWL (trimer and dimer enriched via centrifugal concentration)

2. When performing NMR, it would be advantageous to use a buffer which does not have interfering signals with the sample of interest.
3. The D (diffusion coefficient) values for 1,4-dioxane are slightly variable dependent upon the cosolute, which rightly reflects the different solution microenvironment conditions that both solutes are mutually experiencing for each sample.

References

1. Cavanagh J, Fairbrother WJ, Palmer AG III, Rance M, Skelton NJ (2006) Protein NMR spectroscopy (2nd Ed): principles and practice. Academic, New York, NY
2. Li D, Keresztes I, Hopson R, Williard PG (2009) Characterization of reactive intermediates by multinuclear diffusion-ordered NMR spectroscopy (DOSY). *Acc Chem Res* 42 (2):270–280
3. Stejskal EO, Tanner JE (1965) Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *J Chem Phys* 42:288–292
4. Dehner A, Kessler H (2005) Diffusion NMR spectroscopy: folding and aggregation of domains in p53. *ChemBiochem* 6:1550–1565
5. Pagès G, Gilard V, Martinob R, Malet-Martino M (2017) Pulsed-field gradient nuclear magnetic resonance measurements (PFG NMR) for diffusion ordered spectroscopy (DOSY) mapping. *Analyst* 142:3771
6. Morris KF, Johnson CS Jr (1992) Diffusion-ordered two-dimensional nuclear magnetic resonance spectroscopy. *JACS* 114:3139–3141
7. Barjat H, Morris GA, Smart SC, Swanson AG, Williams SCR (1995) High resolution diffusion ordered 2D spectroscopy (HR-DOSY)—a new tool for the analysis of complex mixtures. *J Magn Reson B* 108:170–171
8. Balayssac S, Delsuc M-A, Gilard V, Prigent Y, Malet-Martino M (2009) Two-dimensional DOSY experiment with excitation sculpting water suppression for the analysis of natural and biological media. *J Magn Reson* 196:78–83
9. Arora B, Tandon R, Attri P, Bhatia R (2017) Chemical crosslinking: role in protein and peptide science. *Curr Protein Pept Sci* 18 (9):946–955
10. Wu D, Chen A, Johnson CS Jr (1995) An improved diffusion-ordered spectroscopy experiment incorporating bipolar-gradient pulses. *J Magn Reson A* 115(2):260–226
11. Jones JA, Wilkins DK, Smith LJ (1997) Characterization of protein unfolding by NMR diffusion measurement. *J Biomol NMR* 10:199–203
12. Pelta MD, Barjat H, Morris GA, Davis AL, Hammond SJ (1998) Pulse sequences for high-resolution diffusion-ordered spectroscopy (HR-DOSY). *Magn Reson Chem* 36:706–714
13. Shimizu A, Ikeguchi M, Sugai S (1994) Appropriateness of DSS and TSP as internal references for (^1H) NMR studies of molten globule proteins in aqueous media. *J Biomol NMR* 4:859–862

Antonín Brisuda, James C. S. Ho, Pancham S. Kandiyal, Justin T-Y. Ng, Ines Ambite, Daniel S. C. Butler, Jaromir Háček, Murphy Lam Yim Wan, Thi Hien Tran, Aftab Nadeem, Tuan Hiep Tran, Anna Hastings, Petter Storm, **Daniel L. Fortunati** *et al.* **Bladder cancer therapy using a conformationally fluid tumoricidal peptide complex.** Nat Commun 12, 3427 (2021).

ARTICLE


<https://doi.org/10.1038/s41467-021-23748-y>

OPEN

Bladder cancer therapy using a conformationally fluid tumoricidal peptide complex

Antonín Brisuda^{1,7}, James C. S. Ho^{2,6,7}, Pancham S. Kandiyal³, Justin T.-Y. Ng⁴, Ines Ambite^{5,2}, Daniel S. C. Butler^{6,2}, Jaromir Háček⁵, Murphy Lam Yim Wan², Thi Hien Tran², Aftab Nadeem², Tuan Hiep Tran², Anna Hastings^{6,3}, Petter Storm², Daniel L. Fortunati³, Parisa Esmaeili², Hana Novotná¹, Jakub Horňák¹, Y. G. Mu⁴, K. H. Mok³, Marek Babjuk^{1,8} & Catharina Svanborg^{6,2,8}✉

Partially unfolded alpha-lactalbumin forms the oleic acid complex HAMLET, with potent tumoricidal activity. Here we define a peptide-based molecular approach for targeting and killing tumor cells, and evidence of its clinical potential (ClinicalTrials.gov NCT03560479). A 39-residue alpha-helical peptide from alpha-lactalbumin is shown to gain lethality for tumor cells by forming oleic acid complexes (alpha1-oleate). Nuclear magnetic resonance measurements and computational simulations reveal a lipid core surrounded by conformationally fluid, alpha-helical peptide motifs. In a single center, placebo controlled, double blinded Phase I/II interventional clinical trial of non-muscle invasive bladder cancer, all primary end points of safety and efficacy of alpha1-oleate treatment are reached, as evaluated in an interim analysis. Intra-vesical instillations of alpha1-oleate triggers massive shedding of tumor cells and the tumor size is reduced but no drug-related side effects are detected (primary end-points). Shed cells contain alpha1-oleate, treated tumors show evidence of apoptosis and the expression of cancer-related genes is inhibited (secondary endpoints). The results are especially encouraging for bladder cancer, where therapeutic failures and high recurrence rates create a great, unmet medical need.

¹Department of Urology, Motol University Hospital, 2nd Faculty of Medicine, Charles University Praha, Prague, Czech Republic; ²Division of Microbiology, Immunology and Glycobiology, Department of Laboratory Medicine, Faculty of Medicine, Lund University, Lund, Sweden; ³Trinity Biomedical Sciences Institute (TBSI), School of Biochemistry & Immunology, Trinity College Dublin, The University of Dublin, Dublin, Ireland; ⁴School of Biological Sciences, Nanyang Technological University, Singapore, Singapore; ⁵Department of Pathology and Molecular Medicine, Motol University Hospital, 2nd Faculty of Medicine, Charles University Praha, Prague, Czech Republic; ⁶Present address: Centre for Biomimetic Sensor Science, Nanyang Technological University, Singapore, Singapore; ⁷These authors contributed equally: Antonín Brisuda, James C.S. Ho; ⁸These authors jointly supervised this work: Marek Babjuk, Catharina Svanborg. ✉email: catharina.svanborg@med.lu.se

NATURE COMMUNICATIONS | (2021)12:3427 | <https://doi.org/10.1038/s41467-021-23748-y> | www.nature.com/naturecommunications

1

Targeted cancer therapies have made significant advances but the lack of tumor specificity remains a significant concern^{1,2}. Few current therapies kill cancer cells without harming healthy tissues, and severe side effects have become accepted as a necessary price to pay for survival or cure. The notion of successfully combining efficacy with increased tumor selectivity is justly regarded with skepticism. Yet, a serendipitous discovery has provided insights into mechanisms of tumor-specific cell death, induced by unfolded polypeptide chains, which acquire tumoricidal activity by forming fatty acid complexes^{3,4}. Extensive observations in tumor models and clinical studies have further defined the protein–lipid complexes as an interesting class of molecules with significant therapeutic potential⁵.

The findings are challenging, as protein unfolding and loss of structural definition is associated with a gain of toxicity, due to the formation of amorphous aggregates and amyloid fibrils^{6,7}. Native protein structure is often regarded as a prerequisite for biological function, by epitope-specific interactions and molecular fitness for a finite number of cellular targets. Yet, a lack of structural definition may, in some cases, result in a gain of function, in part by uncovering different conformations and exposing peptide motifs that are unavailable in the native state^{8,9}. Such effects have been predicted for membrane perturbing α -helices in antimicrobial peptides, where the ability to destabilize lipid bilayers has been proposed to reside in the three-dimensional conformation rather than the amino acid sequence¹⁰.

Alpha-lactalbumin is crucial for the survival of lactating mammals. In its native state, the protein serves as a substrate specifier in the lactose-synthase complex¹¹, defining the nutritional content and fluidity of milk. Partially unfolded alpha-lactalbumin, in contrast, forms an oleic acid complex, named HAMLET, with potent tumoricidal activity^{3,4,8,9}. The HAMLET complex kills a range of tumor cells with rapid kinetics and shows therapeutic efficacy in animal models of colon cancer, glioblastoma, and bladder cancer^{12–15}. Early, investigator-driven clinical studies demonstrated that HAMLET is active topically, against skin papilloma and induces tumor cell shedding into the urine in patients with bladder cancer^{5,16}.

This study presents a synthetic, peptide-based drug candidate derived from alpha-lactalbumin, which reproduces the tumoricidal properties of HAMLET and allows for a full translation of these findings into the clinic. Through complementary nuclear magnetic resonance (NMR) analysis and computational modeling, the molecular basis for this “gain-of-function” is defined, including the three-dimensional structural motifs that determine fatty acid-binding efficiency and tumoricidal activity. The therapeutic efficacy of the complex is demonstrated in patients with non-muscle invasive bladder cancer (NMIBC), in a fully controlled clinical trial.

Results

Peptide-specific tumoricidal activity. To understand the involvement of specific peptide motifs in tumor cell death, we synthesized the N-terminal alpha-helical domain (residues 1–39, alpha1) or the beta-sheet (40–80, beta) domains of human alpha-lactalbumin (Fig. 1a). The alpha1 peptide formed complexes with oleate (alpha1–oleate, 1:5) and circular dichroism (CD) spectra detected an increase in alpha-helical structure content in these complexes (Fig. 1b). The beta–oleate complex remained structurally unchanged (Fig. 1b). Alpha1–oleate triggered a rapid, dose-dependent death response in human lung- and kidney carcinoma cells and in murine bladder cancer cells (Fig. 1c, d). The beta–oleate complex lacked tumoricidal activity and tumor cells

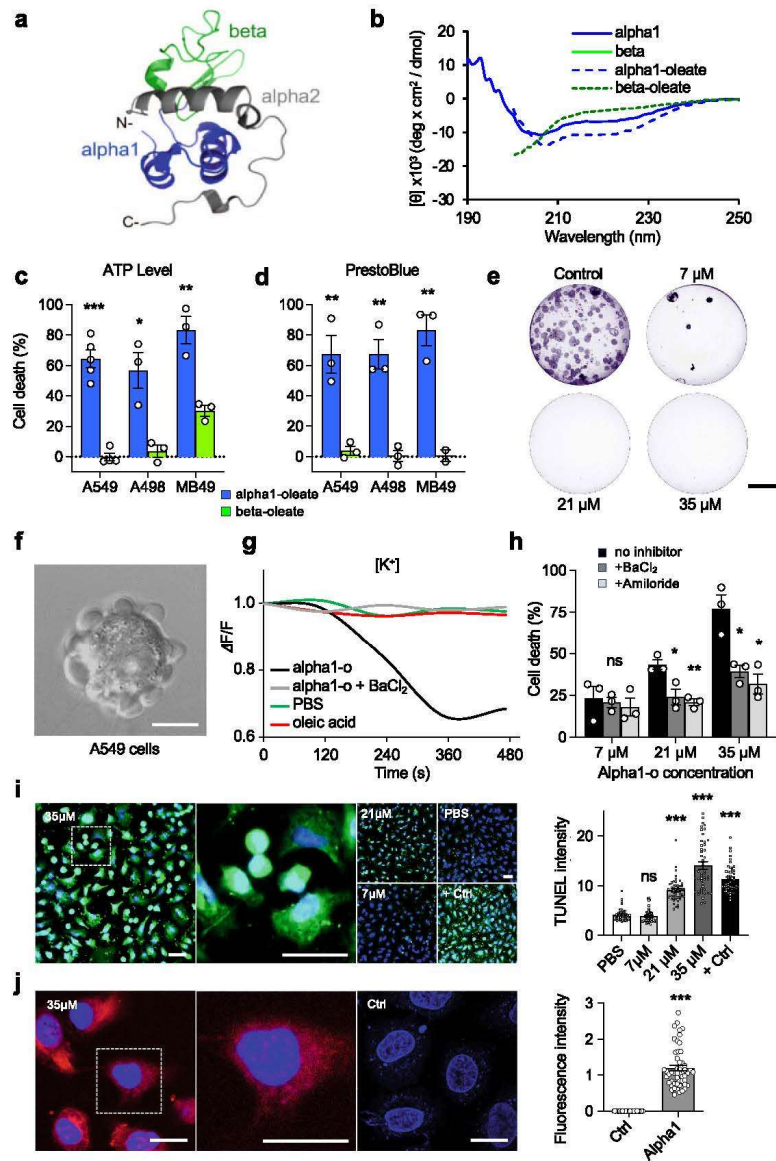
subjected to the naked alpha-helical peptides (35 μ M) or oleate (175 μ M) controls were not tumoricidal (Fig. 1c, d and Supplementary Fig. 1). The loss of cell viability was irreversible, as shown after 10 days, by using colony assays (Fig. 1e and Supplementary Fig. 1). Membrane blebbing occurred in tumor cells within minutes of exposure to alpha1–oleate but the naked peptide- and oleate controls were not active (Fig. 1f and Supplementary Fig. 1). Rapid K^+ fluxes were recorded, further defining the membrane response (Fig. 1g). Pretreatment of the cells with Na^+ and K^+ flux inhibitors reduced cell death by 40–50%, linking the membrane response to tumor cell death (Fig. 1h). The alpha1–oleate complex was rapidly internalized by tumor cells and by TUNEL staining, alpha1–oleate was shown to induce double-strand DNA breaks in the tumor cells, indicative of apoptosis (Fig. 1i, j).

In a screen of proteins with membrane-integrating properties, SARI was found to form tumoricidal complexes with oleic acid, reproducing effects of alpha1–oleate (Supplementary Figs. 1 and 2). SARI is a membrane-integrating protein of the COPII complex that induces membrane tubulation by insertion of its N-terminal amphipathic α -helix^{17–19}. The N-terminal alpha-helical peptide (residues 1–23, sar1alpha) formed an oleate complex, which efficiently killed tumor cells (Supplementary Figs. 1 and 2). The naked peptide- and oleate controls were not active (Supplementary Fig. 1). Sar1alpha–oleate triggered membrane blebbing in tumor cells, rapid K^+ fluxes were recorded, and tumor cell death was partially inhibited by Na^+ and K^+ flux inhibitors, suggesting a similar mode of action of the two complexes, despite low sequence homology (Supplementary Fig. 2). The sar1beta–oleate complexes and naked peptide controls did not trigger tumor cell death, however (Supplementary Figs. 1 and 2).

In preparation for the clinical trial, the safety of alpha1–oleate was investigated in C57BL/6J mice carrying MB49-induced bladder tumors¹³. Therapeutic efficacy in 100% of treated mice compared to the sham group and a lack of toxicity was demonstrated, providing the necessary background to plan the clinical trials¹³.

Biomolecular NMR analysis of the peptide–oleate complexes.

1H NMR spectra of the alpha1–oleate and sar1alpha–oleate complexes detected a shift from sharp signals for the naked peptides to broad signals and poor chemical shift dispersion for the oleate complexes (Fig. 2a, b), suggesting a conformational change from a random-coil fast-exchange time regime to an intermediate millisecond timescale. Broadening in the amide, side-chain methyl and aromatic regions suggests that interactions between fatty acids and peptides occur throughout the molecules. Two-dimensional nuclear Overhauser effect spectroscopy (2D NOESY) spectra identified non-covalent, relatively short through-space interactions between the respective peptides and fatty acids. Important nuclear Overhauser effects (NOEs) were detected between the olefinic protons (5.23 ppm) of oleic acid and the Ha and aromatic protons of alpha1 and between the sar1alpha aromatic region and the oleic acid olefinic protons (Fig. 2c, d). The downfield chemical shift of amide protons observed between 7.6 and 8.8 ppm suggests the presence of secondary structure in alpha1 and alpha1–oleate. Well-resolved signals obtained from the one-dimensional 1H NMR spectra provided a stoichiometry of 3.7 oleate molecules per alpha1 peptide. Chemical shift mapping revealed a cluster of residues with aliphatic side chains that change upon the binding of oleate, providing further evidence of interactions between peptides and fatty acids (Fig. 2e, f).



Hydrodynamic volume measurements carried out with size-exclusion high-performance liquid chromatography (SE-HPLC) and diffusion-ordered NMR spectroscopy (DOSY) showed that the alpha1-oleate complex R_H was considerably larger (27.4 and 29.3 Å, respectively) than the naked peptides (16.1 Å from SE-HPLC and 14.4 Å from DOSY) (Fig. 2g, h and Supplementary

Figs. 3 and 4). The distinctly larger R_2 values (transverse relaxation rate) for the complex than those of alpha1 peptide and human serum albumin (HSA) suggested slower millisecond to microsecond exchange processes (Supplementary Table 1 and Supplementary Figs. 5 and 6). Importantly, the R_2 values for the complex were also different from oleate in an aqueous solution,

Fig. 1 Tumorcidal activity of two non-homologous alpha-helical peptide-oleate complexes. **a** Ribbon representation of the crystallographically determined three-dimensional structure of human α -lactalbumin (PDB ID: 1B9O), indicating the alpha1 (blue), beta (green), and alpha2 (gray) domains. The calcium ion is not shown. **b** Far-UV circular dichroism spectra of synthetic alpha1 peptide, beta peptide, and their respective peptide-oleate complexes. **c, d** Death response in human lung (A549), kidney (A498), and murine bladder (MB49) carcinoma cells, quantified as a reduction in ATP levels (**c**, $P = 3.26 \times 10^{-5}$ for A549, 0.013 for A498 and 0.005 for MB49, alpha1-oleate compared to beta-oleate) or PrestoBlue fluorescence (**d**, $P = 0.007$ for A549, 0.003 for A498 and 0.002 for MB49, alpha1-oleate compared to beta-oleate). Cells were treated with the alpha1-oleate complex (blue) or the beta-oleate complex (green), (3 h, 35 μ M, cell death compared to PBS controls). For controls exposed to the naked peptides or oleate alone, see Supplementary Fig. 1d. **e** Colony assay showing dose-dependent long-term effects of alpha1-oleate. A representative image is shown from two independent experiments. Scale bar = 5 mm. **f** Alpha1-oleate triggers rapid membrane blebbing in A549 lung carcinoma cells (35 μ M, 10 min). Scale bar = 10 μ m. A representative image is shown from three independent experiments. **g** K^+ efflux in A549 lung carcinoma cells exposed to alpha1-oleate and inhibition with BaCl₂. **h** Inhibition of cell death by the ion flux inhibitors Amiloride and BaCl₂ (100 μ M), measured by PrestoBlue fluorescence ($P = 0.031$ for 21 μ M + BaCl₂, 0.005 for 21 μ M + Amiloride, 0.028 for 35 μ M + BaCl₂ and 0.014 for 35 μ M + Amiloride, compared to no inhibitor). **i** DNA strand breaks detected by TUNEL staining in alpha1-oleate-treated A549 lung carcinoma cells ($n = 50$ cells per group). Scale bar = 20 μ m. **j** AlexaFluor568-labeled alpha1-oleate (red) is internalized by A549 lung carcinoma cells. Nuclei are counterstained with DAPI (blue) ($n = 52$ cells per group). Scale bar = 10 μ m. Data are presented as mean $\pm$ SEM from three independent experiments, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, analyzed by two-tailed unpaired t-test (**c, d, h, j**) and 2-way ANOVA using Dunnett's correction (**i**).

suggesting that the dynamics of the complex were clearly different from micelle/vesicle-like particles formed from oleic acid/oleate with no peptide component.

Computational analyses of the peptide- and peptide-oleate system. Computational simulations also pointed toward structural heterogeneity, showing that the naked peptides and peptide-oleate complexes belonged to wide conformational spaces with relatively deep basins (Fig. 3a, b). Representative structures mapped to different free energy surface minima, revealed prominent alpha-helical secondary structural elements and a hydrophobic oleate core for the peptide-oleate complexes (Fig. 3c, e). The peptide-oleates fold upon this core differently from the naked peptides, which exhibit multiple local minima (Fig. 3d, f, and Supplementary Tables 2 and 3). Naked alpha1 ensembles were characterized by various partially folded helix-turn conformations, whereas naked sar1alpha ensembles exhibited a mixture of the random coil, alpha-helical, and beta-sheet structures.

A contact probability analysis revealed that the interactions between alpha1 or sar1alpha and oleate were mainly hydrophobic, with a >0.9 contact probability with olefinic protons (Supplementary Tables 4 and 5). The peptide-oleate complexes displayed relatively wide and deep free energy minima basins, suggesting that a multitude of conformations would be equally possible to visit (Fig. 3c, e). When combined with the R_2 relaxation rates, the possibility of multiple sampling of various conformations within a short period of time provides an argument that rather than targeting specific partners, these alpha-helical complexes may potentially be interacting with multiple putative binding partners available on cancer cell surfaces²⁰.

Based on these extensive investigations and the strong agreement of the experimental aspects with the simulated predicted ensembles, it was clear that the apparently unrelated peptides alpha1 and sar1alpha can form complexes with shared structural characteristics, involving a flexible peptide moiety and a fatty acid cluster. This notion resonates with the sequence-function inconsistency among certain antimicrobial amphipathic alpha helices, where peptides with similar overall properties, such as hydrophobicity or charge, can have dramatically different levels of activity¹⁰.

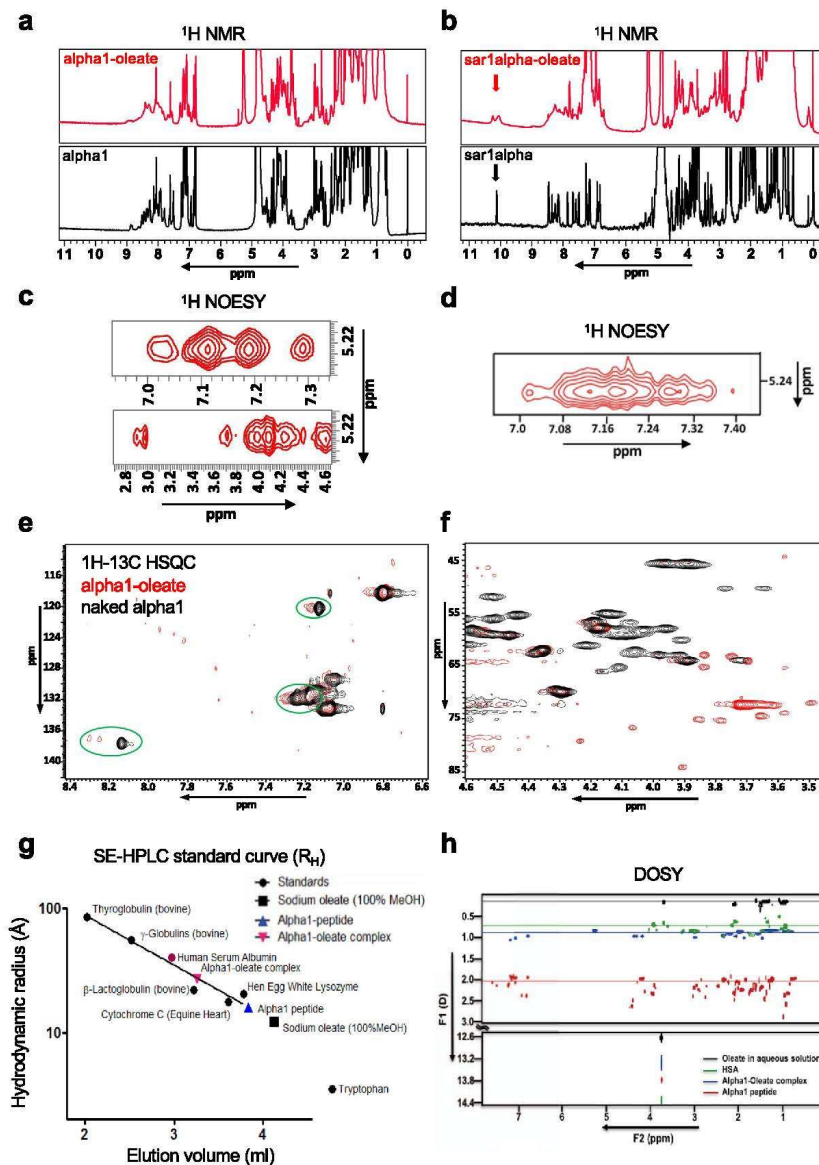
A placebo-controlled, randomized clinical trial of alpha1-oleate in patients with NMIBC. NMIBC is common and despite current treatment protocols, recurrence rates are high^{21,22}. To address if the therapeutic effects observed in the murine MB49 bladder

cancer model can be translated into the clinic, the investigational product alpha1-oleate was produced under GMP conditions. The alpha1-oleate complex was further subjected to formal toxicity testing and the results have been published¹³. Toxicity for bladder tissue was not detected at concentrations ranging from 1.7 to 17 mM¹³.

The clinical safety and therapeutic potential of alpha1-oleate were tested in a single-center, placebo-controlled, double-blind Phase I/II trial (EudraCT No: 2016-004269-14, ClinicalTrials.gov NCT03560479, Supplementary Note 1, Supplementary Table 6). Patients with suspected NMIBC were randomized 1/1 to receive alpha1-oleate or placebo during a period of 22 days, prior to endoscopic removal of the tumor by transurethral resection (TURB), (Fig. 4a, b). Alpha1-oleate (1.7 mM) or placebo (PBS) was administered intravesical on six occasions (30 mL, days 1, 3, 5, 8, 15, and 22). The placebo solution was identical in appearance to the active treatment. Demographic data, medical history, and vital signs did not differ between the treatment and placebo groups (for details see Supplementary Table 7).

Primary study endpoints. Adverse events (AEs) were recorded and coded according to MedDRA (version 21.1) with a safety follow-up after 52 days (Supplementary Table 8). Procedure-related AEs, such as dysuria and bacteriuria, occurred at a similar rate in the treatment and placebo groups. AEs specific for the treatment group were not detected, suggesting low toxicity of the study medication (Fig. 4c). Furthermore, there was no evidence of a toxic response in healthy tissue samples from patients treated with alpha1-oleate, defined by histopathology or TUNEL staining.

Tumor cell shedding and release of tumor cell clusters were recorded. Cells with uroepithelial morphology were quantified in urine at each visit, before instillation and about 2 h after the instillation of alpha1-oleate or placebo. Alpha1-oleate triggered a rapid increase in cell shedding compared to the pre-instillation sample in all treated patients, at all visits (Fig. 5a–c and Supplementary Fig. 7). In addition, tumor cell clusters were released into the urine in the treatment group. The clusters were relatively large and the presence of tumor stroma in some samples supported their tumor origin (Fig. 5d–f). The cells shed in urine were assigned a pathology score as per the Paris classification (classes 1–6, urine cytology was a secondary endpoint). In the treatment group, an increase in the Paris score was detected in post-inoculation samples compared to pre-inoculation samples (Fig. 5g). A low level of cell shedding in the placebo group was attributed to the instillation procedure and changes in pathology



score or shedding of cell clusters were not observed (Fig. 5a–g and Supplementary Fig. 7).

The endoscopic appearance of the tumors was recorded at the time of diagnosis and prior to surgery using a flexible cystoscope with white light-band and narrow-band imaging. Sizes were assessed by an experienced endourologist using fully opened

clamps of the flexible forceps and a measuring device close to the tumor. Paired images from 39 patients were evaluated in a blinded manner, by an independent NMIBC expert using a simplified Delphi method²³ addressing changes in lesion size, superficial necrosis, and tissue vascularization. A reduction in lesion size was detected in the treatment group ($n = 19$, Fig. 5h, i).

Fig. 2 Biomolecular NMR analysis of naked alpha1- and sar1alpha peptides and their oleate complexes. **a, b** One-dimensional ^1H NMR spectra. The naked alpha1- (**a**, black) and sar1alpha- (**b**, black) peptides assume an ensemble of structures that interconvert rapidly and are therefore seen as sharp peaks. The alpha1-oleate (**a**, red) and sar1alpha-oleate complexes (**b**, red) show broader peaks. Arrows indicate the indole ^1H signals arising from the three Trp side chains present in the sar1alpha peptide. **c, d** Two-dimensional ^1H NOESY spectra of alpha1-oleate and sar1alpha-oleate complexes, showing atomic-level proximities of the fatty acid to the respective peptide. The spectra highlight NOEs between the 9,10 olefinic protons (5.23 ppm) of oleic acid with the H α protons and aromatic protons of the alpha1-oleate complex (**c**) and the sar1alpha-oleate complex (**d**). **e, f** Two-dimensional ^1H - ^{13}C Heteronuclear Single Quantum Coherence (HSQC) spectra overlays of the alpha1 peptide (red) and alpha1-oleate complex (black). Chemical shift perturbation is detected in the aromatic side-chain region and the imidazole ring protons (**e**, green circled regions), and in the aliphatic side chain regions (**f**). **g** Size-exclusion HPLC (SE-HPLC) of the alpha1 peptide and the alpha1-oleate complex, mapped onto a standard calibration curve. **h** Diffusion-ordered NMR spectroscopy (DOSY) of the alpha1 peptide, alpha1-oleate complex, human serum albumin (HSA), and oleate suspension.

In 1 patient, paired pre-treatment and post-treatment images could not be collected for technical reasons. No difference in superficial necrosis or vascularization was observed and there was no change in lesion size in the placebo group ($n = 20$).

Secondary end points. Evidence of tumor cell apoptosis was obtained by staining of tumor biopsies obtained at the time of surgery. Biopsies were examined for evidence of treatment-induced apoptosis, quantified by the TUNEL assay (Fig. 6a–c and Supplementary Fig. 8). A significant increase in net mean fluorescence was detected in the treatment group, compared to placebo (see also Fig. 1). Staining was most intense adjacent to the lumen suggesting that a gradient might be formed, from the lumen towards the center of the tumor. TUNEL staining intensity was significantly correlated to cell shedding and alpha1-oleate uptake in individual patients (Fig. 6d) but not to the tumor grade. In healthy tissue biopsies from the treated patients, TUNEL staining was low. Tumors from the placebo group did not show increased TUNEL staining, suggesting that tumor cell apoptosis may be induced by the alpha1-oleate treatment (Fig. 6c).

The alpha1-oleate content of shed cells in urine was quantified by immunohistochemistry, using alpha1-specific antibodies. Alpha1-staining was detected in 70% of post-inoculation samples in the treatment group (Fig. 6e, f and Supplementary Fig. 9). Uptake correlated with cell shedding and cluster grade but not with the tumor grade or stage (Supplementary Fig. 9).

The response to alpha1-oleate was further evaluated by RNA-seq, using RNA from tumor biopsies and comparing the treatment to the placebo group. A strong treatment effect was detected (Fig. 7a–c and Supplementary Fig. 10). Cancer-related genes accounted for about 80% of the significantly regulated genes in the treated patients (cut off fold change > 2.0 , $P < 0.05$), confirming the effect of alpha1-oleate on the tumor environment. Genes regulating tumor growth and invasion were inhibited and Ras signaling was suppressed, consistent with known effects of the complex on Ras family members²⁴ (Fig. 7d and Supplementary Fig. 10). Bladder cancer genes were specifically regulated, including metalloproteinases, solute carriers, WNT complex constituents, and thrombospondin, which affects angiogenesis²⁵ (Fig. 7e). Furthermore, Fatty Acid Desaturase 6 (*FADS6*) and transcriptional activator *CREB3L4* were affected, suggesting that the tumors respond to the constituents of the alpha1-oleate complex. *FADS6* regulates oleate biosynthesis and *CREB3L4* the unfolded protein response to conformationally fluid proteins, such as the alpha1 peptide. Interesting targets also included the gap junction alpha1 protein, which was inhibited, potentially promoting cell detachment (*GJA1/CXA1* encoding Connexin 43, Supplementary Fig. 10). No difference in tumor grade was observed between the treatment or placebo groups, defined by WHO 1973 and 2004/2016 criteria (Supplementary Table 9). Data regarding two secondary end-points are not reported. As this is an interim analysis, the long-term treatment effects will be

evaluated when the entire study has been completed. The urine proteomics data set has not been fully analyzed.

Discussion

Bladder cancer is the fourth most common malignancy in the United States and the fifth in Europe, with a prevalence of about 1/4000²⁶. Due to high recurrence rates and a lack of curative therapies, “bladder cancer has the highest lifetime treatment costs per patient of all cancers, followed by colorectal-, breast-, prostate-, and lung cancer”²⁷. More than 80% recur after complete surgical removal of the first tumor and 15% progress to muscle-invasive disease²⁸. Intravesical chemotherapy and Bacillus Calmette-Guérin (BCG) immunotherapy have limited efficacy and significant side effects^{29,30}. Systemic administration of PD-1 and PD-L1 inhibitors is considered only in BCG unresponsive patients where the experience is limited. Therapeutic options are also limited by the inadequate supply of immunotherapy and chemotherapy drugs worldwide³¹. In this study, we identify conformationally fluid peptide–fatty acid complexes as additional tools in cancer therapy and show that intra-vesical inoculation of alpha1-oleate is safe and effective in patients with bladder cancer.

The tumor response to alpha1-oleate was analyzed in-depth, using cellular and molecular tools to detect changes induced by the complex. Treatment triggered the shedding of cells and tissue fragments into the urine and alpha1-oleate internalization by tumor cells confirmed the affinity of the complex for the tumor. Further analysis of tissue biopsies suggested a lasting effect of the alpha1-oleate instillations, as several tumor samples showed a gradient-like pattern of apoptosis, starting from the bladder lumen. Dysfunctional apoptosis has been identified as a key to tumor development, especially in environments where oncogenes such as MYC drive tumor cell proliferation³². Numerous attempts have been made to develop apoptosis-inducing therapeutics with tumor selectivity, but this has proven challenging, probably due to the heterogeneity of individual tumors as well as their intrinsic resistance to activating cell death pathways. The ability of alpha1-oleate to stimulate apoptosis in the majority of bladder tumors is, therefore, encouraging and consistent with the apparent lack of toxicity for bladder tissue.

RNA sequencing revealed profound molecular changes in treated tissues, attributable to alpha1-oleate. Classical cancer gene networks were strongly inhibited in the treated patients, compared to the placebo group, including Ras, previously identified as a target for HAMLET; the oleate complex formed by the alpha-lactalbumin holoprotein²⁴. HAMLET binds activated Ras at the plasma membrane of tumor cells and inhibits the Ras signaling pathway, in part through effects on b-Raf phosphorylation. Significant effects on adaptive immunity were not detected and innate immunity was largely inhibited, including granulocyte activation pathways. Notably, genes involved in oleate metabolism and the unfolded protein response were affected, possibly reflecting a direct response to the constituents of the alpha1-oleate complex. In addition, treatment inhibited *GJA1*, a

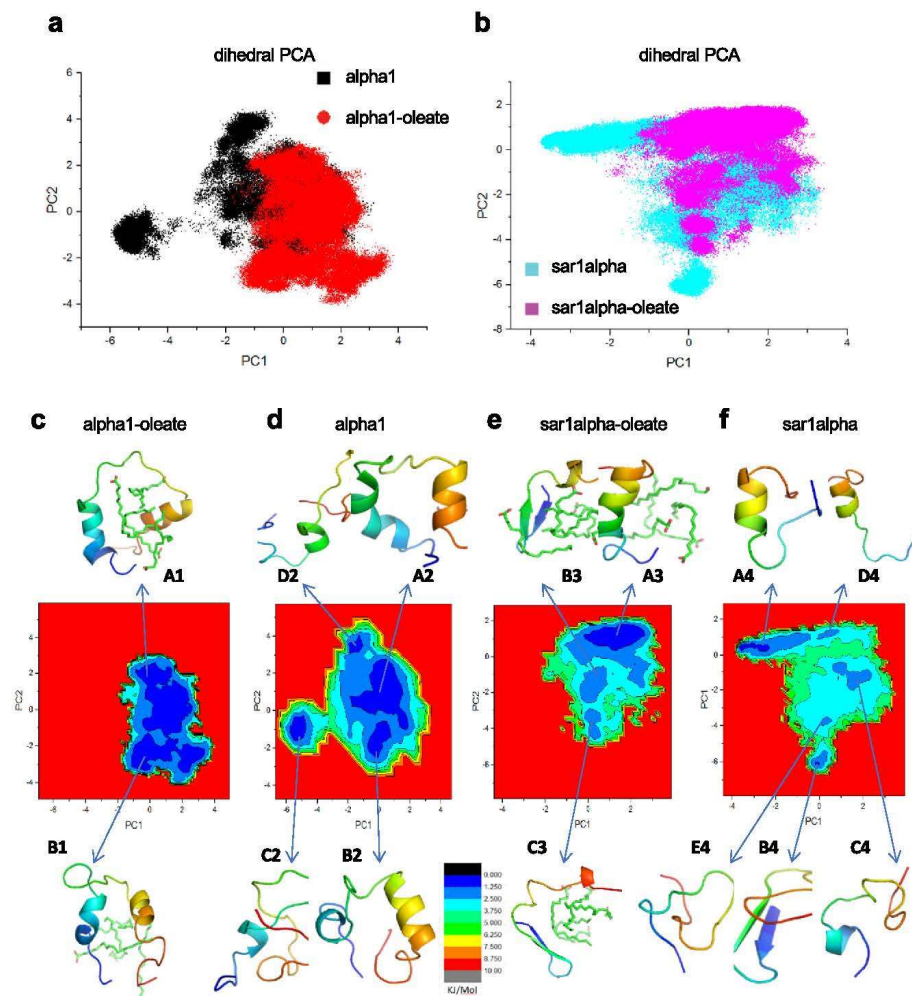


Fig. 3 Free energy surface analyses of the peptide- and peptide-oleate system. **a, b** Superimposition of dihedral principal component analysis (PCA) plots of alpha1 (black points) and alpha1-oleate (red points) systems (**a**), and sar1alpha (cyan points) and sar1alpha-oleate (magenta points) systems (**b**). Principal component (PC)1 and PC2 represent the axes of the two greatest variances after mathematical transformation for dimension reduction. **c-f** Free-energy surfaces as a function of the first two principal components for alpha1-oleate (**c**), naked alpha1 (**d**), sar1alpha-oleate (**e**), and naked sar1alpha (**f**). The representative structures of peptides or peptide-oleate complexes, along with their respective local minima annotations, are colored from the N termini (blue) to the C termini (orange/brown). The free-energy surface of the alpha1-oleate complex contains 2 minima basins, A1 and B1, with A1 representing the major conformational ensemble. The free-energy surface of the sar1alpha-oleate complex contains 3 minima basins, A3, B3, and C3 (with the A3 basin harboring the major structural ensemble), which are characterized by a prominent alpha-helical secondary structural element, as shown from simulation calculated alpha-helical propensities. By contrast, the free-energy surface of the naked sar1alpha shows large structural heterogeneity. Here, minima basins A4 and D4 are represented by helical structures, B4 by beta structure, and C4 and E4 by random coil structures.

gap junction protein that has been proposed to promote cancer development and metastasis^{33,34}. The effect occurred specifically in tumor tissue, potentially providing the alpha1-oleate complex with a mechanism to trigger cell shedding, as observed here. It is

interesting to speculate that cell shedding may serve as a “tip-of the iceberg” marker of the profound changes in tumor biology that include activation of programmed cell death, transcriptional reprogramming, and inhibition of tumor progression.

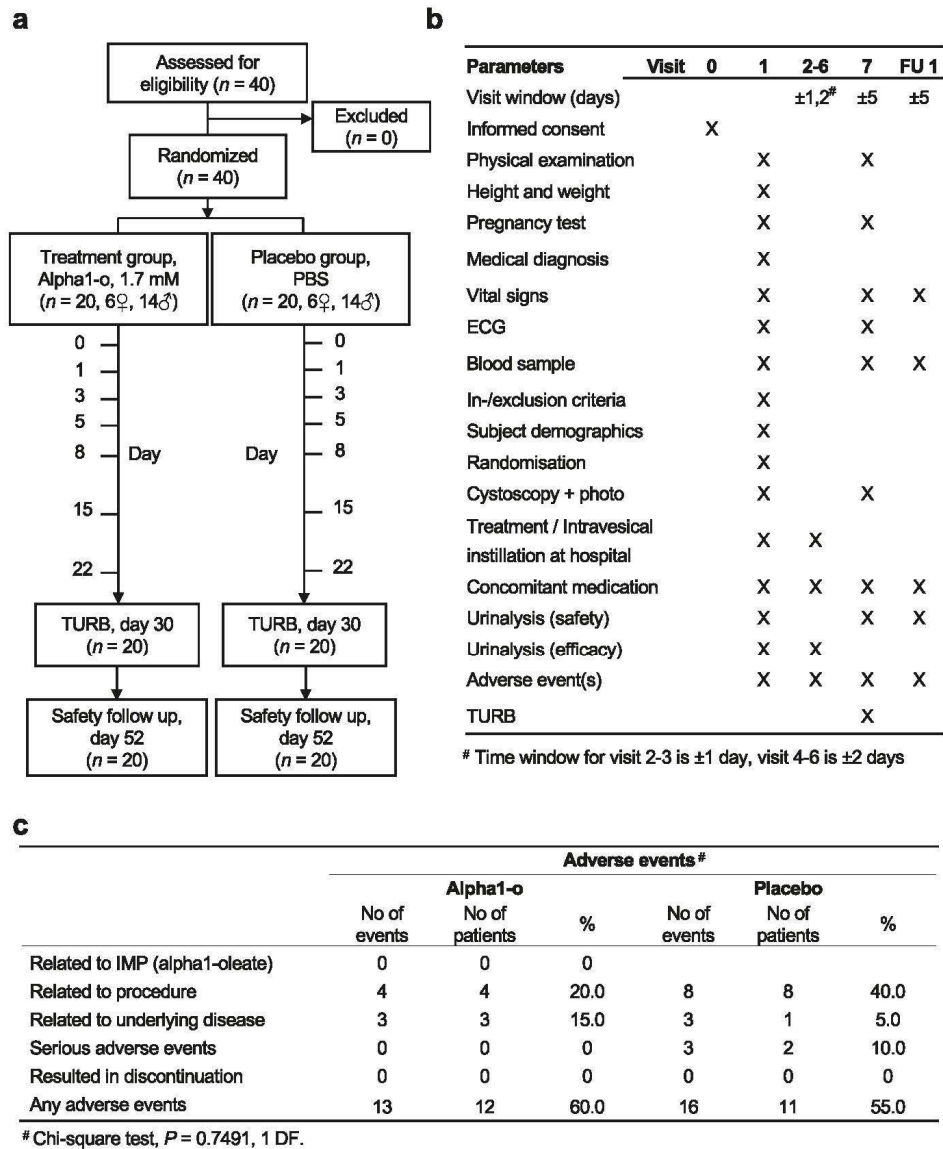
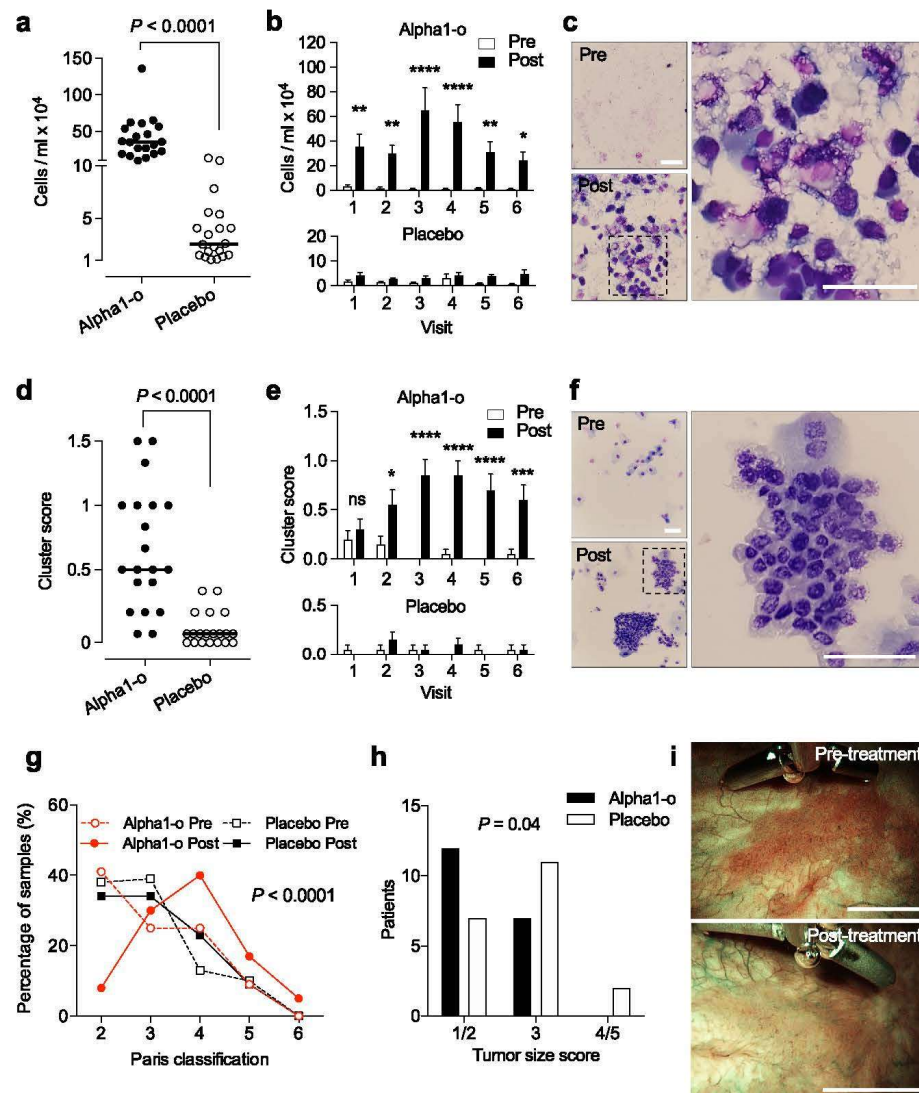


Fig. 4 Clinical study protocol, demographic data, and adverse events. a Study CONSORT diagram. **b** Study protocol. After diagnosis and informed consent, the subjects received intravesical instillations of either alpha1-oleate or placebo on six occasions during one month preceding a scheduled transurethral resection (TURB). A safety follow-up was performed 52 days after the first instillation. **c** Number of adverse events (AEs) in the active and placebo groups. No drug-related adverse events were recorded. There were totally 29 AEs reported by 12 subjects in the active group and by 11 subjects in the placebo group. None of the AEs were related to the investigational product. One AE was severe and two were moderate in the placebo group. The active group had one moderate AE. Two subjects in the placebo group reported severe AE (SAEs). The AEs were evaluated descriptively, and the AE profiles were similar between the placebo and the active groups.



Alternative therapeutic tools are actively being developed and tested in patients with NMIBC, particularly in patients with disease recurrence after BCG treatment^{35–37}. Device-assisted hyperthermia was shown to increase the efficacy of intra-vesical chemotherapy but treatment was accompanied by side effects, reducing compliance^{38–40}. An oncolytic-virus-based intra-vesical therapy was recently reported to achieve a complete response in 53.4% of patients with BCG-unresponsive carcinoma in situ, in a phase III trial⁴¹. The authors discuss the assessment of side

effects and the development of biomarkers to help select patients suitable for this therapy. In patients with BCG unresponsive disease, treated with systemic Pembrolizumab, a 41% response rate was reported but side effects were prevalent, limiting compliance (Keynote-676 trial³⁵). The present study identifies alpha1-oleate as an active drug candidate with low toxicity. Further dose-finding clinical studies and adjuvant therapy protocols will be essential to define the therapeutic window of this complex.

Fig. 5 Primary endpoints: shedding of tumor cells and reduction in tumor size following intra-vesical instillation of alpha1-oleate. **a–c** Cell shedding increased significantly after alpha1-oleate instillation. **a** Scatterplot showing individual means of six visits per patient in the treatment group ($n = 20$) compared to patients receiving placebo ($n = 20$). Line represents the median. **b** Comparison of cell numbers in urine before (pre = white) and after (post = black) alpha1-oleate inoculation on visits 1–6 showing increased cell numbers post-inoculation in the treatment group ($n = 20$ patients per group, $P = 0.0030$ for visit 1, 0.0098 for visit 2, <0.0001 for visits 3 and 4, 0.0073 for visit 5 and 0.0336 for visit 6) but not in the placebo group. Data are presented as mean $\pm$ SEM. **c** Representative images, illustrating the increase in cell shedding after alpha1-oleate instillation. Magnification = $\times 400$. Scale bar = $50 \mu\text{m}$. **d–f** Difference in the shedding of tumor cell clusters between the treatment and placebo group. **d** Scatterplot showing individual means of six visits per patient in the treatment group compared to patients receiving placebo. Line represents the median. **e** Increased numbers of cell clusters in post-inoculation samples of patients receiving alpha1-oleate ($n = 20$ patients per group, $P = 0.9743$ for visit 1, 0.0212 for visit 2, <0.0001 for visits 3, 4, and 5, and 0.0005 for visit 6). Data are presented as mean $\pm$ SEM. **f** Representative images of cancer cell clusters after alpha1-oleate instillation. Magnification = $\times 400$. Scale bar = $50 \mu\text{m}$. **g** Paris grade of shed cells before or after alpha1-oleate instillation. An increase is observed in the treatment group (χ^2 test). **h** Reduction in tumor size in patients receiving alpha1-oleate treatment. Images were compared between the time of diagnosis and the time of TURB ($P = 0.04$, χ^2 test for trend compared to placebo, $n = 19$ for treatment group and $n = 20$ for placebo group). **i** Examples of cystoscopy photographs obtained by A.B. at the time of diagnosis and after treatment at the time of TURB. Scale bars = 5 mm . * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$. The data were analyzed by two-tailed unpaired Mann–Whitney U -test (**a**, **d**) or by repeated-measures two-way ANOVA with Sidak's correction (**b**, **e**).

Cancer cells are aggressive, outcompete healthy cells, and ruin tissue integrity. It is generally assumed that treatments must be equally aggressive and highly toxic substances are often used, despite their lack of selectivity and the severe side effects that they cause. The protein–lipid complexes studied here are attractive to cancer cells, which actively internalized them, but end up being killed. Healthy cells are less responsive and extensive toxicity studies have failed to detect adverse effects in the bladder¹³. This low toxicity was confirmed here, as no drug-related side effects were observed in the treatment group. The results, therefore, identify the alpha1-oleate treatment of NMIBC as an interesting therapeutic option. In view of the low toxicity observed so far, liberal intra-vesical administration in early-stage NMIBC might be an interesting approach to postponing the introduction of more toxic and invasive therapeutic options.

Methods

Clinical trial of alpha1-oleate in NMIBC

Trial design and patient population. This was a single-center, placebo-controlled, double-blinded randomized Phase I/II interventional clinical trial of non-muscle invasive bladder cancer taking place from May 21, 2018 to June 3, 2019. Subjects diagnosed with non-muscle invasive bladder cancer and scheduled for transurethral surgery were included in the study. The study was approved by the State Institute for Drug Control (SUKL) in the Czech Republic; number 273799/17-I and the Ethics Committee of the Motol University Hospital; number EK-786/17 (ClinicalTrials.gov Identifier: NCT03560479). Patients gave their written informed consent.

Peptide synthesis and alpha1-oleate complex generation. Peptide synthesis and the preparation of the investigational product were performed under good manufacturing practice (GMP) conditions and the complex was diluted in phosphate-buffered saline (PBS) to the final concentration (1.7 mM). The placebo group received PBS (sodium chloride, potassium chloride, sodium- and potassium phosphate, and water for injection), which was identical in appearance to the active treatment.

Study protocol. Study subjects were randomized 1/1 and received intra-vesical instillations (30 mL) of either alpha1-oleate (1.7 mM) or PBS on six occasions during a period of 22 days (days 1, 3, 5, 8, 15, and 22). A safety follow-up was included 52 days after the last instillation (EudraCT Number: 2016-004269-14 and ClinicalTrials.gov NCT03560479). The complete Study Protocol is provided as Supplementary Note 1 in the Supplementary Information file. The interim analysis of the clinical trial presented here represents a complete evaluation of the Phase I/II study, as per the original protocol. The protocol was later amended to include a 24-month follow-up. At that time, it was decided to perform the initially planned final analysis as an interim analysis, scheduled after all subjects had completed the 52-day safety follow-up. The study underwent data lock and subsequent unblinding was under third-party control. The primary objective of the trial was to evaluate the safety of alpha1-oleate. No formal sample size calculation evaluating the power of the trial has been performed. However, consideration regarding the sample size was made based on a previous open study of HAMLET instillations in bladder cancer patients¹⁶ and in the murine bladder cancer model¹⁴. For efficacy, the sample size was based on analysis of change in tumor cells assessed before HAMLET

instillation and after 2 h. The mean fold increase of shed cells was 41.3 and the standard deviation was 60.4 in 9 patients¹⁶. A sample size of 20 patients per group was deemed suitable to achieve criterion for significance (alpha) 0.05 and power 90% using the paired samples 1-tailed t -test. The null hypothesis is H_0 : mean change in cell shedding = 0 and the alternative hypothesis is H_A : mean change in cell shedding > 0 .

Demographic data, morbidity and health parameters as well as tumor characteristics were recorded by the study physicians in the electronic Case Report Form (eCRF) and closely monitored by an external monitor. Population characteristics were evaluated by the study statistician. No significant differences between the treatment and placebo groups were registered in terms of age, gender, co-morbidity, or tumor parameters.

Primary endpoints. - Safety as AEs profile (time frame: from the signing of informed consent (day 1) and until end of study (day 52)); Incidence of AEs and classification in terms of severity, causality, and outcome.

- Efficacy as cell shedding (time frame: days 1–22): change in cell shedding into urine (number of epithelial cells per mL of urine).

- Change from baseline in characteristics of papillary tumors (time frame: prior to treatment (baseline) and on day 30, in connection with scheduled surgery); the bladder tumors are characterized by in vivo imaging during examination by cystoscopy.

Secondary endpoints. - Histopathology scoring of the tumor using established parameters for scoring of Grade and Stage/Invasiveness.

- Urine cytology examined before and after instillation, using the Paris scoring system.

- Uptake of alpha1-oleate by tumor cells, defined by staining with specific antibodies.

- Tissue apoptotic response to alpha1-oleate, defined by TUNEL staining.

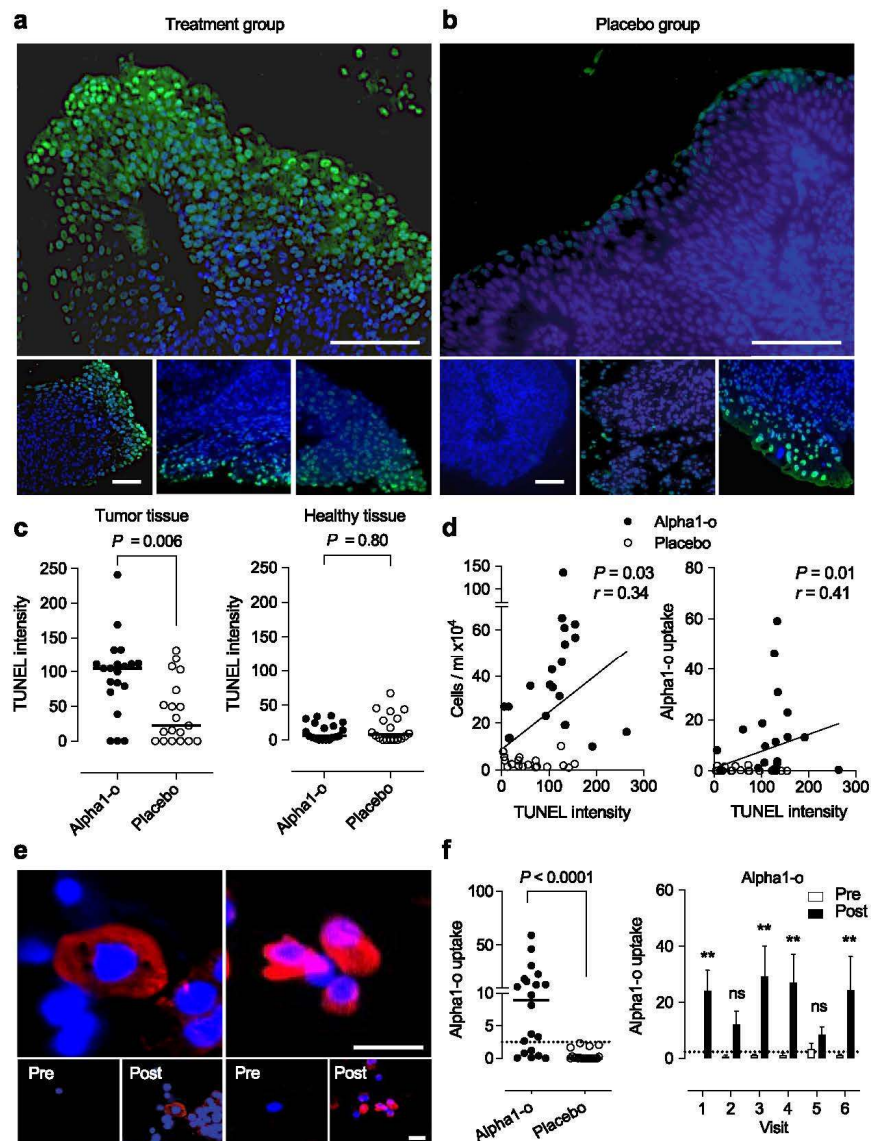
- Tumor response to alpha1-oleate, defined by RNA sequence analysis.

- Proteomic analysis of markers in urine was not completed.

- Long-term effects of the study treatment have not been evaluated.

AEs profile. AEs were collected from the signing of the informed consent form until the end of the study (PU1 Visit, day 52). All diagnoses, symptom(s), sign(s), or finding(s) with a start date after the first dose of the study drug were recorded as AEs or severe AEs (SAEs). (S)AEs related to the study procedure were coded during the course of the trial according to MedDRA by preferred terms and primary system organ class. All AEs recorded during the course of the trial were included in the subject data listings and an overall summary of the number (percentage) of subjects with any treatment-emergent (S)AEs, premature discontinuations from the trial due to AEs, treatment-related AEs, and SAEs were constructed. The number of subjects experiencing each type of adverse event was tabulated regardless of the number of times each adverse event was reported by each subject. The severity of each type of adverse event was also tabulated and graded as the most severe recording for that adverse event.

Cell shedding. To quantify the shedding of cells and cell clusters into the urine, samples were obtained from each patient prior to and after each instillation of alpha1-oleate or placebo (Visit 1–Visit 6). Cell shedding was quantified by counting the total number of epithelial cells in a unit of uncentrifuged urine under light microscopy, using a hemocytometer chamber. Changes in cell shedding were quantified at each visit, by comparing cell numbers in samples obtained before and after each instillation. The cell clusters were scored based on the examination of these samples by an experienced pathologist on a range of 0–2 where 0 = no clusters and 2 = the highest number of clusters.



Characteristics of papillary tumors, tumor size. To examine if the alpha1-oleate treatment affects tumor size, all included subjects underwent outpatient cystoscopy at Visit 0. Tumors were reexamined at Visit 7, prior to scheduled surgery. High-quality photographs were collected endoscopically, using a flexible cystoscope (Olympus) before being removed by TURB according to EAU Guideline recommendations⁴². Changes in tumor size were evaluated intra-individually, using paired images. The results were evaluated using a simplified Delphi

method²³, by an independent NMIBC expert. Changes in lesion size, superficial necrosis, and tissue vascularization were addressed in a blinded manner.

Histopathology scoring. Tumor biopsies, collected at the time of surgery were evaluated by histopathology, using established parameters for scoring of Grade and Stage/Invasiveness. Tissue samples were analyzed by a designated study uro-pathologist. Both grading classifications (WHO 1973 and 2004/2016) were used⁴³.

Fig. 6 Apoptotic response to alpha1-oleate and cellular uptake by tumor cells. Apoptosis was quantified in tumor biopsies, using the TUNEL assay. Arbitrary units were calculated after subtraction of background staining in TUNEL negative healthy tissue samples. **a** Representative image of TUNEL staining (green = TUNEL, blue = DAPI) in tumor tissue from individual patients receiving alpha1-oleate instillations. Scale bars = 200 μ m. **b** Representative images of TUNEL staining in tumor tissue from individual patients receiving placebo. Scale bars = 200 μ m. **c** Scatter plot demonstrating elevated TUNEL staining intensity in tumor biopsies from patients receiving alpha1-oleate instillations compared to placebo. TUNEL staining was not significantly altered in healthy tissue biopsies from patients receiving alpha1-oleate instillations or placebo ($n = 40$ tumors and 38 healthy biopsies, two data points were further removed due to medical conditions from patients and confirmed by Grubbs's outlier test) (two-tailed unpaired Mann-Whitney U -test). Line represents the median. **d** Correlation of TUNEL staining intensity with cell shedding ($P = 0.03$, 95% CI 0.0220–0.6010) and alpha1-oleate uptake ($P = 0.01$, 95% CI 0.0957–0.6461) (Spearman correlation, two-tailed, approximate P -value, $n = 20$ for alpha1-o and $n = 19$ for placebo). **e** Representative images of alpha1-oleate (red) uptake with counter-stained nuclei (blue). Alpha1-oleate uptake by tumor cells was quantified by staining of shed cells in urine with polyclonal anti-alpha1-oleate antibodies. Scale bars = 20 μ m. **f** Scatterplots of cellular uptake in individual patients receiving alpha1-oleate. Each dot represents the mean fluorescence intensity of six post-instillation samples per patient treated with alpha1-oleate. Comparison of alpha1-oleate uptake by the cell in urine before (pre = white) and after (post = black) alpha1-oleate inoculation on visits 1–6 (repeated-measures two-way ANOVA with Sidak's correction, $P = 0.0049$ for visit 1, 0.1913 for visit 2, 0.0067 for visit 3, 0.0025 for visit 4, 0.3807 for visit 5 and 0.0043 for visit 6, $n = 20$ per group and time point). Line represents the median and bars represent mean $\pm$ SEM.

Biopsies from healthy tissue areas distant from the tumor were collected for comparison.

Urine cytology. Urine cells were centrifuged onto L-lysine-coated microscope slides (Cytospin 3, Shandon) at 113 $\times$ g for 5 min, fixed and stored at room temperature until further analyses. Urinary cytology was evaluated using the Paris System for Reporting Urinary Cytology 2016^{45,46} and defined as: 1. No diagnosis/unsatisfactory. 2. Negative for high-grade urothelial carcinoma. 3. Atypical urothelial cells present. 4. Suspicious for high-grade urothelial carcinoma. 5. High-grade urothelial carcinoma. 6. Low-grade urothelial neoplasm. 7. Other positive for malignancies and miscellaneous lesions.

Alpha1-oleate uptake. Alpha1-oleate uptake by tumor cells was quantified by staining with specific antibodies. Cells on cytospin-slides were washed (Tris-buffer saline TBS, 10 min), permeabilized (0.25% TritonX-100 in TBS, 20 min, room temperature) and blocked (5% normal goat serum in TBS, 1 h, room temperature) before the addition of rabbit polyclonal anti-human alpha-lactalbumin antibodies (1:50 in 5% normal goat serum at 4 $^{\circ}$ C, overnight, Mybiosource, Cat# MBS175270). Slides were washed (TBS, 2 $\times$ 5 min) and stained with Alexa-568-labeled secondary antibody (1:200, 1 h, room temperature, ThermoFisher). The nucleus was counterstained using DRAQ5 (1:1000, 15 min) before a final wash (2 $\times$ 5 min in TBS). Slides were mounted (Fluoromount aqueous mounting media), before capturing images by laser scanning confocal microscopy (Carl Zeiss). Fluorescence intensity was quantified by ImageJ and net fluorescence calculated after subtraction of the secondary antibody background.

Apoptosis in tissue biopsies, TUNEL staining. DNA fragmentation was detected using the terminal deoxynucleotidyl transferase dUTP nick end-labeling (TUNEL) assay (Click-IT TUNEL Alexa Fluor 488 imaging assay kit, ThermoFisher). Tissue sections were de-paraffinized with xylene followed by serial dehydration with ethanol (100%, 95%, 75%, and 50%). Dehydrated sections were fixed (4% PFA, 15 min), permeabilized (DNase-free Protease K solution 20 μ g/mL, 15 min) and incubated with TUNEL reaction mixture containing TdT for 60 min at 37 $^{\circ}$ C. After the TUNEL reaction, sections were incubated with Click-IT reaction mixture (30 min, 37 $^{\circ}$ C). Sections were counterstained with DAPI (1 μ g/mL, 5 min), mounted in Fluoromount aqueous mounting media, and analyzed by fluorescence microscopy (Zeiss). Fluorescence intensities were quantified by ImageJ and net mean fluorescence intensity calculated after subtraction of background fluorescence.

RNA sequence analysis of tissue biopsies. RNA was extracted from tissues stabilized in RNAlater using the AllPrep DNA/RNA/miRNA Universal Kit. Disruption was in the TissueLysor system and CK28 Precellys tubes and by homogenization in the QIAshredder homogenizer. The quantity and quality of the RNA samples were evaluated using NanoDrop and Agilent 2100 Bioanalyzer. RNA samples were prepared by Illumina TruSeq Stranded mRNA Library Prep Kit (20020594), and libraries were multiplexed and sequenced using NextSeq 500/550 High Output Kits (v2.5 2 $\times$ 75 Cycles) with an average of 22 million reads per sample. Raw sequencing data were demultiplexed using bcl2fastq (version 2.18) and RSEM (1.3) was used for abundance estimation using the human genome release 37/Ensemble 75. Samples were thoroughly quality checked (QC) and visualized using dimensionality reduction (i.e. PCA), MA-plots as well as RNA-seq intrinsic biases (such as GC bias, transcriptome complexity, and alignment quality). Differential expression analysis was performed using R (version 3.4) and the packages limma and DESeq2. Fold changes were calculated by comparing tumors in the treated to the placebo group. Relative expression levels were analyzed and genes with an absolute fold change >2.0 and $P < 0.05$ were considered as differentially expressed. Heat-maps were constructed using the Gtools 2.1.1 software.

Differentially expressed genes were functionally characterized using the Ingenuity Pathway Analysis version 57662101 (IPA, Qiagen) software.

Statistical analysis. For efficacy, the sample size was based on analysis of change in tumor cells assessed from a previous study⁴⁶. A sample size of 20 patients per group was deemed suitable to achieve criterion for significance (alpha) 0.05 and power 90% using the paired samples 1-tailed t -test. The null hypothesis is H_0 : mean change in cell shedding = 0 and the alternative hypothesis is H_A : mean change in cell shedding > 0. The Gaussian distribution was determined by the D'Agostino and Pearson normality test. For data following a Gaussian distribution, student t -tests were used. Other data sets were analyzed by Mann-Whitney U -test. Correlations were determined by Spearman correlation. Kinetic data were analyzed using the repeated measures two-way ANOVA test. All statistical analysis was done by using Prism version 6.02 (GraphPad Software Inc.). P values < 0.05 were considered statistically significant. All images were created by the study team.

Molecular and cellular studies

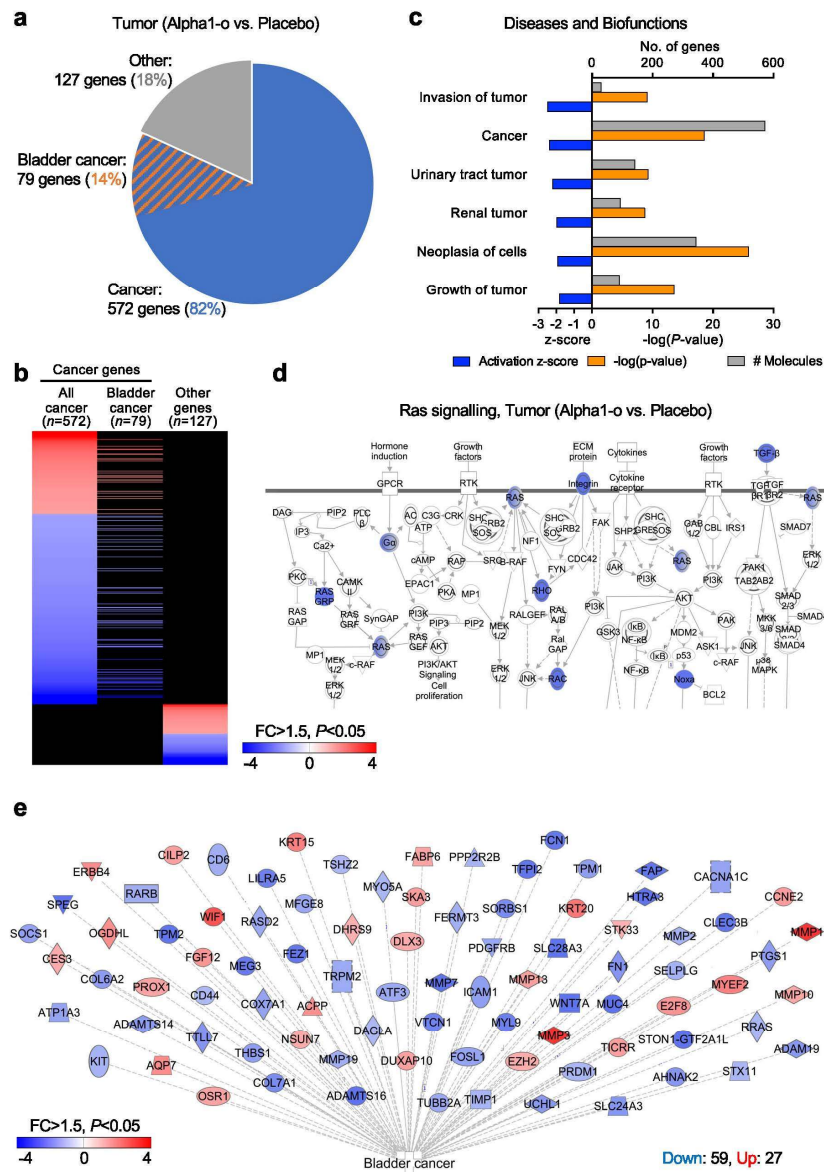
Chemicals and antibodies. Sodium oleate (Sigma-Aldrich, Cat# O7501), Alexa-Fluor568 protein labeling kit (Thermo Scientific, Cat# A10238), AlexaFluor488 protein labeling kit (Thermo Scientific, Cat# A10235), ATPite (Perkin Elmer, Cat# 6016947), Presto Blue Cell Viability Assay (Invitrogen, Cat# A13262), Anti-peptide antibodies (this study, produced by GeneCust), FluxOR Potassium ion channel assay (Invitrogen, Cat# P20015), Barium chloride BaCl₂ (Sigma-Aldrich, Cat# B0750), Amiloride (Sigma-Aldrich, Cat# A7410), Click-IT TUNEL Alexa Fluor 488 imaging assay kit (ThermoFisher Scientific Cat# C10245), DRAQ5 (Abcam, Cat# ab108410), Fluoromount (Sigma-Aldrich, Cat# P4680), DNA/RNA/miRNA Universal Kit (Qiagen, Cat# 80224).

Peptide synthesis and complex generation. Peptides for in vitro and in vivo experiments were synthesized using Fmoc solid-phase chemistry (Mimotopes, Melbourne, Australia). For biotinylated peptides, an aminohexanoic acid (Ahx) spacer was added to ensure adequate separation between the biotin and the peptide moieties. A five-fold stoichiometric concentration of sodium oleate in phosphate-buffered saline was prepared, followed by the addition of each respective peptide. The more hydrophobic peptides were initially dissolved in DMSO, then transferred to the oleate buffer. The sequences for the peptides are as follows:

Alpha1:	Ac-KQFTKAELSQLLKIDIDGYGGIALPELIATMPHTSGYDTQ-OH
Beta:	Ac-IVENNESTEYGLFQISNKLWAKSSQVPQSRNIADISADKFLD DD-OH
Sar1alpha:	Ac-MAGWDIFGWFRDVLASLGLWKNH-OH
Sar1beta:	Ac-DRLATLQPTWHPTEELAIIGNIKETFDLGGHI-OH

Cell lines and cell culture. Human lung carcinoma cells (A549, ATCC Cat# CCL-185, RRID:CVCL_0023), human kidney carcinoma cells (A498, ATCC Cat# HTB-44, RRID:CVCL_1056), and murine bladder carcinoma cells (MB49, RRID: CVCL_7076, provided by Sara Mangsbo, Uppsala University, Sweden) were cultured in RPMI-1640 with non-essential amino acids (1:100), 1 mM sodium pyruvate, 50 μ g/mL gentamicin, and 5–10% fetal calf serum (FCS) at 37 $^{\circ}$ C, 5% CO₂.

Cell death assays. To quantify effects on cell viability, A549, A498, or MB49 cells were seeded in 96-well plates (2 $\times$ 10⁴/well, Tecan Group Ltd.), cultured overnight at 37 $^{\circ}$ C, 5% CO₂ and incubated with peptide-oleate complexes in serum-free



RPMI-1640 at 37 °C. FCS was added after 1 h and cell viability was quantified after 3 h, by the Presto Blue fluorescence assay (Thermo Scientific) or ATP levels (luminescence-based ATPlite kit), using a microplate reader (Infinite F200, Tecan).

The colony assay was used to define long-term effects on cell viability. MB49 cells were seeded in a 24-well tissue culture plate (1000 cells/well in 1 mL in RPMI with FCS), allowed to adhere for 24 h, washed in PBS and treated with alpha1-oleate or sar1alpha-oleate (7, 21, or 35 μ M in 1 mL of RPMI without FCS

for 1 h). After the addition of FCS (5%) the cells were incubated at 37 °C in 5% CO₂ for 10 days or until visible colonies were present in the PBS control. Plates were fixed in cold methanol, stained with hematoxylin. Staining intensity was quantified using ImageJ.

TUNEL staining was performed using Click-IT TUNEL Alexa Fluor 488 Imaging Assay kit (Thermo Scientific). Briefly, the cells were fixed in 2% paraformaldehyde for 15 min followed by permeabilization with 0.25% TritonX-

Fig. 7 Reprogramming of gene expression. RNA sequencing was used to compare gene expression profiles in tumor tissue biopsies from the treatment or placebo groups. **a** Pie chart of genes regulated in response to treatment (cut-off FC > 1.5, $P < 0.05$ compared to the placebo group). In the treatment group, 82% of all regulated genes were cancer-related and 14% were bladder cancer-related. Gene categories were identified by biofunction analysis. **b** Heatmap of specific cancer- and bladder cancer-related genes regulated in tumor biopsies from the treatment group (red = upregulated, blue = downregulated, cut-off FC > 1.5, $P < 0.05$ compared to placebo group). About 60% of all regulated genes were inhibited in the treatment group. **c** Detailed analysis of data in (**a**, **b**). Top regulated, cancer-associated functions are shown. Inhibition is indicated by negative z -scores (blue) and significance by P values (orange). The expression of genes involved in tumor invasion, neoplasia, tumor growth, and urinary tract tumors was strongly inhibited. **d** Inhibition of Ras signaling in the treatment group compared to placebo. **e** Bladder cancer gene network regulated specifically in patients receiving alpha1-oleate treatment compared to placebo.

100 for 20 min. The sections were incubated for 60 min at 37 °C with the TUNEL reaction mixture, then counter-stained with DAPI (20 µg/mL) for 15 min. A positive control slide (cells treated with DNase I for 30 min at 37 °C) was included in each experiment. The TUNEL fluorescent intensity was quantified using Image J.

Membrane blebbing and ion fluxes. Membrane changes in lung carcinoma cells were visualized by transmission light microscopy imaging. Cells were seeded on a glass coverslip and allowed to partially adhere to the glass surface for 10 min at room temperature, prior to exposure to the alpha1-oleate or sar1alpha-oleate complexes. Changes in cell morphology were captured with LSM 510 META confocal microscope (Carl Zeiss) using $\times 40$ oil immersion objective. Naked peptides and oleate were used as negative controls (Supplementary Fig. 1e).

K^+ fluxes were quantified using the FluxOR potassium ion channel assay (Invitrogen). The influx of the indicator Ti^+ measures the opening of K^+ channels⁴⁵. Briefly, 20,000 adherent cells were incubated in loading buffer followed by incubation with assay buffer and stimulus buffer for 60 min. For inhibition, cells were pretreated with the K^+ channel inhibitor BaCl₂ (100 µM) followed by treatment with alpha1-oleate (35 µM). Fluorescence was measured at 535 nm after excitation at 485 nm using a fluorescence plate reader (TECAN infinite F200).

Structural analysis of the peptide-oleate complexes

CD spectroscopy. Far-ultraviolet (UV) CD spectra were collected on alpha1-, beta-, sar1alpha-, and sar1beta-peptides with and without oleate at 25 °C using a Jasco 815 CD Spectropolarimeter. The peptides were dissolved in 50 mM sodium phosphate buffer, pH 7.4, with 10% D₂O, at a final concentration of 0.2 mg/mL. Far-UV CD was performed from 185 to 260 nm for the samples without oleate and from 200 to 260 nm for the samples with oleate and the background was subtracted. The mean residue ellipticity (MRE), $[\theta]$, in deg cm²/dmol, was calculated as described previously³.

Biomolecular NMR spectroscopy. The alpha1- and sar1alpha-naked peptide samples were dissolved in 50 mM sodium phosphate buffer (pH 7.4, 90% H₂O, 10% D₂O), and the peptide-oleate complexes were reconstituted from a lyophilisate of phosphate-buffered saline. All experiments were carried out in the phase-sensitive mode⁴⁶. One-dimensional ¹H, two-dimensional NOESY (Nuclear Overhauser Effect Spectroscopy) and ¹H-¹³C HSQC (Heteronuclear Single Quantum Correlation) spectra were acquired on an Agilent Technologies 18.8T (800 MHz) DD2 Premium Compact spectrometer with a triple-resonance, 5 mm enhanced cold probe. The ¹H-¹³C HSQC spectra were collected at 20 °C with 16 scans, an initial delay of 3.0 s, a 90° pulse width of 7.5 and 9.8 µs, and an acquisition time of 0.4 s with broadband decoupling for alpha1 and sar1alpha peptide samples. For the alpha1-oleate and sar1alpha-oleate complexes, we used a 90° pulse width of 12.80 and 13.30 µs and an acquisition time of 0.4 s. All acquisition parameters were kept constant for all samples. Two-dimensional DPGSE-NOESY (Double Pulse Field Gradient Spin Echo-NOESY) pulse sequences were used to acquire data at 20 °C with 16 scans, with an optimized mixing time of 300 ms for the alpha1 and alpha1-oleate complexes and a delay period of 1.5 s. For the sar1alpha-oleate complex, water-gate NOESY was used with 12 scans, with a mixing time of 150 ms. A trace amount of TSP was added to serve as a chemical shift reference. Each 2D HSQC spectrum consisted of 4 K complex points in the acquisition dimension and 512 complex points in the indirect dimension. For the NOESY spectra, 4 K complex points were used in the acquisition dimension and 1 K complex points in the indirect dimension. The two-dimensional data were processed with Gaussian apodization in both dimensions. The stoichiometry of the peptide with the oleic acid was determined by comparing the peak areas (using the 1D ¹H spectra) or peak volumes (using the 2D ¹H-¹³C HSQC spectra) of well-resolved, isolated regions found in the spectra.

Diffusion-ordered spectroscopy (DOSY) measurements were performed at 293 K. Samples were prepared in 50 mM phosphate buffer at pH 7.4. The DgscstL_dpfgsc DOSY pulse program was used, which consists of gradient compensated stimulated echo with spin lock using the excitation sculpting solvent suppression method⁴⁷. A spectral window of 13,020 Hz was used, with an acquisition time of 2.46 s with a relaxation delay of 3 s. The FIDs were collected with 32,000 complex data points with 64 scans. Logarithmically the gradient pulse strength was increased from 3% to 86% of the maximum strength of 32,767 G/cm

in 60 steps. A diffusion time (Δ) of 100 ms and bipolar half-sine-shaped gradient pulses (δ) of 5 ms was applied. 1,4-Dioxane, which is known to behave independently of protein concentration and the folded state of the protein, was used as an internal chemical shift reference and hydrodynamic radius calibration reference (3.75 ppm; $R_H = 2.12$ Å)^{48,49}. DOSY processing was performed using a two-component fit with a discrete approach, which further processed using a non-uniform gradients approach. Three replicate acquisitions were given for each sample, and the resulting diffusion coefficient (D) values calculated. For alpha1 peptide the average D value was 2.162 and 14.10 m²/s for 1,4-dioxane. In the case of alpha1-oleate complex the average D value was 0.986 m²/s for complex and 13.61 m²/s for 1,4-dioxane. The calculated R_H are as follows: alpha1 peptide $R_H = 13.82 \pm 0.447$ Å, alpha1-oleate complex $R_H = 29.3 \pm 0.606$ Å, HSA $R_H = 40.9 \pm 1.44$ Å, oleate in aqueous solution $R_H = 104.3 \pm 7.22$ Å, oleate in methanol $R_H = 5.58 \pm 0.0649$ Å. Note that the D (diffusion coefficient) values for 1,4-dioxane are slightly variable dependent upon the co-solute (lower panel where D is between 14.4 and 12.6), which rightly reflects the different solution micro-environment conditions that both solutes are mutually experiencing for each sample.

For T_2 relaxation measurements, the standard CPMGT2 pulse sequence was used to run the experiments with 15 relaxations delays, which were chosen logarithmically for different maximum T_2 time intervals: 8 s (alpha1 peptide), 1.2 s (alpha1-oleate complex), 3.0 s (HSA), 7.0 s (oleate in aqueous solution), and 10 s (oleate in methanol), respectively. The data were acquired with 32,000 complex points with a baseline correction of 4. The T_2 analyses were performed on VNMRJ version 4.0 (Agilent Technologies) software by the exponential fitting of these values with their corresponding intensity. All other NMR parameters were kept constant for all samples throughout the experiments. The experiments were acquired at a sample temperature of 293 K. The data are presented in Supplementary Table 1.

Size-exclusion HPLC. Calibration standards and samples were injected onto a TSKgel Super SW3000 HPLC column (4.6 mm $\times$ 30 cm, Particle size 4 µm, pore size 25 nm, Tosoh Bioscience) eluted with 0.05 M sodium phosphate buffer pH 7.0 containing 0.1 M Na₂SO₄ at a flow rate of 0.25 mL/min and detection at 280 nm. The chromatography was performed on a Dionex Ultimate HPLC 3000 Standard System running Chromeleon 6 software (Dionex, Thermo Scientific). The standard calibration curve was generated with the proteins given including HSA, which has a hydrodynamic radius (R_H) of 40 Å (Supplementary Fig. 5). The R_H vs. elution volume linearity of the standard calibration curve is known to vanish after approximately 3.8 mL of elution volume⁵⁰. As a result, the R_H of oleate in methanol eluent (a solvent that ensures that the fatty acid is monomeric) will be less than what is estimated from the standard curve (12.3 Å). The retention times of small R_H -analyses are closely reproduced regardless of eluent, be it aqueous buffer or methanol.

Computational simulations: model building of peptide and peptide-OA complexes. The initial structure of the alpha1 peptide was obtained from the corresponding domain in the crystal structure of human alpha-lactalbumin (PDB ID: 1B9O). All cysteines were mutated to alanines, consistent with findings that a reduced human alpha-lactalbumin mutant in which all cysteines mutated to alanines could form a cytotoxic complex in the presence of the lipid cofactor⁹. The initial structure of the sar1alpha peptide was obtained from an I-TASSER-built homology model⁵¹. The sequence similarity of the Sar1alpha peptide with the sequence of the top-ranked threading template used by I-Tasser (PDB ID: 1R7G) was 0.35. However, this is eventually irrelevant as we used extensive H-REMD sampling to sample the conformations of the peptide, both in its oleate bound and apo forms. The alpha1 and the sar1alpha peptide were centered in a cubic box with box edges 1.2 nm from the peptide. For the oleate-containing systems of alpha1 and sar1alpha, 4 molecules of oleate are placed randomly in the box surrounding the alpha1 peptide to obtain a peptide-oleic acid ratio of 1:4. The Amber 99SB-ildn⁵² force field and the TIP3P⁵³ water model were used. For the coordinates, the starting structure was built using Discovery Studio 4.1 (Accelrys). Geometry optimization for the ligand was performed using Gaussian09⁵⁴ at the level of HF-6-31G*, and the partial charges were determined by the RESP⁵⁵ method implemented in the antechamber tool of AmberTools16 (AMBER 2016). Topologies for the oleate were built using the General Amber Force Field⁵⁶. The respective atom labels, corresponding atom type, and partial charges are shown in Supplementary Table 10. Additional parameters

for the molecule are derived from the General Amber Force Field. For reference, the structure of the oleate molecule and its atom labels is shown in Supplementary Fig. 11. All systems were neutralized and Na^+ and Cl^- ions were added to a concentration of 0.15 M. Energy minimization was performed using the steepest descent algorithm for 1000 steps to remove any initial bad contacts. Long-range electrostatics were treated with the particle mesh Ewald algorithm⁵⁷, with a real-space cutoff of 1.2 nm, and Van der Waals interactions were truncated at 1.2 nm. All systems with oleate-containing peptides or naked peptides were initially heated at 500 K for 40 ns to eliminate starting structure bias and provide a partially unfolded state for the peptides. Temperature coupling of the system was performed using a velocity rescaling thermostat⁵⁸.

Hamiltonian replica exchange molecular dynamics simulations. The Gromacs 5.1.2 molecular dynamics package⁵⁹ with the Plumed 2.3 plugin for Hamiltonian Replica Exchange Molecular Dynamics⁶⁰ was used to perform the simulations. All atoms of the alpha1- and sar1alpha peptide residues, along with the oleate residues of the two oleate-containing systems, were selected for Hamiltonian scaling. Twenty replicas were used for each system, and scaling factors were generated for an effective temperature range of 300–800 K. Temperatures for scaling were selected based on a geometric progression. The temperature factors were 300, 315.893, 332.629, 350.251, 368.807, 388.346, 408.919, 430.583, 453.395, 477.415, 502.707, 529.34, 557.384, 586.913, 618.006, 650.747, 685.223, 721.525, 759.75, and 800 K. Each replica was simulated for 400 ns, resulting in an effective simulation of 8 μs . Exchanges were attempted every 2 ps, and the result was an average acceptance probability of approximately 30%.

Simulation analysis. Analysis of simulation data was performed using the built-in Gromacs tools of the Gromacs package⁵⁹. The ensemble for each system with the canonical unscaled potential energy was used for the analysis, and data analysis was performed on the last 300 ns for each system. Dihedral principal component analysis⁶⁰ was performed using the gmx angles, gmx covar, and gmx anaig tools to prepare and diagonalize the covariance matrix and analyze eigenvectors and eigenvalues. The free-energy surface was constructed through projection onto the first and second principal components with the formula $F_i = -RT \ln(P_i/P_0)$, where R is the gas constant, T is temperature (300 K), P_i is the population in each bin and P_0 is the population of the most populated bin. The gmx cluster tool of Gromacs 5.1.2 was used to identify the representative structure of each minima for geometric clustering for the Gromacs algorithm. We used the define secondary structure of proteins (DSSP)⁶¹ algorithm to calculate secondary structure propensities. For our analysis, we classified the 3_{10} helix, α helix, and π helix structures as helices; the β -sheet and residue in isolated β -bridge structures as sheets and the remaining structures as others. The contact probability was calculated using the gmx mindist tool in the Gromacs package. The minimum distance between protons of side chains for each residue and oleic acid was calculated for each frame. To calculate the contact probability, a contact was defined if the measured distance was less than 0.55 nm. The contact probabilities between Aromatic ring protons and Olefinic protons of alpha1- and sar1alpha-oleate-containing systems were also calculated similarly. Proton distances were calculated to facilitate the comparison of simulation data to Nuclear Overhauser Spectroscopy data.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data supporting the structural and cellular findings of this study are available within the article and its supplementary information files. The structural data referenced during the study are available in a public repository from the Protein Data Bank website (www.rcsb.org, DOI:10.2210/pdb1B9O/pdb, DOI:10.2210/pdb1R7G/pdb). The RNA sequencing data generated in this study have been deposited in the Gene Expression Omnibus (GEO) database under accession number GSE172112. Source data are provided with this paper.

Received: 8 April 2020; Accepted: 15 April 2021;
Published online: 08 June 2021

References

- Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
- Scott, A. M., Wolchok, J. D. & Old, L. J. Antibody therapy of cancer. *Nat. Rev. Cancer* **12**, 278–287 (2012).
- Håkansson, A., Zhivotovskiy, B., Orrenius, S., Sabharwal, H. & Svanborg, C. Apoptosis induced by a human milk protein. *Proc. Natl Acad. Sci. USA* **92**, 8064–8068 (1995).
- Svensson, M., Håkansson, A., Mossberg, A.-K., Linse, S. & Svanborg, C. Conversion of α -lactalbumin to a protein inducing apoptosis. *Proc. Natl Acad. Sci. USA* **97**, 4221–4226 (2000).
- Gustafsson, L., Leijonhufvud, L., Aronsson, A., Mossberg, A. K. & Svanborg, C. Treatment of skin papillomas with topical α -lactalbumin-oleic acid. *N. Engl. J. Med.* **350**, 2663–2672 (2004).
- Chiti, F. & Dobson, C. M. Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. *Annu. Rev. Biochem.* **86**, 27–68 (2017).
- Knowles, T. P., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15**, 384–396 (2014).
- Ho, C. S. J., Rydstrom, A., Manimekalai, M. S. S., Svanborg, C. & Gruber, G. Low resolution solution structure of HAMLET and the importance of its α -domains in tumoricidal activity. *PLoS ONE* **7**, e53051 (2012).
- Petersson-Kastberg, J. et al. α -Lactalbumin, engineered to be nonnative and inactive, kills tumor cells when in complex with oleic acid: a new biological function resulting from partial unfolding. *J. Mol. Biol.* **394**, 994–1010 (2009).
- Fjell, C. D., Hiss, J. A., Hancock, R. E. & Schneider, G. Designing antimicrobial peptides: form follows function. *Nat. Rev. Drug Discov.* **11**, 37–51 (2011).
- Mok, K. H., Nagashima, T., Day, L. J., Hore, P. J. & Dobson, C. M. Multiple subsets of side-chain packing in partially folded states of α -lactalbumins. *Proc. Natl Acad. Sci. USA* **102**, 8899–8904 (2005).
- Fischer, W. et al. Human α -lactalbumin made lethal to tumor cells (HAMLET) kills human glioblastoma cells in brain xenografts by an apoptosis-like mechanism and prolongs survival. *Cancer Res.* **64**, 2105–2112 (2004).
- Hien, T. T. et al. Bladder cancer therapy without toxicity—a dose-escalation study of α 1-oleate. *Int. J. Can.* **147**, 2479–2492 (2020).
- Mossberg, A. K., Hou, Y., Svensson, M., Holmqvist, B. & Svanborg, C. HAMLET treatment delays bladder cancer development. *J. Urol.* **183**, 1590–1597 (2010).
- Puthia, M., Storm, P., Nadeem, A., Hsiung, S. & Svanborg, C. Prevention and treatment of colon cancer by peroral administration of HAMLET (human α -lactalbumin made lethal to tumour cells). *Gut* **63**, 131–142 (2014).
- Mossberg, A. K. et al. Bladder cancers respond to intravesical instillation of (HAMLET) human α -lactalbumin made lethal to tumor cells. *Int. J. Can.* **121**, 1352–1359 (2007).
- Barlowe, C. et al. COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* **77**, 895–907 (1994).
- Jensen, D. & Schekman, R. COPII-mediated vesicle formation at a glance. *J. Cell Sci.* **124**, 1–4 (2010).
- Sato, K. & Nakano, A. Dissection of COPII subunit-cargo assembly and disassembly kinetics during Sar1p-GTP hydrolysis. *Nat. Struct. Mol. Biol.* **12**, 167–174 (2005).
- Nadeem, A. et al. Protein receptor-independent plasma membrane remodeling by HAMLET: a tumoricidal protein–lipid complex. *Sci. Rep.* **5**, 16432 (2015).
- Kang, M., Jeong, C. W., Kwak, C., Kim, H. H. & Ku, J. H. Single, immediate postoperative instillation of chemotherapy in non-muscle invasive bladder cancer: a systematic review and network meta-analysis of randomized clinical trials using different drugs. *Oncotarget* **7**, 45479–45488 (2016).
- Sylvester, R. J. et al. Systematic review and individual patient data meta-analysis of randomized trials comparing a single immediate instillation of chemotherapy after transurethral resection with transurethral resection alone in patients with stage pTa–pT1 urothelial carcinoma of the bladder: which patients benefit from the instillation? *Eur. Urol.* **69**, 231–244 (2016).
- Khare, S. R. et al. Quality indicators in the management of bladder cancer: a modified Delphi study. *Urol. Oncol.* **35**, 328–334 (2017).
- Ho, J., Nadeem, A., Rydstrom, A., Puthia, M. & Svanborg, C. Targeting of nucleotide-binding proteins by HAMLET—a conserved tumor cell death mechanism. *Oncogene* **35**, 897–907 (2016).
- Lawler, P. R. & Lawler, J. Molecular basis for the regulation of angiogenesis by thrombospondin-1 and -2. *Cold Spring Harb. Perspect. Med.* **2**, a006627 (2012).
- Antoni, S. et al. Bladder cancer incidence and mortality: a global overview and recent trends. *Eur. Urol.* **71**, 96–108 (2017).
- Stevart, K. D. et al. Economic aspects of bladder cancer: what are the benefits and costs? *World J. Urol.* **27**, 295–300 (2009).
- van Rhijn, B. W. et al. Recurrence and progression of disease in non-muscle-invasive bladder cancer: from epidemiology to treatment strategy. *Eur. Urol.* **56**, 430–442 (2009).
- Hu, Q., Sun, W., Wang, C. & Gu, Z. Recent advances of cocktail chemotherapy by combination drug delivery systems. *Adv. Drug Deliv. Rev.* **98**, 19–34 (2016).
- Sharma, P. & Allison, J. P. Immune checkpoint targeting in cancer therapy: toward combination strategies with curative potential. *Cell* **161**, 205–214 (2015).

31. Ourfali, S. et al. Recurrence rate and cost consequence of the shortage of Bacillus Calmette-Guérin connaught strain for bladder cancer patients. *Eur. Urol. Focus* **7**, 111–116 (2021).
32. Dang, C. V. MYC on the path to cancer. *Cell* **149**, 22–35 (2012).
33. Chevallier, D., Carrette, D., Segretain, D., Gilleron, J. & Pointis, G. Connexin 43 a check-point component of cell proliferation implicated in a wide range of human testis diseases. *Cell. Mol. Life Sci.* **70**, 1207–1220 (2013).
34. Ghosh, S., Kumar, A. & Chandna, S. Connexin-43 downregulation in G2/M phase enriched tumour cells causes extensive low-dose hyper-radiosensitivity (HRS) associated with mitochondrial apoptotic events. *Cancer Lett.* **363**, 46–59 (2015).
35. Kamat, A. M. et al. KEYNOTE-676: Phase III study of BCG and pembrolizumab for persistent/recurrent high-risk NMIBC. *Future Oncol.* **16**, 507–516 (2020).
36. Lamm, D. L. Intravesical therapy for superficial bladder cancer: slow but steady progress. *J. Clin. Oncol.* **21**, 4259–4260 (2003).
37. Kamat, A. M. et al. Evidence-based assessment of current and emerging bladder-sparing therapies for non-muscle-invasive bladder cancer after Bacillus Calmette-Guérin therapy: a systematic review and meta-analysis. *Eur. Urol. Oncol.* **3**, 318–340 (2020).
38. de Jong, J. J., Hendricksen, K., Rosier, M., Mostafid, H. & Boormans, J. L. Hyperthermic intravesical chemotherapy for BCG unresponsive non-muscle invasive bladder cancer patients. *Bladder Cancer* **4**, 395–401 (2018).
39. Colombo, R. et al. Multicentric study comparing intravesical chemotherapy alone and with local microwave hyperthermia for prophylaxis of recurrence of superficial transitional cell carcinoma. *J. Clin. Oncol.* **21**, 4270–4276 (2003).
40. Arends, T. J. H. et al. Results of a randomised controlled trial comparing intravesical chemohyperthermia with mitomycin C versus Bacillus Calmette-Guérin for adjuvant treatment of patients with intermediate- and high-risk non-muscle-invasive bladder cancer. *Eur. Urol.* **69**, 1046–1052 (2016).
41. Boorjian, S. A. et al. Intravesical nadofaragene farnesin gene therapy for BCG-unresponsive non-muscle-invasive bladder cancer: a single-arm, open-label, repeat-dose clinical trial. *Lancet Oncol.* **22**, 107–117 (2021).
42. Babjuk, M. et al. EAU guidelines on non-muscle-invasive urothelial carcinoma of the bladder: update 2016. *Eur. Urol.* **71**, 447–461 (2017).
43. Rosenthal, D. L., Wojcik, E. M. & Kurtycz, D. F. In *The Paris System for Reporting Urinary Cytology* (eds Rosenthal, D. L., Wojcik, E. M. & Kurtycz, D. F.) (Springer, 2016).
44. VandenBussche, C. A review of the paris system for reporting urinary cytology. *Cytopathology* **27**, 153–156 (2016).
45. Storm, P. et al. A unifying mechanism for cancer cell death through ion channel activation by HAMLET. *PLoS ONE* **8**, e58578 (2013).
46. States, D. J., Haberkorn, R. A. & Ruben, D. J. A two-dimensional nuclear overhauser experiment with pure absorption phase in four quadrants. *J. Magn. Reson.* **48**, 286–292 (1982).
47. Pelta, M. D., Barjat, H., Morris, G. A., Davis, A. L. & Hammond, S. J. Pulse sequences for high-resolution diffusion-ordered spectroscopy (HR-DOSY). *Magn. Reson. Chem.* **36**, 706–714 (1998).
48. Jones, J. A., Wilkins, D. K., Smith, L. J. & Dobson, C. M. Characterisation of protein unfolding by NMR diffusion measurements. *J. Biomol. NMR* **10**, 199–203 (1997).
49. Shimizu, A., Ikeguchi, M. & Sugai, S. Appropriateness of DSS and TSP as internal references for (1)H NMR studies of molten globule proteins in aqueous media. *J. Biomol. NMR* **4**, 859–862 (1994).
50. SEC: Size Exclusion Chromatography (Tosoh Bioscience GmbH, 2017), <https://www.separations.eu.tosohbioscience.com/solutions/hplc-products/size-exclusion>.
51. Yang, J. et al. The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* **12**, 7–8 (2015).
52. Lindorff-Larsen, K. et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958 (2010).
53. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
54. Gaussian 09 (Gaussian, Inc., 2016).
55. Bayly, C. L., Cieplak, P., Cornell, W. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem. A* **97**, 10269–10280 (1993).
56. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
57. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: an $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
58. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
59. Wang, L., Friesner, R. A. & Berne, B. J. Replica exchange with solute scaling: a more efficient version of replica exchange with solute tempering (REST2). *J. Phys. Chem. B* **115**, 9431–9438 (2011).
60. Mu, Y., Nguyen, P. H. & Stock, G. Energy landscape of a small peptide revealed by dihedral angle principal component analysis. *Proteins* **58**, 45–52 (2005).
61. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983).

Acknowledgements

The authors thank Petr Bouška and NEOX for the Monitoring of the clinical study; Jeanette Valchik and the Center for Translational Genomics (CTG) at Lund University for the mRNA Library prep and Sequencing; Susanne Strömblad, Lina Gefors and the Lund University Biomimicry Centre (LBC) for providing experimental resources for tissue analysis; Arunima Chaudhuri for input and comments on the manuscript; Sara Mangsbo, Uppsala University, Sweden for the MB49 cells (RRID:CVCL_7076). We gratefully acknowledge the support of the Swedish Research Council, the Swedish Cancer Society (Cancerfonden) and HAMLET Pharma. Support to the Svanborg group was further provided from the European Union's Horizon 2020 research and innovation program under grant agreement No. 954360. The funding sources had no role in the design of this study or in its execution, analyses, interpretation of the data, or decision to submit the results for publication.

Author contributions

All authors met the ICMJE criteria for authorship. C.S., M.B., J. Ho, Y.G.M., K.H.M., conceived and designed the study. J. Ho, P.S.K., A.H., D.L.F., K.H.M. performed structural biology experiments and analyses. J.T.Y.-N. performed and analyzed the simulations under Y.G.M. supervision. J. Ho, A.N., T. Hiep T., P.E. performed cellular experiments and analyses. A.B., J. Håček, I.A., D.S.C.B., M.L.Y.W., T. Hiep T., H.N., J. Horňák, M.B. and C.S. performed the human trial and analyses. M.B., A.B., J. Håček, I.A., D.S.C.B., T. Hiep T., T. Hien T., M.L.Y.W., P.S., M.B., C.S. analyzed the data and J. Ho, I.A., D.S.C.B., T. Hiep T., K.H.M., M.L.Y.W., A.B., M.B., C.S. wrote the paper.

Funding

Open access funding provided by Lund University.

Competing interests

C.S. holds shares in HAMLET Pharma, as a representative of scientists in the HAMLET group. Patents protecting the use of the alpha1 peptide were filed previously (Biologically active complexes and therapeutic uses thereof, GB 201707715 priority date 14/05/2017, PCT/EP2018/062396 filing date 14/05/2018; inventors: C.S., A.N., J.Ho). No specific patents have been filed based on this study. Other authors declare no competing or conflicts of interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-23748-y>.

Correspondence and requests for materials should be addressed to C.S.

Peer review information *Nature Communications* thanks Keith Syson Chan and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

Arnas Petrauskas* & Daniel L. Fortunati* *et al.* **Structured and disordered regions of Ataxin-2 contribute differently to the specificity and efficiency of mRNP granule formation.** bioRxiv (2022).

Structured and disordered regions of Ataxin-2 contribute differently to the specificity and efficiency of mRNP granule formation.

Arnas Petrauskas^{1*}, Daniel L. Fortunati^{1*}, Amanjot Singh², Arvind Reddy Kandi⁵, Sai Shruti Pothapragada², Khushboo Agrawal^{3,4}, Joern Huelsmeier¹, Jens Hillebrand¹, Georgia Brown¹, Dhananjay Chaturvedi², Jongbo Lee⁶, Chunghun Lim⁶, Georg Auburger⁷, K. VijayRaghavan², Mani Ramaswami^{1,2†} and Baskar Bakthavachalu^{3,5†}.

¹Trinity College Institute of Neuroscience, School of Genetics and Microbiology, Smurfit Institute of Genetics and School of Natural Sciences, Trinity College Dublin, Dublin-2 Ireland. ²National Centre for Biological Sciences, TIFR, Bangalore 560065, India. ³Tata Institute for Genetics and Society Centre at inStem, Bellary Road, Bangalore 560065, India. ⁴School of Biotechnology, Amrita Vishwa Vidyapeetham University, Kollam-690525, Kerala, India. ⁵School of Basic Sciences, Indian Institute of Technology, Mandi 175005, India. ⁶Department of Biological Sciences, Ulsan National Institute of Science and Technology (UNIST), 50 UNIST-gil, Ulsan 44919, Republic of Korea. ⁷Experimental Neurology, Medical School, Goethe University, 60590, Frankfurt, Germany.

*These authors contributed equally to this work.

† Correspondence:

mani.ramaswami@tcd.ie; Phone +353 (1) 896 8400
baskar@iitmandi.ac.in; Phone +91-1905-267705

Running Title: Ataxin-2 PAM2:PABP interactions specify granule composition

Keywords: Ataxin-2, TRIBE, mRNA, PABP, Disordered Regions, Subcellular Organization, Stress Granule, Neurodegeneration, RNP Granule, *Drosophila*.

31 **ABSTRACT:**

32

33 Ataxin-2 (*ATXN2*) is a gene implicated in spinocerebellar ataxia type II
 34 (SCA2), amyotrophic lateral sclerosis (ALS) and Parkinsonism. The encoded protein is a
 35 therapeutic target for ALS and related conditions. ATXN2 (or Atx2 in insects) can function in
 36 translational activation, translational repression, mRNA stability and in the assembly of
 37 mRNP-granules, a process mediated by intrinsically disordered regions (IDRs). Previous work
 38 has shown that the LSm (Like-Sm) domain of Atx2, which can help stimulate mRNA
 39 translation, antagonizes mRNP-granule assembly. Here we advance these findings through a
 40 series of experiments on *Drosophila* and human Ataxin-2 proteins. Results of Targets of RNA-
 41 Binding Proteins Identified by Editing (TRIBE), co-localization and immunoprecipitation
 42 experiments indicate that a polyA-binding protein (PABP) interacting, PAM2 motif of Ataxin-
 43 2 may be a major determinant of the mRNA and protein content of Ataxin-2 mRNP
 44 granules. Transgenic experiments in *Drosophila* indicate that while the Atx2-LSm
 45 domain may protect against neurodegeneration, structured PAM2- and unstructured IDR-
 46 interactions both support Atx2-induced cytotoxicity. Taken together, the data lead to a proposal
 47 for how Ataxin-2 interactions are remodelled during translational control and how structured
 48 and non-structured interactions contribute differently to the specificity and efficiency of RNP
 49 granule condensation as well as to neurodegeneration.

50

51

52 INTRODUCTION:

53

54 mRNP granules are intriguing, dynamic membrane-less organelles containing translationally
 55 repressed mRNAs, RNA-binding proteins (RBPs), molecular chaperones and a variety of other
 56 cellular proteins (Buchan, 2014; Formicola *et al.*, 2019; Kiebler & Bassell, 2006; Knowles *et al.*, 1996; Martin & Ephrussi, 2009). The formation and composition of mRNP assemblies are
 57 determined by base-pairing interactions between mRNAs, protein-protein interactions and
 58 RBP-RNA interactions, whose respective contributions may vary across granule types and
 59 physiological states (Bevilacqua *et al.*, 2022; Matheny *et al.*, 2021; Van Treeck & Parker, 2018;
 60 Van Treeck *et al.*, 2018). Stress granules (SGs) are particularly well-studied granules that form
 61 when cellular stress signals mediated by eIF2 α kinase activation (Kedersha *et al.*, 1999) cause
 62 individual mRNPs to arrest in translation and condense into multi-mRNP assemblies (Ivanov
 63 *et al.*, 2019; Kedersha & Anderson, 2007; Youn *et al.*, 2019). Mutations in mRNP granule
 64 proteins, including TDP-43, FUS, Ataxin-2, hnRNPA1, hnRNPA2B1, EWSR1, have been
 65 associated with ALS and/or other forms of neurodegenerative disease (Cirulli *et al.*, 2015;
 66 Couthouis *et al.*, 2012; Elden *et al.*, 2010; Ginsberg *et al.*, 1998; Kim *et al.*, 2013; Liu *et al.*, 2017;
 67 Taylor *et al.*, 2016; Wolozin & Ivanov, 2019). For this reason, and because TDP-43 and other
 68 stress-granule protein aggregates are components of protein inclusions found in ALS and
 69 Frontotemporal dementia (FTD), the regulation and cellular functions of stress granules have
 70 been topics of considerable fundamental and clinical interest (Cao *et al.*, 2020; Li *et al.*, 2013;
 71 Mallucci *et al.*, 2020; Protter & Parker, 2016; Wang *et al.*, 2020; Wheeler *et al.*, 2016; Wolozin
 72 & Ivanov, 2019).

73

74
 75 The cast of intermolecular interactions required for mRNP-granule assembly and the precise
 76 sequence with which they occur are not yet elucidated (Khong & Parker, 2020; Van Treeck &
 77 Parker, 2018). However, many studies show that intrinsically disordered regions (IDRs) found
 78 on mRNP-granule proteins contribute substantially to RNP granule assembly (Andrusiak *et al.*,
 79 2019; Ash *et al.*, 2021; Calabretta & Richard, 2015; Decker *et al.*, 2007; Gilks *et al.*, 2004;
 80 Järvelin *et al.*, 2016; Kim *et al.*, 2021; Yang *et al.*, 2020). In biochemical experiments, such IDRs
 81 show the ability to phase separate into liquid-like assemblies (Babinchak & Surewicz, 2020;
 82 Han *et al.*, 2012; Hyman *et al.*, 2014; Kato *et al.*, 2012; Lin *et al.*, 2017; Murray *et al.*, 2017;
 83 Murthy *et al.*, 2019; Shin & Brangwynne, 2017; Strome & Wood, 1982; Toretzky & Wright,
 84 2014; Yang *et al.*, 2020) The accessibility or activities of IDRs can be tightly regulated by
 85 posttranslational modifications, allowing rapid physiological and spatial control over granule

86 assembly and disassembly (Ash *et al.*, 2021; Bah & Forman-Kay, 2016; Bah *et al.*, 2015;
87 Berlow *et al.*, 2015; Hofweber & Dormann, 2019; Kwon *et al.*, 2013; Rayman *et al.*, 2018; Saito
88 *et al.*, 2019; Yang *et al.*, 2020).

89

90 An important observation is that most IDRs also have the ability to transition from liquid-like
91 states into solid, beta-sheet rich, amyloid-fibrils *in vitro*, particularly at high concentrations
92 achieved in the liquid-phase (Alberti *et al.*, 2019; Li *et al.*, 2013; Murray *et al.*, 2017; Patel *et al.*
93 *et al.*, 2015; Ramaswami *et al.*, 2013). This, and studies showing that inhibitors of eIF2 α kinase or
94 downstream events including SG formation can be protective in animal models of
95 neurodegenerative disease (Chou *et al.*, 2017; Halliday *et al.*, 2017; Sidrauski *et al.*, 2015; Wong
96 *et al.*, 2018; Zyryanova *et al.*, 2021) have led to a conceptual framework in which: (a) mRNP
97 granules provide a microenvironment where pathogenic protein seeds can form and grow
98 (Bakthavachalu *et al.*, 2018; Mandrioli *et al.*, 2020; Patel *et al.*, 2015); (b) increased misfolded-
99 protein loads result in inclusion formation, chronic stress signalling and reduced protein
100 translation (Hetz *et al.*, 2020; Preissler & Ron, 2019); (c) increased demand on protein handling
101 systems results in multiple cellular defects, notably in the functions of membrane-less
102 organelles (Alberti *et al.*, 2017; Azkanaz *et al.*, 2019; Jiang *et al.*, 2020; Latonen, 2019; Schuller
103 *et al.*, 2021). In particular, aberrant SG formation also results in nuclear transport defects which
104 may contribute to cell death and toxicity (Hochberg-Laufer *et al.*, 2019; Zhang *et al.*, 2018).

105

106 Particularly strong support for the role of RNP granule formation in promoting disease comes
107 from studies of Ataxin-2. Loss of Ataxin-2 is cytoprotective in yeast TDP-43 and *Drosophila*
108 TDP-43 or C9ORF72 or Tau models of cytotoxicity (Bakthavachalu *et al.*, 2018; Becker *et al.*,
109 2017; Elden *et al.*, 2010; Huelsmeier *et al.*, 2021; Kim *et al.*, 2014; Lee *et al.*, 2016; Shulman &
110 Feany, 2003). In mouse models for SCA2 or ALS, either genetic loss of *ATXN2* or delivery of
111 antisense oligonucleotides (ASOs) targeting *ATXN2* in the central nervous system, reduced
112 aggregation of TDP-43, increased animal survival and improved motor function (Becker *et al.*,
113 2017; Scoles *et al.*, 2017). These observations have led to ASOs against human *ATXN2* being
114 developed and approved for clinical trials (Biogen, 2021).

115

116 Given Ataxin-2's therapeutic significance and multiple roles in biology, it is particularly
117 important to determine which molecular activities of the protein are relevant to disease and to
118 its various biological functions (Kim *et al.*, 2020). Across species, Ataxin-2 has three conserved
119 regions: a Like-Sm (LSm) domain, an LSm-associated domain (LSm-AD) and a PAM2 motif,

120 which is flanked by extended intrinsically disordered regions (IDRs) (Boeynaems *et al.*, 2021).
 121 Detailed work in *Drosophila* has shown that a c-terminal IDR of Atx2 is selectively required
 122 for mRNP assembly into granules (Bakthavachalu *et al.*, 2018). Parallel experiments showing
 123 that the IDR is also required for cytotoxicity in *Drosophila* FUS, C9ORF72 and Huntington's
 124 disease models suggest RNP-granule formation to be a significant mechanism by which Atx2
 125 promotes neurodegeneration (Bakthavachalu *et al.*, 2018; Huelsmeier *et al.*, 2021). A recent
 126 discovery that the Atx2-LSm domain antagonizes IDR-function has led to a model in which
 127 the Atx2 cIDR: (a) does not support mRNP assembly when Atx2 is associated with actively
 128 translating mRNAs through Atx2-LSm domain interactions; (b) becomes accessible and active
 129 in mediating mRNP assembly when LSm-domain interactions break and mRNA translation
 130 stalls (Boeynaems *et al.*, 2021; Singh *et al.*, 2021).

131
 132 We now present a series of experiments further elaborating mechanisms by Ataxin-2 functions
 133 in mRNA translation and mRNP assembly. These show that the PAM2 motif of Ataxin-2 and
 134 its interactions with PABP are not essential for granule assembly but are required to efficiently
 135 recruit Atx2-target mRNAs and specific protein components into Ataxin-2 granules. When
 136 taken together with other findings (Boeynaems *et al.*, 2021; Kim *et al.*, 2014; Satterfield &
 137 Pallanck, 2006; Singh *et al.*, 2021), our observations indicate that PAM2 binding to PABP on
 138 the polyA tail of mRNAs helps specify the composition of Ataxin-2 granules. We propose an
 139 early role for PAM2:PABP interactions working in coordination with the LSm domain to
 140 support mRNA translation and thereby oppose the mRNP formation (Boeynaems *et al.*, 2021);
 141 as well as a later role in escorting translationally-stalled PAM2:PABP linked mRNAs into
 142 mRNP granules. *In vivo* experiments analysing motor decline in transgenic *Drosophila* indicate
 143 that the PAM2:PABP interactions also support the progression of the neurodegenerative
 144 process. We provide new evidence for fresh insight into the enigmatic role of mRNP assembly
 145 in neurodegeneration.

146

147 RESULTS:

148

149 *The structured PAM2 domain of Atx2 is necessary for the correct mRNA and protein content*
 150 *of Atx2 granules.*

151

152 A recent eLife publication used Targets of RNA-Binding Proteins Identified by Editing
 153 (TRIBE) technology to identify mRNAs associated with Atx2 in the *Drosophila* adult brain

(Singh *et al.*, 2021). *In vivo*, the ability of an Atx2-fusion with ADARcd (the catalytic domain of an RNA-editing enzyme, ADAR), to edit a group of 256 target mRNAs was found to be dependent on the presence of the Atx2-cIDR, previously shown to be necessary for the formation of neuronal mRNP granules *in vivo*. In contrast, Atx2-ADARcd mutants lacking the LSm domain, both edited Atx2 TRIBE target RNAs and formed mRNP granules in cultured *Drosophila* S2 cells more efficiently than the wild-type. Thus, Atx2-ADARcd editing of target mRNAs occurs in and is reflective of mRNP granule assembly. While demonstrating a role for LSm-domain interactions in antagonizing cIDR mediated granule assembly, these observations did not address mechanisms by which Atx2 target mRNAs are selected, or whether and how Atx2 played any role in determining the composition of RNP granules. New experiments presented here address these outstanding questions.

Previous TRIBE analyses showed that LSm and LSm-AD regions have no major role in the recognition or selection of the Atx2-target mRNAs (Singh *et al.*, 2021). We therefore tested whether the third conserved region of Ataxin-2, a PAM2 motif known to associate with PABP (polyA binding protein), played any role in this process (Jiménez-López & Guzmán, 2014; Kaehler *et al.*, 2012).

We used Gal80^{ts}-controlled *elav-Gal4* to express Atx2ΔPAM2-ADARcd (deleted for the PAM2 motif) in brains of adult *Drosophila* for 5 days and used RNA-Seq to identify edited RNAs in polyA selected brain mRNA and compare it with Atx2-ADARcd using procedures described earlier (Figure 1A) (McMahon *et al.*, 2016). ADAR-edits, which converts Adenosine to Inosine on RNAs, are identified as A to G changes in TRIBE analyses. Each sample was sequenced to obtain 20 million reads (Supplementary table 1). The edits were only considered from the regions of the transcriptome that contained at least 20 reads. Genes with edits identified at a threshold above 15% in two biological replicates were considered as high-confidence true targets. We compared edit frequency and edited-gene identity in the brains of flies expressing Atx2ΔPAM2-ADARcd with those in brains expressing Atx2-ADARcd.

In contrast to Atx-2 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 1B, C and Supplementary table 2). More striking, 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 (Figure 1C). Of

the 108 genes edited by Atx2 Δ PAM2-ADARcd, 36 were also targets of Atx2-ADARcd, the remaining 72 were unique. (Figure 1C, D and Supplementary table 2). 50 edit sites were common between the Atx2 Δ PAM2 and Atx2WT targets. Those sites were edited with much lower efficiency in Atx2 Δ PAM2 as compared to Atx2WT (Figure 1E).

The location of edits made by Atx2 Δ PAM2-ADARcd also differed dramatically as to where they occurred relative to the coding sequences of the target mRNAs (Figure 1F). While edits made by wild-type and Δ LSm forms of Atx2-ADARcd were greatly enriched in the 3'UTR of the mRNAs, Atx2 Δ PAM2 targets were edited indiscriminately all along the mRNA length (Figure 1F).

Taken together, these data identify the PAM2 motif as necessary for Atx2 engagement with its correct mRNA targets. 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 for subsequent inclusion of these mRNAs in Atx2-containing granules.

If Atx2-ADARcd edits of target mRNAs occur predominantly in the mRNP granules (Singh *et al.*, 2021), then the ability of Atx2 Δ PAM2-ADARcd to edit some target mRNAs would suggest that the PAM2 motif is not essential for mRNP granule formation *per se*. To examine this, we expressed wild-type and Δ PAM2 mutant forms of GFP-tagged Atx2 under control of the native genomic promoter in *Drosophila* S2R+ cells. Atx2 overexpression in S2 cells induced the formation of mRNP granules closely related to SGs, containing endogenous Atx2 and various SG proteins as previously reported (Figure 2A) (Bakthavachalu *et al.*, 2018; Singh *et al.*, 2021). Similar expression of Atx2 Δ PAM2-GFP also induced granule formation. However, these granules were compositionally distinct from those induced by Atx2-GFP. While they clearly contained some SG markers present on Atx2-granules, e.g., Me31B and Rox8 (*Drosophila* homologs of DDX6 and TIA1), they did not contain others such as PABP, Caprin and dFMRP (Figure 2B).

RNP-granules induced by expression of wild-type, LSm and PAM2 deficient forms of Atx2-GFP required the presence of the c-terminal IDR (Figure 2C). Thus, while largely dispensable for efficient mRNP assembly, the PAM2 domain plays a significant role in determining both mRNA and protein components of mRNP granules. One possibility is that the PAM2 motif

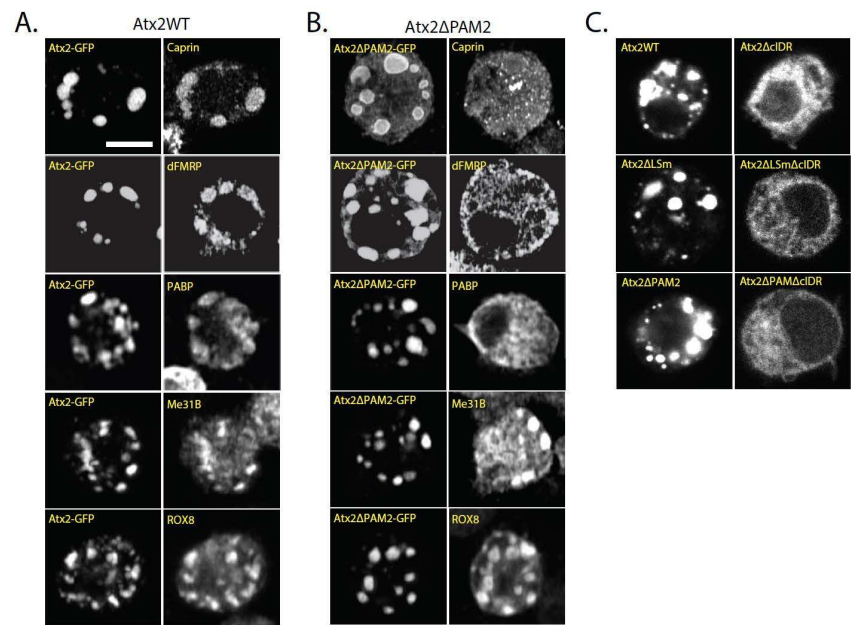


Figure 2: Presence of the PAM2 domain affects the protein composition of Atx2-GFP granules in S2 cells. (A) Over-expression of Atx2-GFP in unstressed *Drosophila* S2 cells induces the formation of Atx2-GFP granules to which various SG markers co-localize. (B) Deletion of the PAM2 affects the Atx2-GFP granule composition. Over-expression of Atx2ΔPAM2-GFP in S2 cells still induces the formation of granules, but some SG markers fail to co-localize in these, notably dFMRP, Caprin and PABP. (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. See Supplemental Figure 1, A-B, for quantification. The scale bar in (A) applies to (B) and (C). Scale bar = 5 μm,

PAM2:PABP interactions are sufficient for Atx2 to associate with stress granules

We wanted to directly confirm Ataxin-2 PAM2 motif interactions with PABP and analyse their relevance to RNP granule assembly. For this, we generated constructs encoding SNAP-epitope tagged variants of Atx2. These were radically truncated forms of *Drosophila* and human Ataxin-2 proteins containing only the LSm, LSm-AD and PAM2 elements and lacking all unstructured regions of the protein. The structured elements are connected via flexible linkers (Figure 3A). These “Mini-Ataxin-2” constructs and their domain-deleted forms allowed us to separate functions of the structured regions of Ataxin-2 from those of the remaining extended

disordered regions. A similar approach has been previously shown for MeCP2 (Tillotson *et al*, 2017). We identified key residues involved in *Drosophila* Atx2-PAM2:PABP interactions based on a previously solved crystal structure of a strongly conserved mammalian PAM2:PABPC1-MLLE domain complex (Kozlov *et al*, 2010; Xie *et al*, 2014) (Figure 3B). Residues leucine 914 and phenylalanine 921 (L914 and F921) in the human ATXN2-PAM2 motif are predicted to contact the PABPC-MLLE domain and of these, F921 has been shown to be required for the PABPC-ATXN2 interaction (Inagaki *et al*, 2020). These residues (L859 and F866, respectively) are perfectly conserved in fly Atx2 (Supplementary Figure 3). In order to allow more precise disruption of PAM2:PABP interactions and avoid potential unknown secondary effects of larger PAM2 motif deletions, we additionally generated mini Ataxin-2 constructs where these PABP-contacting residues were singly or doubly altered to alanine. We used these constructs for co-immunoprecipitation (Figure 3) and co-localization (Figure 4) analyses to examine the contribution of PAM2:PABP/PABPC1 interactions in RNP-granule formation.

We expressed SNAP-tagged wild-type and mutant forms of mammalian and *Drosophila* mini-Ataxin-2 in HEK293T and S2 cells, respectively. We tested which Ataxin-2 domains were required for SNAP substrate beads to successfully immunoprecipitate Ataxin-2 complexes containing PABPC1/PABP from cell lysates (Figure 3C-E). Both LSM12 and PABPC1 proteins were co-immunoprecipitated with mammalian mini-ATXN2. However, PABPC1 co-immunoprecipitation was selectively lost when the PAM2 domain was deleted or if predicted PABP-contact residues in the PAM2 domain were mutated (Figure 3C). Similar to the human homolog, fly mini-Atx2-SNAP also required the presence of its PAM2 motif with both predicted contact residues intact for immunoprecipitation of PABP from *Drosophila* S2 cell lysates (Figure 3D). Taken together with previous observations, these data support a potential sequence of molecular events. In unstressed cells, PAM2 domain interaction with PABP help position Ataxin-2 at the 3'-end of mRNAs while LSM-domain association with LSM12 stimulate translation of these mRNAs; under stress conditions (or Ataxin-2 overexpression), translation is arrested and the cIDR domain freed to mediate interactions that facilitate the formation of condensed RNP granules (see Discussion)

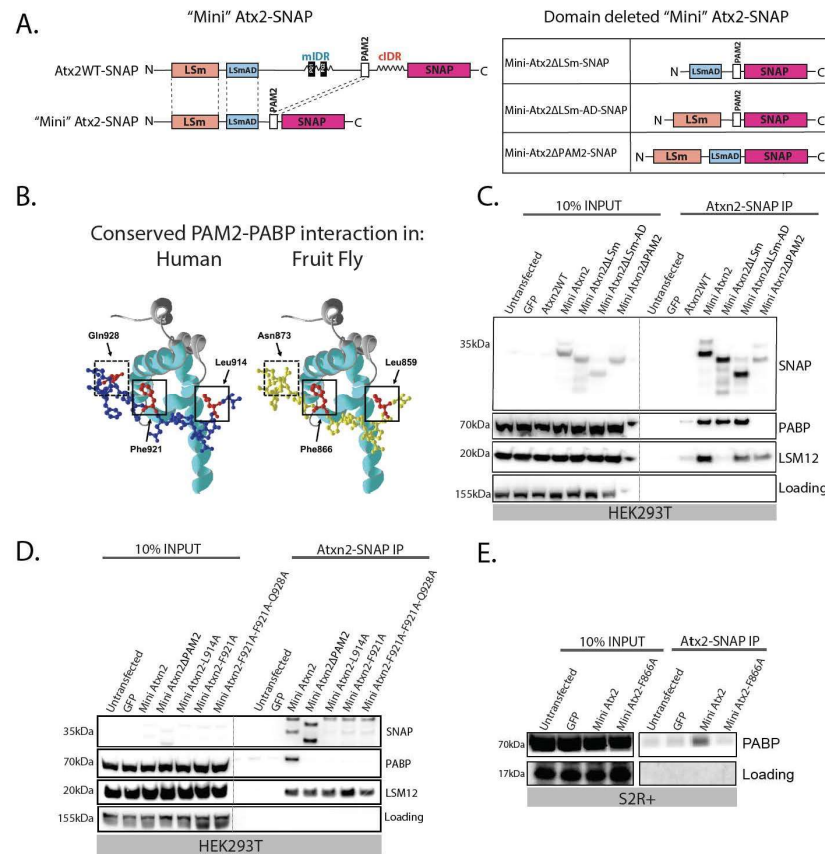


Figure 3: A minimized Ataxin-2 construct containing only the known structured domains maintains the ability to interact with PABP and LSM12. (A) Schematic of SNAP-tagged full length, minimal, and minimal domain-deleted constructs of fly Atx2 used to isolate the function of the structured domains of the protein without interference from IDR-mediated interactions. Human and *Drosophila* Ataxin-2 LSm, LSm-AD and PAM2 domains show high amino acid sequence similarity (Clustal Ω) of 70%, 82% and 87% respectively. This suggests conserved and specific function of these structured domains. (B) Structural model of the PABP MLE 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. (C) 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. (D) 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. (E) *Drosophila* minimized Atx2-SNAP IP-WB from S2 cells. An analogous PAM2 domain point mutation on F866 blocks the Atx2-PABP interaction.

Single mRNAs usually associate with multiple PABP molecules because their polyA tails are considerably longer than the ~24 bases required for PABP binding (Mangus *et al.*, 2003). Thus, in cells expressing endogenous and mini Ataxin-2, mRNAs could have both forms associated with their polyA tails. In response to stress, mini Ataxin-2 would be expected to move into SGs whose formation is facilitated by IDRs on endogenous Ataxin-2 associated with the common target RNAs. We examined this possibility in cells before and after oxidative stress.

Drosophila and human mini-Ataxin-2-SNAP, expressed in fly S2 or human U2OS cells respectively, were diffusely localized in the cytoplasm and neither induced formation of Ataxin-2 foci. However, when cells were exposed to sodium arsenite to induce SG formation, SNAP-tagged mini-Atx2 (Figure 4A) and mini- ATXN2 (Figure 4H) were robustly recruited to stress granules. Thus, association of Ataxin-2 with mRNP-granule components may be achieved by structured domain interactions alone, independently of IDRs required for mRNP assembly into granules.

Further experiments examined which of the LSm, LSm-AD and/or PAM2 domains were necessary for mini-Ataxin-2 to associate with stress granules. Mammalian and *Drosophila* mini-Ataxin-2 forms missing the LSm or LSm-AD domains could still be found in stress granules (Figure 4B-C and I-J). In contrast, mutants lacking the PAM2 domain remained cytoplasmic after stress in both S2 (Figure 4D) and U2OS cells (Figure 4K). Notably, point mutations in the Ataxin-2-PAM2 domain that specifically disrupt PAM2:PABP interaction similarly prevent localization to stress granules (Figure 4E-G and L-N). Thus, interactions between Ataxin-2's PAM2 domain and PABP appear important for the presence of Ataxin-2 in native mRNP granules, whose assembly is driven by the distinct (IDR) region of the protein (Figure 2C). The ability of otherwise full-length Ataxin-2 lacking PAM2 to form compositionally distinct mRNP assemblies (Figure 2B) suggests that PAM2:PABP binding also serves to limit non-physiological interactions by Ataxin-2 (See Discussion).

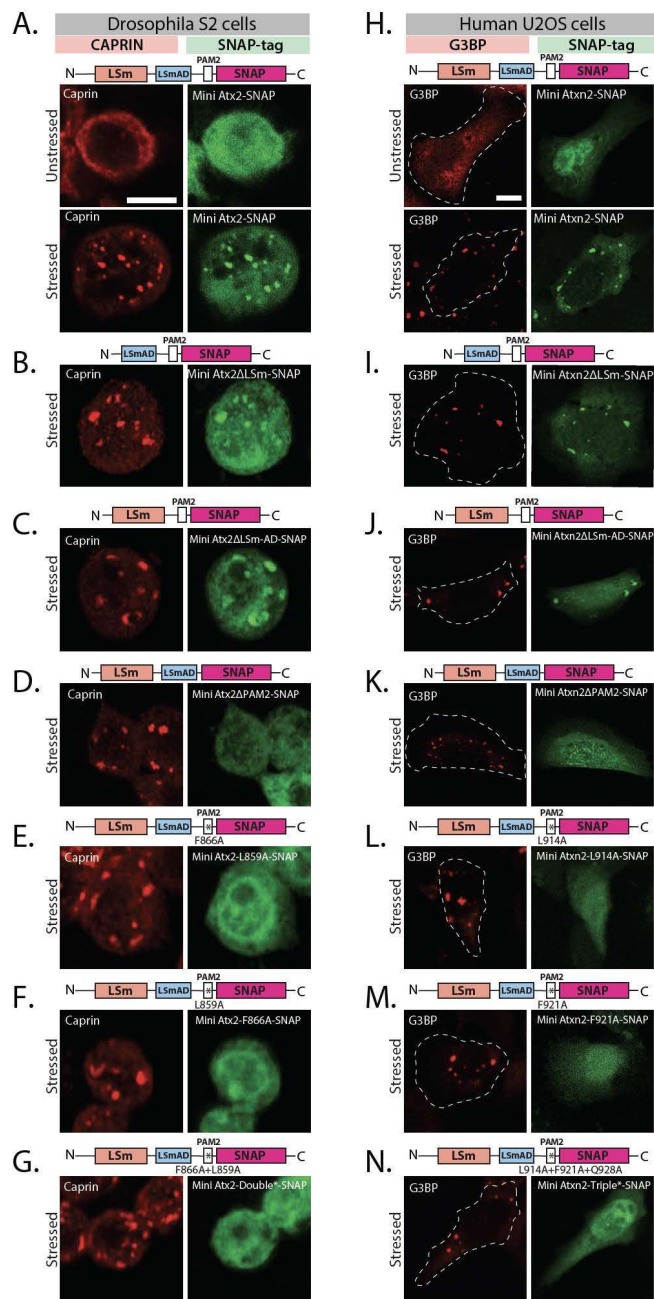


Figure 4: The structured PAM2 domain is necessary and sufficient for Ataxin-2 recruitment to Stress Granules in both *Drosophila* and human cells. (A) In *Drosophila* S2 cells mini-Atx2-SNAP (green) is recruited to SGs induced by arsenite. (B) Deletion of LSm or (C) LSm-AD domains has no significant effect in the arsenite-induced SG recruitment. (D) The presence of PAM2 domain, and specifically its key PABP-interacting amino acids (E-G), is necessary for Atx2 recruitment to SGs. Caprin (red) was used as SG granule marker, scale bar = 5 μ m, (H) Human mini-ATXN2-SNAP (green) is recruited to arsenite-induced SGs in human U2OS cells. (I) Deletion of LSm or (J) LSm-AD domains has no effect on the arsenite-induced SG recruitment of ATXN2. (K) Deletion of the PAM2 domain, and specifically its PABP-interacting amino acids (L-N), are necessary for ATXN2 recruitment to SGs. G3BP1 (red) was used as SG marker; scale bar = 10 μ m. Schematics above images indicate the domain deletions or amino acid mutations that were made in the different Ataxin-2 constructs. See Supplemental Figure 1C for quantification.

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

Three different Ataxin-2 domain deletions tested showed three distinct effects on mRNP granule assembly in S2 cells. IDR domain deletions prevent Ataxin-2 granule formation. LSm-domain deletion enhances the formation of Ataxin-2 granules. PAM2 domain deletions result in the formation of unusual mRNP assemblies (Bakthavachalu *et al.*, 2018; Singh *et al.*, 2021) (Figure 2B/C). Prior observations showing that Atx2 IDR deletions suppress cytotoxicity in *Drosophila* models for neurodegeneration indicate that mRNP granules support events that lead to degenerative disease (Bakthavachalu *et al.*, 2018; Becker *et al.*, 2017; Huelsmeier *et al.*, 2021; Scoles *et al.*, 2017). If true, the expression of Atx2 Δ LSm, which enhances granule assembly, would promote or potentially accelerate the degeneration, while the expression of Atx2 Δ cIDR would not. The expression of Atx2 Δ PAM2 would be expected to support mRNP assemblies of different compositions from the ones containing wild-type or Δ LSm forms of Atx2. The effects on degeneration for this condition would be hard to predict.

To examine how the different Atx-2 domain deletions affect nervous system integrity and function over time, we combined a Gal4-responsive *UAS-Atx2* transgene with *elav-Gal4* and *TubGal80^{ts}*. This allows us to use a temperature shift from 18°C to 30°C to induce *UAS-Atx2* transgene expression, specifically in the brains of adult flies (Figure 5A). We then analysed the rate at which flies climbed the walls of a glass cylinder, a surrogate measure of motor ability, one day and 15 days after transgene expression. All genotypes tested showed robust and

comparable levels of climbing ability on day 1. Interesting variations were identified on day 15. The 15-day old flies expressing Atx2WT or Atx2 Δ LSm showed a strong decline in climbing ability. In contrast, Atx2 Δ cIDR flies showed a minimal decline (Figure 5B). These observations were in line with the effects of these Atx2 types on granule formation. Strikingly, flies expressing the Atx2 Δ PAM variant, which formed compositionally distinct granules in S2 cells, showed no significant decline in climbing ability, suggesting that Atx2's ability to promote progressive decline of neural function depends less on Atx2 granule formation and aggregation, but a bit more on its sequestration of critical translation factors such as PABP (and associated RNAs) (Figure 5B). These observations support and extend prior work showing that heterologous overexpression of full-length, but not PAM2-domain deleted forms of mammalian ATXN2 enhances mammalian TDP-43-induced degeneration of the *Drosophila* compound eye (Kim *et al.*, 2014). They are also consistent with work in mice showing that PABPC1 sequestration in inclusions correlates strongly with the progression of neurodegeneration (Damrath *et al.*, 2012).

The conclusion, that PAM2 mediated interactions were required for progressive cytotoxicity, is further supported by a parallel series of experiments in which we used *mef2-Gal4* in place of *elav-Gal4*, to target *UAS-Atx2* transgene expression to *Drosophila* adult muscles (Figure 5C). Micro-computed tomography (micro-CT) scanning to visualize the integrity of flight muscle fibers in whole-mount preparations (see Methods) revealed degeneration of muscles expressing wild-type Atx2 in 20-day old flies. While there was more severe degeneration in Atx2 Δ LSm expressing muscle, muscles similarly expressing Atx2 Δ cIDR or Atx2 Δ PAM forms showed no morphological defects (Figure 5C and Supplementary Figure 4).

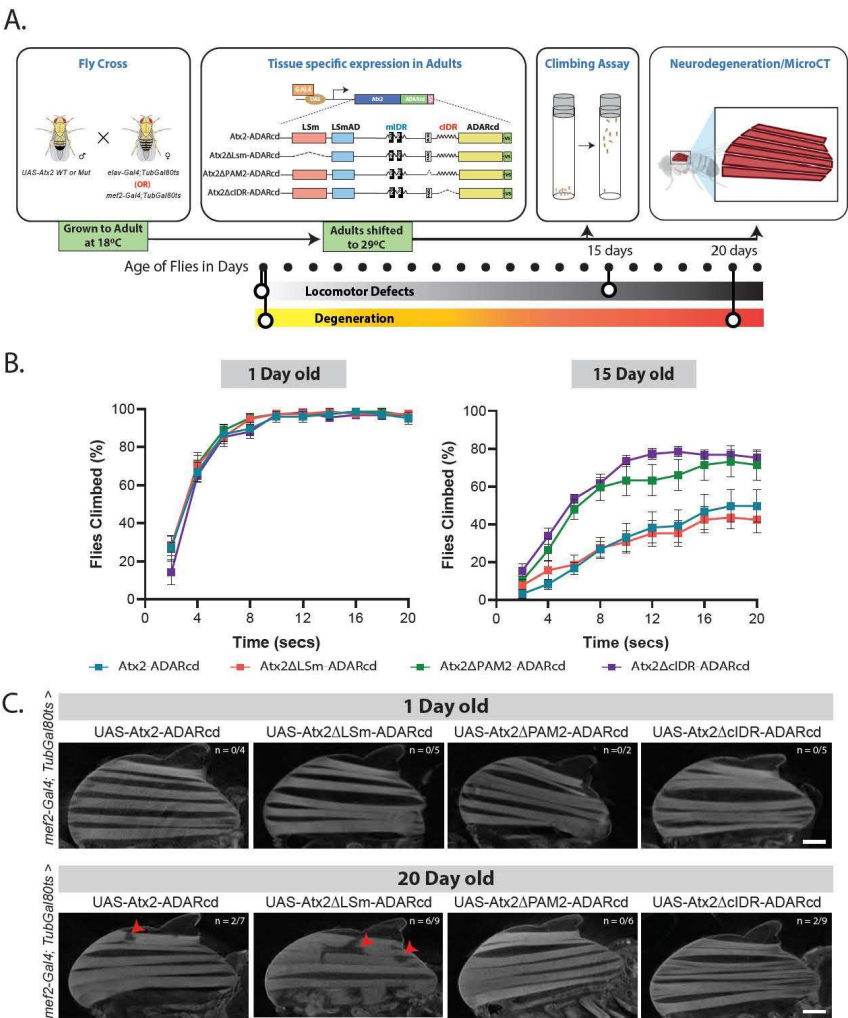


Figure 5: The IDR and PAM2 domains promote and the LSm domain inhibits neurodegeneration in *Drosophila*. (A) A schematic of the experimental design is shown. UAS-transgenes were crossed with *elav-Gal4; tub-Gal80^{ts}* or *mef2-Gal4, tub-Gal80^{ts}* and kept at 18°C till the adult flies emerged. The 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 (Atx2WT, 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 (Atx2WT, Atx2ΔcIDR, Atx2ΔPAM or Atx2ΔLsm) with *mef2-*

402 *Gal4*. Fly indirect flight muscles were imaged using micro-CT and the loss of muscle fibers is shown with
403 solid red arrowheads.

404

405 **DISCUSSION**

406 The results described above provide three significant lines of insight. First, they support a
407 detailed model for sequential protein-protein interactions through which Ataxin-2 can
408 modulate different translational states of a single mRNA. Second, they show that the Ataxin-2
409 polypeptide contains distinct activities that promote or protect against neurodegeneration,
410 pointing to the value of developing therapeutics that target specific Ataxin-2 interactions,
411 beyond those that reduce overall levels of the protein. Third, the work identifies a novel
412 molecular mechanism involving the PAM2 domain and PABP that contributes to the assembly
413 of mRNP granules.

414

415 **Molecular mechanisms of Ataxin-2 function**

416 Some RNA-binding proteins can remain associated with mRNAs across multiple stages: RNA
417 processing, transport, translation, or translational control (Formicola *et al.*, 2019; Gomes &
418 Shorter, 2019; Hachet & Ephrussi, 2004; Harlen & Churchman, 2017; Lin *et al.*, 2015; Maniatis
419 & Reed, 2002). Ataxin-2 may be one such protein. It is a translational activator of the
420 *Drosophila period* mRNA, a repressor of several miRNA reporters, a facilitator of neuronal
421 mRNP-granule and stress-granule formation as well as a broad stabilizer of Ataxin-2 associated
422 mRNAs (Bakthavachalu *et al.*, 2018; Inagaki *et al.*, 2020; Lim & Allada, 2013; McCann *et al.*,
423 2011; Nonhoff *et al.*, 2007; Sudhakaran *et al.*, 2014; Yokoshi *et al.*, 2014; Zhang *et al.*, 2013).
424 While these different functions could represent different modes of engagement with distinct
425 sets of mRNAs, the data are also consistent with another model. Sequential interactions
426 mediated by different protein regions during mRNP modelling allow Ataxin-2 to contribute in
427 multiple ways to translational control to a single mRNA.

428

429 Previous work has shown that Atx-2 enhances *period* mRNA translation through a mechanism
430 requiring LSm-domain interactions with a complex of LSM12 and TYF (Twenty Four) proteins
431 associated with the 5' cap of the translating mRNA (Lee *et al.*, 2017; Lim & Allada, 2013; Zhang
432 *et al.*, 2013). Given considerable supportive evidence for direct binding between the LSm-
433 domain and LSM12, we postulate that LSm-domain-LSM12 interactions occur in translating
434 polysomes (Satterfield & Pallanck, 2006) and contribute to increased efficiency of translation.

435 This proposal is consistent with the observation that the LSm domain opposes the formation of
 436 mRNP granules, which usually contain translationally repressed mRNAs (Singh *et al.*, 2021).
 437
 438 However, the LSm domain must also contribute to LSM12-independent functions, because
 439 while LSm-domain deletions from *Drosophila* Atx2 cause lethality and LSM12 null mutants,
 440 while arrhythmic, are viable and fertile (Lee *et al.*, 2017). One possibility is that LSm domains
 441 additionally contribute, perhaps indirectly, to interactions with the DEAD-box helicase
 442 Me31B/DDX6 in a translational repressor complex (Brandmann *et al.*, 2018; Lee *et al.*, 2017).
 443 Thus, we suggest that in the case of actively translating mRNAs, the Atx2 function is driven
 444 by LSm-domain association with LSM12 and translational initiators, and that LSM12
 445 disengages from a translational initiation complex as the mRNA transitions into a repressed
 446 state driven by Me31B.
 447
 448 While polyA tails and PABP are known to support translation and the Ataxin-2 PAM2 domain
 449 is involved in targeting the protein to polysomes (Satterfield and Pallanck, 2006), existing data
 450 do not directly address how Ataxin-2 PAM2 motif interactions contribute to translational
 451 activation. One possibility, supported by observations on the *period* mRNA is that the PAM2-
 452 domain guides Ataxin-2 to the 3'UTR of its target mRNAs (Lim & Allada, 2013). Our
 453 observation that PABP co-immunoprecipitates with mini-Ataxin2, show that Atx2-
 454 PAM2:PABP interactions occur independently of and prior to mRNP granule formation.
 455 Recent findings that this association antagonizes the Ataxin-2 condensation (Boeynaems *et al.*,
 456 2021) are consistent with a model in which the Atx2-PAM2 motif interacts with PABP in
 457 translating mRNAs to support efficient translation driven by the LSm-LSM12 complex.
 458 However, in addition to supporting translation, PABP is also known to associate with
 459 translational repressors that could drive either mRNA deadenylation and/or storage (Machida
 460 *et al.*, 2018; Yoshida *et al.*, 2006). Our data support such a dual role for Ataxin-2 associated
 461 with PABP in translational repression. First, when Ataxin-2 target mRNAs are not actively
 462 translated, then the mRNP through Me31B/DDX6 and PABP may recruit deadenylases to
 463 transition into either a translationally dormant or degradative state (Lee *et al.*, 2017; Machida
 464 *et al.*, 2018; Yi *et al.*, 2018; Yoshida *et al.*, 2006). Second, Atx2 associated mRNA may move
 465 into mRNP granules whose formation is facilitated by Atx2 IDR-mediated condensation. We
 466 postulate that mRNAs in such assemblies are stored in a form that is protected from
 467 degradation. While the above model, shown in Figure 6, 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.

Implications for Ataxin-2 as a therapeutic target

Antisense Oligonucleotide (ASO) based therapeutic strategies that lower levels of Atxn-2 are being developed for the treatment of ALS and spinocerebellar ataxia type 2 (SCA2) (Becker *et al.*, 2017; Scoles *et al.*, 2017). Our experiments provide a much finer grained analysis of activities of Ataxin-2, suggesting that the function of the LSm domain should be spared, and that IDR mediated assembly mechanisms and perhaps PAM2:PABP interactions should be most usefully targeted by therapeutics.

Our previous work showed that *Atx2* mutants lacking the cIDR required for Ataxin-2 granule formation in *Drosophila* neurons and S2 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 5C). 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, they 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 (De Graeve & Besse, 2018; Vogler *et al.*, 2018).

Structured interactions may determine mRNP granule composition.

502 Many lines of evidence argue that specific molecular interactions, e.g. mediated by structured
503 domains of the P-body component Edc3 or the stress-granule components G3BP and Caprin,
504 contribute to the mRNP granule formation (Decker *et al.*, 2007; Kedersha *et al.*, 2016). In
505 engineered systems, the condensation of RNA-binding proteins and mRNAs into granules has
506 been clearly shown to depend on both traditional protein-protein interactions and on more
507 promiscuous interactions between intrinsically disordered regions (Protter *et al.*, 2018). Our
508 work now identifies the interactions between Ataxin-2's PAM2 motif and PABP as a critical
509 contributor to the assembly of Ataxin-2 containing mRNP granules. This suggests a mechanism
510 by which the interaction helps select mRNA and protein components of mRNP granules.

511

512 We suggest that Ataxin-2, guided by PAM2:PABP interactions and LSm domain interactions,
513 recruits target mRNAs and associated proteins into translating mRNPs (Satterfield & Pallanck,
514 2006). Under conditions where the translation is arrested, LSm-domain interactions are altered
515 (Lee *et al.*, 2017), and transcripts are released from stalled ribosomes. Base-pairing interactions
516 between exposed mRNA side chains, as well as interactions between Ataxin-2's now
517 accessible intrinsically disordered regions, contribute to the assembly of these mRNPs into
518 granules. This logical sequence of events is consistent with: (a) TRIBE data showing a reduced
519 number of edits of native Ataxin-2 target mRNAs by Atx2ΔPAM2-ADARcd; (b) the inability
520 of ΔPAM2-miniAtx2 constructs to associate with stress granules; and (c) the aberrant protein
521 composition of granules induced by Atx2ΔPAM2 in S2 cells. The additional observation that
522 Atx2ΔPAM2-ADARcd expression results in a large number of non-native mRNA edits,
523 indicates that the PAM2:PABP interaction not only selects correct target mRNAs but also
524 prevents Ataxin-2 engagement with incorrect mRNA target regions.

525

526 Our conclusion that Ataxin2-PAM2:PABP interactions are involved in the selection of mRNA
527 components of RNP granules is superficially inconsistent with the observation that RNA
528 components of native stress granules can be predicted with remarkable accuracy on the basis
529 of mRNA size. This argues for a primary role for RNA-RNA interactions in the stress granule
530 assembly (Jain & Vale, 2017; Matheny *et al.*, 2021; Van Treeck & Parker, 2018). However,
531 we note that experiments presented here do not address mechanisms by which mRNAs are
532 selected into stress granules. Instead, the TRIBE data address how Atx2-target mRNAs are
533 selected into neuronal mRNP granules that exist in non-stressed cells *in vivo*, and microscopic
534 studies analyse protein components of mRNP granules formed following Atx2 expression in
535 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*.

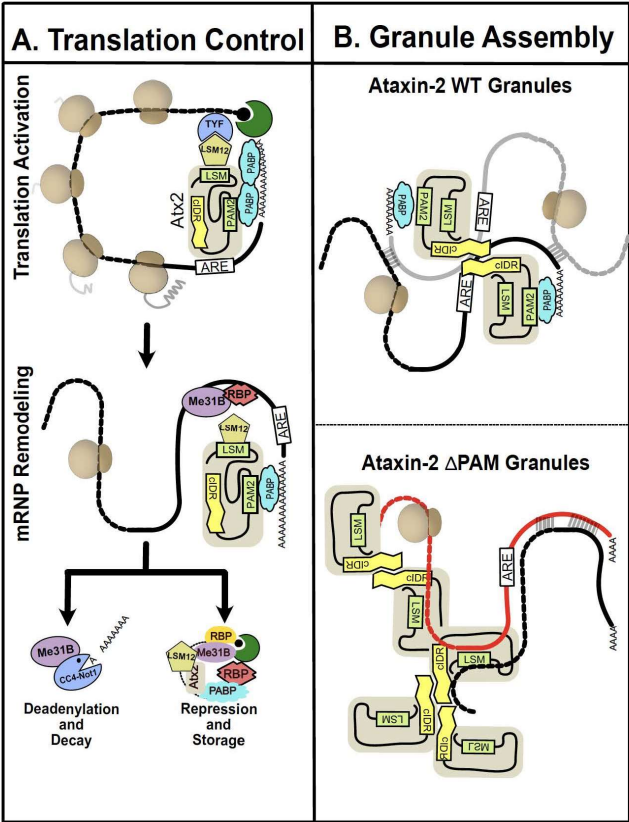


Figure 6: A model for Ataxin-2 RNP dynamics and the role of PAM2 domain in determining its RNP composition and mRNA selection. (A) Ataxin-2 is recruited to mRNAs by RBPs during different stages of the mRNA life cycle. Ataxin-2 activates translation of subsets of mRNA by recruiting LSM12, TYF and other 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

551 deadenylation can lead to degradation or translation repression and storage in RNP granules. (B) Ataxin-
552 2-PAM2 domain determines protein and RNA partners of the RNP granules. PAM2 domain is essential for
553 recruitment of Ataxin-2 to stress granules that also contains other RBPs (eg. Me31B, FMRP, Rox8, Rin
554 and Caprin). Ataxin-2-cIDR along with RNA-RNA interaction stabilise the stress induces RNP
555 condensation. In the absence of the PAM2 domain, Ataxin-2 fails to recruit specific target mRNA and
556 proteins. Remodelling of Ataxin-2 exposes the cIDR to induce phase separation and aberrant RNP
557 condensation. The Ataxin-2 Δ PAM2 granules are non-toxic and lack several known stress granule proteins
558 (eg. FMRP, Caprin and PABP).
559

Materials and methods

561

Key resources table

563

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Genetic reagent (<i>Drosophila melanogaster</i>)	<i>UAS-Atx2-WT-ADARed</i>	Singh <i>et al.</i> 2021	N/A	
Genetic reagent (<i>Drosophila melanogaster</i>)	<i>UAS-Atx2-ΔLSm-ADARed</i>	Singh <i>et al.</i> 2021	N/A	
Genetic reagent (<i>Drosophila melanogaster</i>)	<i>(UAS-Atx2-ΔPAM2-ADARed)</i>	This paper	N/A	
Genetic reagent (<i>Drosophila melanogaster</i>)	<i>UAS-Atx2-ΔcIDR-ADARed</i>	Singh <i>et al.</i> 2021	N/A	
Genetic reagent (<i>Drosophila melanogaster</i>)	<i>UAS-Atx2-only LSm/LSm-AD -ADARed</i>	Singh <i>et al.</i> 2021	N/A	
Genetic reagent (<i>Drosophila melanogaster</i>)	<i>mef2-Gal4; Tub-Gal80^o</i>	Bloomington <i>Drosophila</i> Stock center	Derived from #50742	
Genetic reagent (<i>Drosophila melanogaster</i>)	<i>elav-Gal4; Tub-Gal80^o</i>	Bloomington <i>Drosophila</i> Stock center	Derived from #458	
Cell line (<i>Drosophila melanogaster</i>)	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	pUAS-Atx2-SNAP_fly (Plasmid)	This paper	N/A	Construct to express fly Atx2WT-SNAP
Recombinant DNA reagent	pUAS-ATXN2-SNAP_hum (Plasmid)	This paper	N/A	Construct to express human WT ATXN2-SNAP
Recombinant DNA reagent	pUAS-mini-Atx2-SNAP_fly (Plasmid)	This paper	N/A	Construct to express fly mini Atx2-SNAP
Recombinant DNA reagent	pUAS-mini-ATXN2-SNAP_hum (Plasmid)	This paper	N/A	Construct to express human mini ATXN2-SNAP
Recombinant DNA reagent	pUAS-ΔLSm-mini-Atx2-SNAP_fly (Plasmid)	This paper	N/A	Construct to express fly ΔLSm mini Atx2-SNAP
Recombinant DNA reagent	pUAS-ΔLSm-mini-ATXN2-SNAP_hum (Plasmid)	This paper	N/A	Construct to express human ΔLSm mini ATXN2-SNAP
Recombinant DNA reagent	pUAS-ΔLSmAD-mini-Atx2-SNAP_fly (Plasmid)	This paper	N/A	Construct to express fly ΔLSm-AD mini Atx2-SNAP
Recombinant DNA reagent	pUAS-ΔLSmAD-mini-ATXN2-SNAP_hum (Plasmid)	This paper	N/A	Construct to express human ΔLSm-AD mini ATXN2-SNAP
Recombinant DNA reagent	pUAS-ΔPAM2-mini-Atx2-SNAP_fly (Plasmid)	This paper	N/A	Construct to express fly ΔPAM2 mini Atx2-SNAP
Recombinant DNA reagent	pUAS-ΔPAM2-mini-ATXN2-SNAP_hum (Plasmid)	This paper	N/A	Construct to express human ΔPAM2 mini ATXN2-SNAP

Recombinant DNA reagent	pUAS-L859A-mini-Atx2-SNAP_fly (Plasmid)	This paper	N/A	Construct to express fly PAM2* L859A mini Atx2-SNAP
Recombinant DNA reagent	pUAS-L914A-mini-ATXN2-SNAP_hum (Plasmid)	This paper	N/A	Construct to express human PAM2* L914A mini ATXN2-SNAP
Recombinant DNA reagent	pUAS-F866A-mini-Atx2-SNAP_fly (Plasmid)	This paper	N/A	Construct to express fly PAM2* F866A mini Atx2-SNAP
Recombinant DNA reagent	pUAS-F921A-mini-ATXN2-SNAP_hum (Plasmid)	This paper	N/A	Construct to express human PAM2* F921A mini ATXN2-SNAP
Recombinant DNA reagent	pUAS-L859A-F866A-mini-Atx2-SNAP_fly (Plasmid)	This paper	N/A	Construct to express fly PAM2* L859A and F866A mini Atx2-SNAP
Recombinant DNA reagent	pUAS-L914A-F921A-Q928A-mini-ATXN2-SNAP_hum (Plasmid)	This paper	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 Atx2-GFP
Recombinant DNA reagent	pAcman-ΔLSm-Atx2-GFP (Fosmid)	This paper	N/A	Genomic construct to express fly ΔLSm Atx2-GFP
Recombinant DNA reagent	pAcman-ΔPAM2-Atx2-GFP (Fosmid)	This paper	N/A	Genomic construct to express fly ΔPAM2 Atx2-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	Cat# ab21060	WB (1:1000)
Antibody	LSM12 (rabbit anti human polyclonal)	Abcam	Cat# ab173292	WB (1:1000)
Antibody	G3BP (mouse monoclonal)	BD Bioscience	Cat# 611126	IF (1:1000)
Antibody	Histone H3 (rabbit polyclonal)	Cell signaling tech	Cat# 9715	WB (1:1000)
Antibody	Baf155 (rabbit monoclonal)	Cell signaling tech	Cat# 11956	WB (1:1000)

Antibody	HRP Goat anti-rabbit	Invitrogen	Cat# A16104	WB (1:10,000)
Antibody	HRP Goat anti-mouse	Invitrogen	Cat# 31430	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	Cat# S9105S	IF (1:1000)
Chemical compound	SNAP-Surface 488	New England Biolabs	Cat# S9124S	IF (1:1000)
Software, algorithm	TRIBE	McMahon et al., 2016	https://github.com/roshashlab/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	AREScore	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	

564

565 Cell culture, transfection and stress induction

566 *Drosophila* S2R+ cells were obtained from the DGRC, Indiana University, and were grown in
567 Gibco Schneider's S2 media with 10% FBS and 1% penicillin and streptomycin, at 25°C.
568 Transfections were performed using either EugeneHD (Active Motif) or TransIT-X2 (Mirus)
569 reagents at 2:1 ratio µl reagent to µg plasmid DNA for 24-72 h depending on downstream use.
570 HEK293T cells from Adrian Bracken, Trinity College Dublin, were grown in Gibco
571 Dulbecco's Modified Eagle Media with 10% FBS, 2 mM l-glutamine addition, 1% penicillin
572 and streptomycin, at 37°C and 5% CO₂. U2OS cells from Martina Schroeder, Maynooth
573 University, were grown at the same conditions as HEK293T. Mammalian cell transfections
574 were carried out with 1 mM PEI (Polysciences) solution at 2:1 ratio µl reagent to µg plasmid
575 DNA for 24-72H depending on downstream use. For confocal imaging applications cells were
576 grown in 24-well plates on glass cover-slips for 24 h before transfection for up to 48 h. For
577 Western blotting and IP, cells were grown in 75 cm² flasks until >80% confluent before
578 transfection for up to 72H before harvesting. Oxidative stress was induced in *Drosophila* S2R+
579 cells with addition of sodium arsenite solution to a final concentration of 50 µM in media for
580 3 h. In mammalian cells, oxidative stress was induced in the same way except for only 1 h.

581

582 Western blotting and protein immunoprecipitation

583 Total protein extracts were prepared from S2 and HEK293 cells as described earlier
584 (Sudhakaran et al., 2014). Up to 10 µg total protein was loaded per well for detecting Atx2-
585 SNAP constructs, partner proteins and loading controls on 8-12% SDS-PAGE gels and
586 transferred to nitrocellulose membranes. The blots were probed in 5% skim milk in PBS using
587 rabbit anti-SNAP (1:1000), rabbit anti-PABP (1:1000), rabbit anti-LSM12 (1:1000) antibodies,
588 and mouse anti-histone H3 (1:5000) and mouse anti-BAF155 (1:2000) loading control
589 antibodies. Corresponding HRP-conjugated secondary antibodies were used at 1:10,000
590 dilution and developed using Pierce ECL western blotting substrate (ThermoFisher) as per the
591 manufacturer's instructions.

592 For Atx2-SNAP construct immunoprecipitation, transfected cell lysates were normalised to the
593 same volume and concentration, 10% of the volume was saved and diluted as an input control,
594 and Chromotek anti-SNAP-tag conjugated agarose beads and IP kits were used according to
595 the manufacturer's specifications. Pulled-down proteins together with corresponding sample
596 input controls were blotted as described above.

597

598 Immunohistochemistry and imaging of cultured cells

599 Transfected cells on coverslips were fixed with 4% paraformaldehyde in PBS solution for 15
600 min, followed by three 5 min washes in PBS. Permeabilization was performed on all cells with
601 0.5% TritonX100 in PBS solution for 3 min, before three more 5 min washes in PBS. Cells
602 were blocked with 3% BSA in PBS for 1 h at room temperature before staining with primary
603 antibodies at appropriate dilutions in 3% BSA overnight at 4°C. Corresponding fluorescent
604 secondary antibodies in 3% BSA were used to stain the sample cells for 1 h at room temperature
605 after primaries were washed off. Where SNAP-tagged proteins were being visualized, SNAP-
606 ligand TMR-Star (NEB) or SNAP-surface-Alexa488 (NEB) were added at the secondary
607 antibody staining stage. Following staining and washing, cells were mounted upside-down on
608 microscopy slides in MOWIOL, allowed to cure for >12 h at 4 °C, and imaged on a Zeiss
609 LSM880 Airyscan/AiryscanFast confocal microscope with a 20x air objective.

610

611 Bioimage analysis

612 Where relevant, Airyscan images were processed with Zen Black software (Zeiss) with
613 recommended settings. Confocal microscopy images were analysed using macros within
614 ImageJ/FIJI and Excel. Quantification of co-localisation was performed by comparing stress
615 granule marker staining intensity profiles across a randomised selection of Atx2 granules
616 within transfected cells, with the intensity profile of the Atx2 staining. Any signal 10% or
617 higher than background (adjusted for fluorophore bleed through) was deemed evidence of co-
618 localisation within that particular granule. For quantifying the exclusion of mini-Atx2-SNAP
619 constructs from stress induced granules the Caprin or G3BP1 staining was used as independent
620 identifier of stress granules and Atx2 profiles were compared to them. 48-120 granules were
621 quantified in each co-staining (Figure 2), and 28-70 granules were quantified for each construct
622 transfection (Figure 4).

623

624 Crystal structure threading

625 Threading of the *Drosophila* PAM2 peptide bound to the MLLE domain of PABPC1 was
626 performed using the Swiss-PdbViewer software, based on the human crystal structure of the
627 complex obtained from PDB, identified as 3KTR (Kozlov et al., 2010)

628

629 Experimental fly crosses

630 *Drosophila* stocks were maintained at 25°C in corn meal agar. Strains homozygous for UAS-
631 transgenes were crossed with *elav-Gal4* and *tub-Gal80ts* at 18 °C till the adult fly emerged.
632 The flies were shifted to 29 °C for 5 days before processing for RNA extraction for TRIBE
633 experiments. The climbing behaviour experiments were performed on flies kept at 29 °C for
634 either 1 or 15 days. For microCT experiments, the UAS-transgenes were crossed with *mef2-*
635 *Gal4* and *tub-Gal80ts* at 18°C and the adult flies were transferred to 29 °C for 1 day or 20 days.

636

637 RNA extraction from brain and NGS

638 Around 10-12 adult brains were dissected in RNA Later for total RNA isolation. RNA was
639 isolated using TRIzol reagent (Invitrogen) as per the manufacturer's protocol. Poly(A)-
640 enriched mRNA was used to prepare Illumina libraries using the NEBNext Ultra II Directional
641 RNA Library Prep kit (E7765L). Atx2-ΔPAM2-ADARcd samples were sequenced with
642 Illumina HiSeq PE Rapid Cluster Kit v2 (PE-402-4002) to generate 2 × 100 paired-end strand-
643 specific data using the Illumina HiSeq 2500 sequencing platform.

644

645 TRIBE data analysis

646 The sequencing reads obtained had a mean quality score (Q-Score) ≥ 37. Analysis of the
647 TRIBE data was performed as described previously (McMahon et al., 2016, Singh et al., 2021).
648 Briefly, the reference genome and gtf file of *Drosophila melanogaster*, version dm6, were
649 downloaded from the UCSC genome browser. Raw sequencing reads were mapped using
650 TopHat2 (Trapnell et al., 2009) with the parameters ‘--library-type fr-firststrand -m 1 N 3 --
651 read-edit-dist 3 p 5 g 2 -I 50000 --microexon-search --no-coverage-search -G dm6_genes.gtf’.
652 Only uniquely mapped reads are considered for editing analysis. A table of raw and mapped
653 reads is included in Supplementary Table 1. A threshold file was created by ensuring only edits
654 with coverage of at least 20 reads and 15% edits were retained. All the TRIBE experiments
655 were performed in duplicates, and only the edits identified in both replicates above the edit
656 threshold are reported.

657 Climbing Assay:

Appropriately aged adult *Drosophila* was transferred to a 50 ml graduated glass measuring cylinder for the climbing assay and sealed with a cotton plug. A digital video camera was positioned to record the vials. The assay was initiated by tapping the cylinder against a foam pad to collect the flies to the bottom of the cylinder and the flies were allowed to climb the cylinder with video being recorded for ~30 s. The number of flies that crossed the 20 ml mark (~5.5cm) was counted over time and the data was plotted against the time using GraphPad prism. Average of 3 trials were used for each biological replicate. 7-10 biological replicates were used for each genotype.

Sample preparation and scanning for microCT:

Drosophila indirect flight muscle microCT was carried out as described in Chaturvedi et. al, 2019. Briefly, animals were anesthetized on ice and fixed in PBS containing 4% paraformaldehyde (PFA). Thoraces were dissected and stained using 1% elemental iodine (1.93900.0121, Emparta, Merck) with 2% potassium iodide (no. 15 724, Qualigens) dissolved in PBS. The stained samples were washed in PBS and embedded in petroleum jelly. MicroCT scanning was carried out at 40 kV, 250 μ A, on Bruker Skyscan-1272.

Data availability

The RNA sequencing data have been deposited to GEO under the accession code GSE196739.

Contact for reagent and resource sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contacts Mani Ramaswami (mani.ramaswami@tcd.ie) and Baskar Bakthavachalu (baskar@iitmandi.ac.in).

ACKNOWLEDGEMENTS:

We thank members of the Ramaswami, Vijay Raghavan and Bakthavachalu labs and Roy Parker for useful discussions and/or comments on the manuscript. We thank Marlina Mucha, Adrian Bracken and Amir Khan for advice and help with biochemical experiments, Michael Rosbash for *Drosophila* TRIBE plasmids, and colleagues acknowledged in the Key Resources table for reagents and advice. The fly facility at Bangalore Life Science Cluster (BLiSC)

provided support with fly stock supply as well as generation of transgenic; CIFF at BLiSC provided essential confocal microscopy support; and Awadhesh Pandit and next-generation genomics facility at BLiSC provided NGS service. Daniel Fortunati thanks Kenneth Mok for his mentorship. We acknowledge Drosophila Genomics Resource Centre (supported by NIH grant 2P40OC010949) for *Drosophila* S2 cells.

696

697 **FUNDING:**

The work was supported by a Science Foundation Ireland (SFI) Investigator grant to MR and NCBS-TIFR intramural funding to KVR. BB is supported by the DBT/Wellcome Trust India Alliance Fellowship (IA/I/19/1/504286). We acknowledge the support from an Irish Research Council Postgraduate Fellowship to DF, the J. C. Bose Fellowship of the Government of India (KVR), INSA Young Scientist Project (INSA/SP/YSP/143/2017) (AS), SERB to MR from a collaborative VAJRA award to Dr. Raghu Padinjat, and a DST INSPIRE fellowship (KA). We thank C-CAMP for logistical support for SSP.

705

706 **AUTHOR CONTRIBUTIONS:**

Conceptualization, A.P., D.F, A.S., J.Huelsmeier, K.V.R., M.R., and B.B.; Methodology, A.P., D.F, A.S., J.Huelsmeier, A.R.K., S.S.P., J.Hillebrand, K.A., D.J., G.B., J.L., C.L., G.A., K.H.M., K.V.R., M.R., and B.B.; Investigation, A.S., J.Huelsmeier, A.R.K., S.S.P., J.Hillebrand, A.P., K.A., K.V.R., M.R., and B.B.; Writing–Original Draft, A.P., D.F, A.S., J.Huelsmeier, K.V.R., M.R., and B.B.; Writing–Review & Editing, A.P., D.F, A.S., J.Huelsmeier, A.R.K., S.S.P., J.Hillebrand, K.A., D.J., G.B., J.L., C.L., G.A., K.H.M., K.V.R., M.R., and B.B.; Funding Acquisition, K.V.R., M.R., and B.B.; Resources, Fly community.

714

DECLARATION OF INTERESTS: The authors declare no conflicts of interest.

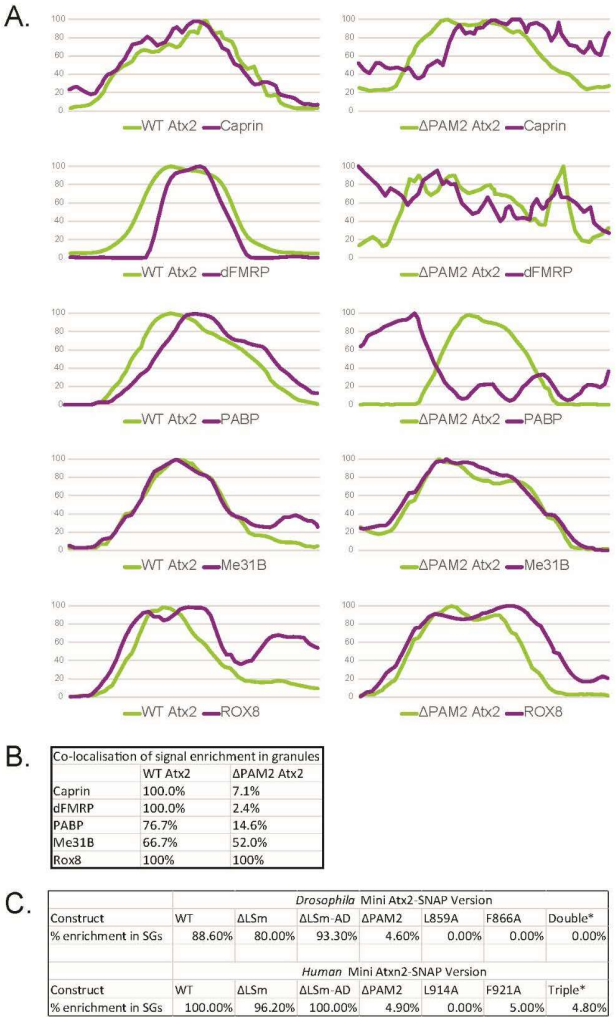
716

717

718

SUPPLEMENTARY FIGURES:

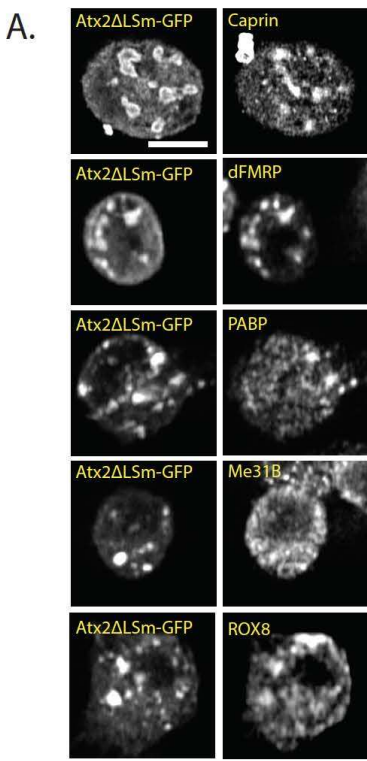
Supplementary Figure 1:



Supplementary Figure 1: Co-localisation quantification for figure 2, figure 4. (A) Normalised profile plots of Atx2-GFP granules in S2 cells as shown in Figure 2. 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). In Atx2ΔPAM2-GFP granules,

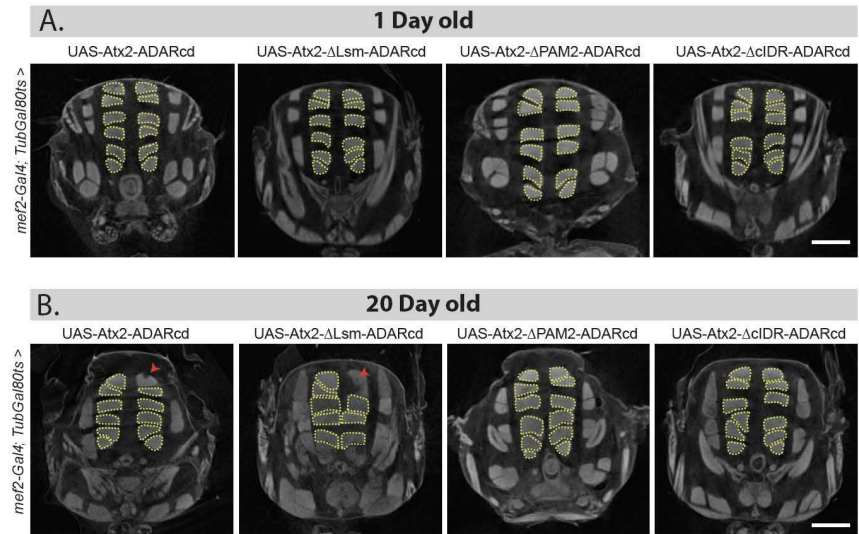
727 this colocalization of fluorescence signals is not seen in the case of Caprin, dFMRP and PABP,
728 suggesting these components are not enriched in these granules above background level. (B)
729 Quantification of co-localization for Figure 2. N = 48-120 images of Atx2-GFP granules were
730 randomly selected for each co-staining and analysed for signal co-enrichment (see methods) in
731 the case of each component assayed. (C) Quantification of Atx2 construct inclusion in SGs for
732 Figure 4. N = 28-70 images of stress granules in arsenite stressed S2 cells (marked by anti-
733 Caprin staining) and U2OS cells (marked by anti-G3BP staining) were randomly selected for
734 each Atx2 construct assayed and were analysed for Mini Atx2-SNAP allele signal co-
735 enrichment (see methods).
736

737 **Supplementary Figure 2:**



738
739 **Supplementary Figure 2:** Atx2ΔLSm granules in S2 cells do not show significantly altered
740 protein contents compared to wild-type Atx2. Caprin, dFMRP, PABP, Me31B, and Rox8 co-
741 localize with overexpressed Atx2ΔLSm GFP, suggesting that the granules formed contain a
742 similar set of components as Atx2 granules. It should be noted that Atx2 granules do not
743 sequester the majority of the endogenous components stained for, leading to a high, diffuse
744 background staining.
745

758 **Supplementary Figure 4:**



759
760 **Supplementary Figure 4:** Transverse view of *Drosophila* indirect flight muscle imaged using
761 micro-CT show cellular toxicity. (A) Driving UAS-transgene (Atx2WT, Atx2ΔcIDR,
762 Atx2ΔPAM or Atx2ΔLsm) with *mef2-Gal4* show normal muscles on day 1. (B) Expression of
763 wild-type and Atx2ΔLsm transgene for 20 days show loss of muscle fibers, indicated with
764 solid red arrowheads. Expression of Atx2ΔPAM2 and Atx2ΔcIDR for 20 days show no visible
765 phenotype.
766

Supplementary Table 1:

767
768

S.NO	Sample Name	Million reads (Paired end)		Total Number of Million reads	% >= Q30	Mean Quality Score	Mapping Percentage Read 1	Mapping Percentage Read 2
		Read 1	Read 2					
1	Atx2ΔPAM2-ADARcd rep_1	23.75	23.75	47	97.04	38.26	92.20%	92.10%
2	Atx2ΔPAM2-ADARcd rep_2	22.66	22.66	45	97.03	38.26	91.30%	91.20%

769

770

771

772

773

Supplementary Table 2: The targets common between Atx2 wild-type and del-PAM2 are shown in bold text.

Chr	Start	End	Genes	Replicate 1 edit percentage	Replicate 2 edit percentage	Average edit percentage	Chr coordinate
chr2L	1009080	1009081	IA-2	17.1	15.5	16.3	chr2L_1009081_IA-2
chr2L	1009440	1009441	IA-2	31.7	38.8	35.25	chr2L_1009441_IA-2
chr2L	1011482	1011483	IA-2	31.7	34.7	33.2	chr2L_1011483_IA-2
chr2L	13505104	13505105	B4	22.4	17.48	19.94	chr2L_13505105_B4
chr2L	17188307	17188308	beat-IIIc	18	21.7	19.85	chr2L_17188308_beat-IIIc
chr2L	19746161	19746162	CG10631	25.65	19.4	22.52	chr2L_19746162_CG10631
chr2L	20056018	20056019	sNPF	16	24.3	20.15	chr2L_20056019_sNPF
chr2L	7891375	7891376	Snoo	18.5	16.7	17.6	chr2L_7891376_Snoo
chr2L	8113160	8113161	Bsg	17.4	19.05	18.23	chr2L_8113161_Bsg
chr2L	8113894	8113895	Bsg	16.82	28.7	22.76	chr2L_8113895_Bsg
chr2L	8116005	8116006	Bsg	18.2	27.4	22.8	chr2L_8116006_Bsg
chr2L	9256901	9256902	Ggamma30A	19.5	32.6	26.05	chr2L_9256902_Ggamma30A
chr2L	9292530	9292531	Ggamma30A	42	54.4	48.2	chr2L_9292531_Ggamma30A
chr2L	9292829	9292830	Ggamma30A	42.6	53.55	48.08	chr2L_9292830_Ggamma30A
chr2L	9295031	9295032	Ggamma30A	23.5	26	24.75	chr2L_9295032_Ggamma30A
chr2R	12085463	12085464	jeb	16	28.4	22.2	chr2R_12085464_jeb
chr2R	13513577	13513578	Vmat	64.2	63.9	64.05	chr2R_13513578_Vmat
chr2R	13513654	13513655	Vmat	27.7	33.45	30.58	chr2R_13513655_Vmat
chr2R	13513732	13513733	Vmat	24	29.75	26.88	chr2R_13513733_Vmat
chr2R	13513762	13513763	Vmat	44.55	47.3	45.92	chr2R_13513763_Vmat
chr2R	13514350	13514351	Vmat	37.7	38.5	38.1	chr2R_13514351_Vmat
chr2R	13519080	13519081	Vmat	18.1	20.8	19.45	chr2R_13519081_Vmat
chr2R	23690701	23690702	Pal2	17.3	17.2	17.25	chr2R_23690702_Pal2
chr2R	24213943	24213944	CG30419	27.8	22.98	25.39	chr2R_24213944_CG30419
chr2R	24214234	24214235	CG30419	24.2	26.1	25.15	chr2R_24214235_CG30419
chr2R	24214648	24214649	CG30419	16.4	15.9	16.15	chr2R_24214649_CG30419
chr2R	24215193	24215194	CG30419	43.5	29.7	36.6	chr2R_24215194_CG30419
chr2R	24229341	24229342	CG30419	20.8	16.3	18.55	chr2R_24229342_CG30419
chr2R	6914804	6914805	CG30158	16.4	16	16.2	chr2R_6914805_CG30158
chr2R	6920815	6920816	CG30158	33.92	40.48	37.2	chr2R_6920816_CG30158
chr2R	6921435	6921436	CG30158	17.5	41.9	29.7	chr2R_6921436_CG30158
chr2R	6921976	6921977	CG30158	36	40.2	38.1	chr2R_6921977_CG30158
chr2R	7718070	7718071	CG18812	15.9	19.7	17.8	chr2R_7718071_CG18812

chr2R	9473772	9473773	Camta	21.75	30.18	25.96	chr2R_9473773_Camta
chr2R	9479421	9479422	Camta	23.1	26	24.55	chr2R_9479422_Camta
chr2R	9480019	9480020	Camta	32	41	36.5	chr2R_9480020_Camta
chr2R	9910328	9910329	FMRFa	42.5	25	33.75	chr2R_9910329_FMRFa
chr3L	11498209	11498210	chrh	15.82	20.5	18.16	chr3L_11498210_chrh
chr3L	12267895	12267896	CG32100	20	16.1	18.05	chr3L_12267896_CG32100
chr3L	1504295	1504296	Psa	18.6	15.9	17.25	chr3L_1504296_Psa
chr3L	1521650	1521651	Psa	18.5	34.6	26.55	chr3L_1521651_Psa
chr3L	1543675	1543676	CG7852	15.5	22	18.75	chr3L_1543676_CG7852
chr3L	17062355	17062356	Rbp6	15.2	18.5	16.85	chr3L_17062356_Rbp6
chr3L	17147382	17147383	Rbp6	18.23	26.62	22.42	chr3L_17147383_Rbp6
chr3L	17345219	17345220	Mip	29.9	27.2	28.55	chr3L_17345220_Mip
chr3L	17345290	17345291	Mip	17.9	16.9	17.4	chr3L_17345291_Mip
chr3L	19066983	19066984	Mkp3	18.9	15.7	17.3	chr3L_19066984_Mkp3
chr3L	21494821	21494822	Hr78	27.1	18.4	22.75	chr3L_21494822_Hr78
chr3L	21831417	21831418	CG7148	15.8	28.6	22.2	chr3L_21831418_CG7148
chr3L	21930851	21930852	mub	15	17.2	16.1	chr3L_21930852_mub
chr3L	21931110	21931111	mub	15.1	20.1	17.6	chr3L_21931111_mub
chr3L	22061206	22061207	Oet-TyrR	15.4	16.7	16.05	chr3L_22061207_Oet-TyrR
chr3L	22877661	22877662	Chro	35.3	45.7	40.5	chr3L_22877662_Chro
chr3L	23148124	23148125	CG32350	27.4	35.3	31.35	chr3L_23148125_CG32350
chr3L	23747549	23747550	CG17698	40.27	28.95	34.61	chr3L_23747550_CG17698
chr3L	23934990	23934991	CG40470	23.5	22.05	22.77	chr3L_23934991_CG40470
chr3L	3910071	3910072	Eip63F-1	20	24.5	22.25	chr3L_3910072_Eip63F-1
chr3L	3954338	3954339	CG12605	35.2	40.8	38	chr3L_3954339_CG12605
chr3L	3954933	3954934	CG12605	35.42	51.1	43.26	chr3L_3954934_CG12605
chr3L	3957068	3957069	CG12605	21.8	26.7	24.25	chr3L_3957069_CG12605
chr3L	3957671	3957672	CG12605	18.9	22.2	20.55	chr3L_3957672_CG12605
chr3L	3961590	3961591	CG12605	18.3	18.5	18.4	chr3L_3961591_CG12605
chr3L	3992789	3992790	sert	21.08	27.28	24.18	chr3L_3992790_sert
chr3L	4092142	4092143	CG14989	15.88	18.6	17.24	chr3L_4092143_CG14989
chr3L	4113123	4113124	Ack	18.4	16.7	17.55	chr3L_4113124_Ack
chr3L	4113297	4113298	Ack	18.4	17.9	18.15	chr3L_4113298_Ack
chr3L	572527	572528	hipk	29.8	31	30.4	chr3L_572528_hipk
chr3L	572530	572531	hipk	42.7	43	42.85	chr3L_572531_hipk
chr3L	575712	575713	hipk	36.12	52.92	44.52	chr3L_575713_hipk
chr3L	575753	575754	hipk	31.7	49.6	40.65	chr3L_575754_hipk
chr3L	576730	576731	hipk	17.1	15.2	16.15	chr3L_576731_hipk
chr3L	577020	577021	hipk	20	16.9	18.45	chr3L_577021_hipk
chr3L	577417	577418	hipk	51.42	62.12	56.77	chr3L_577418_hipk
chr3L	577970	577971	hipk	21.4	25.9	23.65	chr3L_577971_hipk
chr3L	578186	578187	hipk	18.2	24.9	21.55	chr3L_578187_hipk
chr3L	578453	578454	hipk	25.7	32.7	29.2	chr3L_578454_hipk
chr3L	579335	579336	hipk	18.1	17	17.55	chr3L_579336_hipk
chr3L	579634	579635	hipk	20.5	19.8	20.15	chr3L_579635_hipk

chr3L	580103	580104	hipk	63.1	67.7	65.4	chr3L_580104_hipk
chr3L	580500	580501	hipk	15.6	18.1	16.85	chr3L_580501_hipk
chr3L	580932	580933	hipk	17.9	24.1	21	chr3L_580933_hipk
chr3L	8970787	8970788	CG5026	17.9	16.7	17.3	chr3L_8970788_CG5026
chr3L	8993382	8993383	smg	25.9	15.3	20.6	chr3L_8993383_smg
chr3L	9074752	9074753	Tequila	29.8	19.5	24.65	chr3L_9074753_Tequila
chr3L	9103274	9103275	bol	23.9	26.3	25.1	chr3L_9103275_bol
chr3L	9136454	9136455	Use1	25	16	20.5	chr3L_9136455_Use1
chr3L	9669496	9669497	frv	18.2	18.2	18.2	chr3L_9669497_frv
chr3L	9945905	9945906	CG34356	19.4	32.1	25.75	chr3L_9945906_CG34356
chr3R	10158991	10158992	Invadolysin	17.9	18.2	18.05	chr3R_10158992_Invadolysin
chr3R	10862820	10862821	CG6574	20.8	33.3	27.05	chr3R_10862821_CG6574
chr3R	10877257	10877258	CR45195	32.7	19.5	26.1	chr3R_10877258_CR45195
chr3R	10889941	10889942	Leash	18.2	18.8	18.5	chr3R_10889942_Leash
chr3R	13224637	13224638	Ace	24.8	33.3	29.05	chr3R_13224638_Ace
chr3R	13227870	13227871	Ace	26.6	35.3	30.95	chr3R_13227871_Ace
chr3R	13227970	13227971	Ace	33.42	47.95	40.69	chr3R_13227971_Ace
chr3R	14330745	14330746	NK7.1	16.2	16.7	16.45	chr3R_14330746_NK7.1
chr3R	14660839	14660840	Hexim	17.1	20	18.55	chr3R_14660840_Hexim
chr3R	14669984	14669985	Meltrin	16.95	20.7	18.82	chr3R_14669985_Meltrin
chr3R	14746335	14746336	jvl	36.4	20.9	28.65	chr3R_14746336_jvl
chr3R	14746546	14746547	smg-30	45.1	24	34.55	chr3R_14746547_smg-30
chr3R	14804349	14804350	btsz	16.5	26.5	21.5	chr3R_14804350_btsz
chr3R	15255687	15255688	CG42404	18.2	17.6	17.9	chr3R_15255688_CG42404
chr3R	15356475	15356476	Atg4b	15.4	20.7	18.05	chr3R_15356476_Atg4b
chr3R	15414730	15414731	Atx2	22.82	26.25	24.54	chr3R_15414731_At2
chr3R	15414731	15414732	Atx2	19	22.6	20.8	chr3R_15414732_At2
chr3R	15849417	15849418	cv-d	18.8	22.2	20.5	chr3R_15849418_cv-d
chr3R	16611267	16611268	NPF	22.98	23.52	23.25	chr3R_16611268_NPF
chr3R	16645743	16645744	CG10324	26.1	35	30.55	chr3R_16645744_CG10324
chr3R	17090737	17090738	cal1	25	15.4	20.2	chr3R_17090738_cal1
chr3R	17731884	17731885	Lgr1	20	19	19.5	chr3R_17731885_Lgr1
chr3R	17802763	17802764	CG17806	20.7	18.8	19.75	chr3R_17802764_CG17806
chr3R	19155252	19155253	CG11779	15.6	19.1	17.35	chr3R_19155253_CG11779
chr3R	20783186	20783187	Syp	21.1	17	19.05	chr3R_20783187_Syp
chr3R	20797469	20797470	Syp	15.3	20	17.65	chr3R_20797470_Syp
chr3R	20820785	20820786	CG17271	31.2	19.75	25.48	chr3R_20820786_CG17271
chr3R	20862212	20862213	CG3822	17	26.3	21.65	chr3R_20862213_CG3822
chr3R	20992668	20992669	Calx	17.8	27.5	22.65	chr3R_20992669_Calx
chr3R	21213891	21213892	SNF4Agamm a	17.2	20	18.6	chr3R_21213892_SNF4Agamm a
chr3R	21354284	21354285	mod(mdg4)	15	19	17	chr3R_21354285_mod(mdg4)
chr3R	21527750	21527751	CG7956	30.3	26.5	28.4	chr3R_21527751_CG7956
chr3R	21625731	21625732	E2f	18	23.7	20.85	chr3R_21625732_E2f
chr3R	23261698	23261699	orb	15.2	16.1	15.65	chr3R_23261699_orb

chr3R	23681818	23681819	eIF-3p66	20.4	21.5	20.95	chr3R_23681819_eIF-3p66
chr3R	23698883	23698884	pri	15.4	23.4	19.4	chr3R_23698884_pri
chr3R	23723418	23723419	CG10365	16.1	21.1	18.6	chr3R_23723419_CG10365
chr3R	23732353	23732354	Rpn9	20.4	16.7	18.55	chr3R_23732354_Rpn9
chr3R	24664388	24664389	slo	23.1	30	26.55	chr3R_24664389_slo
chr3R	24802164	24802165	polybromo	29.2	20.98	25.09	chr3R_24802165_polybromo
chr3R	24820485	24820486	Saf-B	15	22	18.5	chr3R_24820486_Saf-B
chr3R	25234647	25234648	CG10420	22.2	34.5	28.35	chr3R_25234648_CG10420
chr3R	26233920	26233921	CG12290	25.6	18.4	22	chr3R_26233921_CG12290
chr3R	28050111	28050112	CG34362	15.6	25.6	20.6	chr3R_28050112_CG34362
chr3R	28838531	28838532	Ape	20.4	18.8	19.6	chr3R_28838532_Ape
chr3R	29659085	29659086	Dop1R2	15.8	39.5	27.65	chr3R_29659086_Dop1R2
chr3R	29674515	29674516	Bub3	27.6	17.8	22.7	chr3R_29674516_Bub3
chr3R	31457698	31457699	Gprk2	23.5	24.5	24	chr3R_31457699_Gprk2
chr3R	31841367	31841368	RhoGAP100F	17.73	19	18.37	chr3R_31841368_RhoGAP100F
chr3R	5811503	5811504	CG11000	21.3	32.4	26.85	chr3R_5811504_CG11000
chr3R	5811505	5811506	CG11000	20.8	25.7	23.25	chr3R_5811506_CG11000
chr3R	7126383	7126384	CG10098	18.6	19.1	18.85	chr3R_7126384_CG10098
chr3R	8244435	8244436	CG18749	15.4	16.1	15.75	chr3R_8244436_CG18749
chr3R	8244435	8244436	CG33722	15.4	16.1	15.75	chr3R_8244436_CG33722
chr3R	9416762	9416763	alpha-Man-II	30.8	32.3	31.55	chr3R_9416763_alpha-Man-II
chr3R	9441505	9441506	ps	37.92	33.52	35.72	chr3R_9441506_ps
chr3R	9471135	9471136	CG16779	43.05	34.85	38.95	chr3R_9471136_CG16779
chr3R	9525979	9525980	CG8176	16.2	24.4	20.3	chr3R_9525980_CG8176
chr3R	9539464	9539465	mura	21.75	20.05	20.9	chr3R_9539465_mura
chr3R	9794186	9794187	CG8516	19.4	30.8	25.1	chr3R_9794187_CG8516
chr4	478956	478957	Asator	15.1	20	17.55	chr4_478957_Asator
chr4	532906	532907	zfh2	20	19.7	19.85	chr4_532907_zfh2
chr4	92946	92947	pan	22.5	20.8	21.65	chr4_92947_pan
chrX	10309093	10309094	alpha-Man-I	24.5	29	26.75	chrX_10309094_alpha-Man-I
chrX	12331302	12331303	Ten-a	34.8	34	34.4	chrX_12331303_Ten-a
chrX	16075472	16075473	Tob	23.68	29.32	26.5	chrX_16075473_Tob
chrX	16075887	16075888	Tob	22.2	22	22.1	chrX_16075888_Tob
chrX	16075888	16075889	Tob	21.6	24	22.8	chrX_16075889_Tob
chrX	16076959	16076960	Tob	20	23.3	21.65	chrX_16076960_Tob
chrX	16077193	16077194	Tob	17.85	19.15	18.5	chrX_16077194_Tob
chrX	16089317	16089318	Tob	15.2	16.1	15.65	chrX_16089318_Tob
chrX	3321068	3321069	dnc	22.7	28.9	25.8	chrX_3321069_dnc
chrX	3342369	3342370	dnc	17.4	22.2	19.8	chrX_3342370_dnc
chrX	6325433	6325434	CG15894	32	20.8	26.4	chrX_6325434_CG15894
chrX	9172798	9172799	mei-P26	17.3	23.1	20.2	chrX_9172799_mei-P26
chrX	9179452	9179453	mei-P26	17.35	22.23	19.79	chrX_9179453_mei-P26
chrX	9188891	9188892	mei-P26	15.8	18.4	17.1	chrX_9188892_mei-P26

774
775

BIBLIOGRAPHY:

- Alberti S, Gladfelter A, Mittag T (2019) Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* 176: 419-434
- Alberti S, Mateju D, Mediani L, Carra S (2017) Granulostasis: Protein Quality Control of RNP Granules. *Frontiers in Molecular Neuroscience* 10
- Andrusiak MG, Sharifnia P, Lyu X, Wang Z, Dickey AM, Wu Z, Chisholm AD, Jin Y (2019) Inhibition of Axon Regeneration by Liquid-like TIAR-2 Granules. *Neuron* 104: 290-304.e298
- Ash PEA, Lei S, Shattuck J, Boudeau S, Carlomagno Y, Medalla M, Mashimo BL, Socorro G, Al-Mohanna LFA, Jiang L *et al* (2021) TIA1 potentiates tau phase separation and promotes generation of toxic oligomeric tau. *Proceedings of the National Academy of Sciences* 118: e2014188118
- Azkanaz M, Rodríguez López A, De Boer B, Huiting W, Angrand P-O, Vellenga E, Kampinga HH, Bergink S, Martens JH, Schuringa JJ *et al* (2019) Protein quality control in the nucleolus safeguards recovery of epigenetic regulators after heat shock. *eLife* 8
- Babinchak WM, Surewicz WK (2020) Liquid-Liquid Phase Separation and Its Mechanistic Role in Pathological Protein Aggregation. *Journal of Molecular Biology* 432: 1910-1925
- Bah A, Forman-Kay JD (2016) Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications. *Journal of Biological Chemistry* 291: 6696-6705
- Bah A, Vernon RM, Siddiqui Z, Krzeminski M, Muhandiram R, Zhao C, Sonenberg N, Kay LE, Forman-Kay JD (2015) Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* 519: 106-109
- Bakthavachalu B, Huelsmeier J, Sudhakaran IP, Hillebrand J, Singh A, Petrauskas A, Thiagarajan D, Sankaranarayanan M, Mizoue L, Anderson EN *et al* (2018) RNP-Granule Assembly via Ataxin-2 Disordered Domains Is Required for Long-Term Memory and Neurodegeneration. *Neuron* 98: 754-766.e754
- Becker LA, Huang B, Bieri G, Ma R, Knowles DA, Jafar-Nejad P, Messing J, Kim HJ, Soriano A, Auburger G *et al* (2017) Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice. *Nature* 544: 367-371
- Berlow RB, Dyson HJ, Wright PE (2015) Functional advantages of dynamic protein disorder. *FEBS Lett* 589: 2433-2440
- Bevilacqua PC, Williams AM, Chou H-L, Assmann SM (2022) RNA multimerization as an organizing force for liquid-liquid phase separation. *RNA* 28: 16-26
- Biogen, 2021. <https://clinicaltrials.gov/ct2/show/NCT04494256>.
- Boeynaems S, Dorone Y, Marian A, Shabardina V, Huang G, Kim G, Sanyal A, Şen N-E, Docampo R, Ruiz-Trillo I *et al* (2021) Poly(A)-binding protein is an ataxin-2 chaperone that emulsifies biomolecular condensates. *bioRxiv*: 2021.2008.2023.457426

826
827 Brandmann T, Fakim H, Padamsi Z, Youn JY, Gingras AC, Fabian MR, Jinek M (2018)
828 Molecular architecture of LSM14 interactions involved in the assembly of mRNA silencing
829 complexes. *Embo j* 37
830
831 Buchan JR (2014) mRNP granules. *RNA Biology* 11: 1019-1030
832
833 Calabretta S, Richard S (2015) Emerging Roles of Disordered Sequences in RNA-Binding
834 Proteins. *Trends Biochem Sci* 40: 662-672
835
836 Cao X, Jin X, Liu B (2020) The involvement of stress granules in aging and aging-associated
837 diseases. *Aging Cell* 19
838
839 Chou A, Krukowski K, Jopson T, Zhu PJ, Costa-Mattioli M, Walter P, Rosi S (2017) Inhibition
840 of the integrated stress response reverses cognitive deficits after traumatic brain injury. *Proc*
841 *Natl Acad Sci U S A* 114: E6420-e6426
842
843 Cirulli ET, Lasseigne BN, Petrovski S, Sapp PC, Dion PA, Leblond CS, Couthouis J, Lu YF,
844 Wang Q, Krueger BJ *et al* (2015) Exome sequencing in amyotrophic lateral sclerosis identifies
845 risk genes and pathways. *Science* 347: 1436-1441
846
847 Couthouis J, Hart MP, Erion R, King OD, Diaz Z, Nakaya T, Ibrahim F, Kim H-J, Mojsilovic-
848 Petrovic J, Panossian S *et al* (2012) Evaluating the role of the FUS/TLS-related gene EWSR1
849 in amyotrophic lateral sclerosis. *Human Molecular Genetics* 21: 2899-2911
850
851 Damrath E, Heck MV, Gispert S, Azizov M, Nowock J, Seifried C, Rüb U, Walter M, Auburger
852 G (2012) ATXN2-CAG42 Sequesters PABPC1 into Insolubility and Induces FBXW8 in
853 Cerebellum of Old Ataxic Knock-In Mice. *PLoS Genetics* 8: e1002920
854
855 De Graeve F, Besse F (2018) Neuronal RNP granules: from physiological to pathological
856 assemblies. *Biological Chemistry* 399: 623-635
857
858 Decker CJ, Teixeira D, Parker R (2007) Edc3p and a glutamine/asparagine-rich domain of
859 Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *J Cell Biol* 179:
860 437-449
861
862 Deo RC, Bonanno JB, Sonenberg N, Burley SK (1999) Recognition of Polyadenylate RNA by
863 the Poly(A)-Binding Protein. *Cell* 98: 835-845
864
865 Elden AC, Kim HJ, Hart MP, Chen-Plotkin AS, Johnson BS, Fang X, Armakola M, Geser F,
866 Greene R, Lu MM *et al* (2010) Ataxin-2 intermediate-length polyglutamine expansions are
867 associated with increased risk for ALS. *Nature* 466: 1069-1075
868
869 Formicola N, Vijayakumar J, Besse F (2019) Neuronal ribonucleoprotein granules: Dynamic
870 sensors of localized signals. *Traffic* 20: 639-649
871
872 Gilks N, Kedersha N, Ayodele M, Shen L, Stoecklin G, Dember LM, Anderson P (2004) Stress
873 granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell* 15: 5383-5398

874 Ginsberg SD, Galvin JE, Chiu TS, Lee VM, Masliah E, Trojanowski JQ (1998) RNA
875 sequestration to pathological lesions of neurodegenerative diseases. *Acta Neuropathol* 96: 487-
876 494
877
878 Gomes E, Shorter J (2019) The molecular language of membraneless organelles. *J Biol Chem*
879 294: 7115-7127
880
881 Hachet O, Ephrussi A (2004) Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic
882 localization. *Nature* 428: 959-963
883
884 Halliday M, Radford H, Zents KAM, Molloy C, Moreno JA, Verity NC, Smith E, Ortori CA,
885 Barrett DA, Bushell M *et al* (2017) Repurposed drugs targeting eIF2 α -P-mediated translational
886 repression prevent neurodegeneration in mice. *Brain* 140: 1768-1783
887
888 Han TW, Kato M, Xie S, Wu LC, Mirzaei H, Pei J, Chen M, Xie Y, Allen J, Xiao G *et al*
889 (2012) Cell-free formation of RNA granules: bound RNAs identify features and components
890 of cellular assemblies. *Cell* 149: 768-779
891
892 Harlen KM, Churchman LS (2017) The code and beyond: transcription regulation by the RNA
893 polymerase II carboxy-terminal domain. *Nature Reviews Molecular Cell Biology* 18: 263-273
894
895 Hetz C, Zhang K, Kaufman RJ (2020) Mechanisms, regulation and functions of the unfolded
896 protein response. *Nature Reviews Molecular Cell Biology* 21: 421-438
897
898 Hochberg-Laufer H, Schwed-Gross A, Neugebauer KM, Shav-Tal Y (2019) Uncoupling of
899 nucleo-cytoplasmic RNA export and localization during stress. *Nucleic Acids Research* 47:
900 4778-4797
901
902 Hofweber M, Dormann D (2019) Friend or foe—Post-translational modifications as regulators
903 of phase separation and RNP granule dynamics. *Journal of Biological Chemistry* 294: 7137-
904 7150
905
906 Huelsmeier J, Walker E, Bakthavachalu B, Ramaswami M (2021) A C-terminal ataxin-2
907 disordered region promotes Huntingtin protein aggregation and neurodegeneration in
908 *Drosophila* models of Huntington's disease. *G3 (Bethesda)* 11
909
910 Hyman AA, Weber CA, Jülicher F (2014) Liquid-Liquid Phase Separation in Biology. *Annual*
911 *Review of Cell and Developmental Biology* 30: 39-58
912
913 Inagaki H, Hosoda N, Tsuiji H, Hoshino S-I (2020) Direct evidence that Ataxin-2 is a
914 translational activator mediating cytoplasmic polyadenylation. *Journal of Biological*
915 *Chemistry* 295: 15810-15825
916
917 Ivanov P, Kedersha N, Anderson P (2019) Stress Granules and Processing Bodies in
918 Translational Control. *Cold Spring Harb Perspect Biol* 11
919
920 Jain A, Vale RD (2017) RNA phase transitions in repeat expansion disorders. *Nature* 546: 243-
921 247
922

923 Järvelin AI, Noerenberg M, Davis I, Castello A (2016) The new (dis)order in RNA regulation.
924 *Cell Commun Signal* 14: 9
925
926 Jiang S, Fagman JB, Chen C, Alberti S, Liu B (2020) Protein phase separation and its role in
927 tumorigenesis. *eLife* 9
928
929 Jiménez-López D, Guzmán P (2014) Insights into the evolution and domain structure of ataxin-
930 2 proteins across eukaryotes. *BMC Research Notes* 7: 453
931
932 Kaehler C, Isensee J, Nonhoff U, Terrey M, Hucho T, Lehrach H, Krobisch S (2012) Ataxin-
933 2-like is a regulator of stress granules and processing bodies. *PLoS One* 7: e50134
934
935 Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, Mirzaei H, Goldsmith EJ, Longgood J, Pei J
936 *et al* (2012) Cell-free formation of RNA granules: low complexity sequence domains form
937 dynamic fibers within hydrogels. *Cell* 149: 753-767
938
939 Kedersha N, Anderson P (2007) Mammalian stress granules and processing bodies. *Methods*
940 *Enzymol* 431: 61-81
941
942 Kedersha N, Panas MD, Achorn CA, Lyons S, Tisdale S, Hickman T, Thomas M, Lieberman
943 J, McInerney GM, Ivanov P *et al* (2016) G3BP-Caprin1-USP10 complexes mediate stress
944 granule condensation and associate with 40S subunits. *J Cell Biol* 212: 845-860
945
946 Kedersha NL, Gupta M, Li W, Miller I, Anderson P (1999) RNA-Binding Proteins Tia-1 and
947 Tiar Link the Phosphorylation of Eif-2 α to the Assembly of Mammalian Stress Granules.
948 *Journal of Cell Biology* 147: 1431-1442
949
950 Khong A, Parker R (2020) The landscape of eukaryotic mRNPs. *RNA* 26: 229-239
951
952 Kiebler MA, Bassell GJ (2006) Neuronal RNA granules: movers and makers. *Neuron* 51: 685-
953 690
954
955 Kim G, Gautier O, Tassoni-Tsuchida E, Ma XR, Gitler AD (2020) ALS Genetics: Gains,
956 Losses, and Implications for Future Therapies. *Neuron* 108: 822-842
957
958 Kim H-J, Raphael AR, Ladow ES, Megurk L, Weber RA, Trojanowski JQ, Lee VM-Y,
959 Finkbeiner S, Gitler AD, Bonini NM (2014) Therapeutic modulation of eIF2 α phosphorylation
960 rescues TDP-43 toxicity in amyotrophic lateral sclerosis disease models. *Nature Genetics* 46:
961 152-160
962
963 Kim HJ, Kim NC, Wang YD, Scarborough EA, Moore J, Diaz Z, MacLea KS, Freibaum B, Li
964 S, Molliex A *et al* (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1
965 cause multisystem proteinopathy and ALS. *Nature* 495: 467-473
966
967 Kim TH, Payliss BJ, Nosella ML, Lee ITW, Toyama Y, Forman-Kay JD, Kay LE (2021)
968 Interaction hot spots for phase separation revealed by NMR studies of a CAPRIN1 condensed
969 phase. *Proc Natl Acad Sci U S A* 118
970
971 Knowles RB, Sabry JH, Martone ME, Deerinck TJ, Ellisman MH, Bassell GJ, Kosik KS (1996)
972 Translocation of RNA Granules in Living Neurons. *The Journal of Neuroscience* 16: 7812

973
974 Kozlov G, Safaei N, Rosenauer A, Gehring K (2010) Structural basis of binding of P-body-
975 associated proteins GW182 and ataxin-2 by the Mle domain of poly(A)-binding protein. *J Biol*
976 *Chem* 285: 13599-13606
977
978 Kwon I, Kato M, Xiang S, Wu L, Theodoropoulos P, Mirzaei H, Han T, Xie S, Corden Jeffry L,
979 McKnight Steven L (2013) Phosphorylation-Regulated Binding of RNA Polymerase II to
980 Fibrous Polymers of Low-Complexity Domains. *Cell* 155: 1049-1060
981
982 Latonen L (2019) Phase-to-Phase With Nucleoli – Stress Responses, Protein Aggregation and
983 Novel Roles of RNA. *Frontiers in Cellular Neuroscience* 13: 151
984
985 Lee J, Yoo E, Lee H, Park K, Hur JH, Lim C (2017) LSM12 and ME31B/DDX6 Define
986 Distinct Modes of Posttranscriptional Regulation by ATAXIN-2 Protein Complex in
987 *Drosophila* Circadian Pacemaker Neurons. *Mol Cell* 66: 129-140.e127
988
989 Lee KH, Zhang P, Kim HJ, Mitrea DM, Sarkar M, Freibaum BD, Cika J, Coughlin M, Messing
990 J, Molliex A *et al* (2016) C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and
991 Function of Membrane-Less Organelles. *Cell* 167: 774-788.e717
992
993 Li YR, King OD, Shorter J, Gitler AD (2013) Stress granules as crucibles of ALS pathogenesis.
994 *J Cell Biol* 201: 361-372
995
996 Lim C, Allada R (2013) ATAXIN-2 activates PERIOD translation to sustain circadian rhythms
997 in *Drosophila*. *Science* 340: 875-879
998
999 Lin Y, Currie SL, Rosen MK (2017) Intrinsically disordered sequences enable modulation of
1000 protein phase separation through distributed tyrosine motifs. *J Biol Chem* 292: 19110-19120
1001
1002 Lin Y, Protter SW, David, Rosen K, Michael, Parker R (2015) Formation and Maturation of
1003 Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Molecular Cell* 60: 208-219
1004
1005 Liu EY, Cali CP, Lee EB (2017) RNA metabolism in neurodegenerative disease. *Dis Model*
1006 *Mech* 10: 509-518
1007
1008 Machida K, Shigeta T, Yamamoto Y, Ito T, Svitkin Y, Sonenberg N, Imataka H (2018)
1009 Dynamic interaction of poly(A)-binding protein with the ribosome. *Scientific Reports* 8: 17435
1010 Mallucci GR, Klenerman D, Rubinsztein DC (2020) Developing Therapies for
1011 Neurodegenerative Disorders: Insights from Protein Aggregation and Cellular Stress
1012 Responses. *Annual Review of Cell and Developmental Biology* 36: 165-189
1013
1014 Mandrioli J, Mediani L, Alberti S, Carra S (2020) ALS and FTD: Where RNA metabolism
1015 meets protein quality control. *Semin Cell Dev Biol* 99: 183-192
1016
1017 Mangus DA, Evans MC, Jacobson A (2003) Poly(A)-binding proteins: multifunctional
1018 scaffolds for the post-transcriptional control of gene expression. *Genome Biol* 4: 223
1019
1020 Maniatis T, Reed R (2002) An extensive network of coupling among gene expression
1021 machines. *Nature* 416: 499-506
1022

1023 Martin KC, Ephrussi A (2009) mRNA localization: gene expression in the spatial dimension.
1024 *Cell* 136: 719-730
1025
1026 Matheny T, Van Treeck B, Huynh TN, Parker R (2021) RNA partitioning into stress granules
1027 is based on the summation of multiple interactions. *RNA* 27: 174-189
1028
1029 McCann C, Holohan EE, Das S, Dervan A, Larkin A, Lee JA, Rodrigues V, Parker R,
1030 Ramaswami M (2011) The Ataxin-2 protein is required for microRNA function and synapse-
1031 specific long-term olfactory habituation. *Proc Natl Acad Sci U S A* 108: E655-662
1032
1033 McMahon AC, Rahman R, Jin H, Shen JL, Fieldsend A, Luo W, Rosbash M (2016) TRIBE:
1034 Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding
1035 Proteins. *Cell* 165: 742-753
1036
1037 Murray DT, Kato M, Lin Y, Thurber KR, Hung I, McKnight SL, Tycko R (2017) Structure of
1038 FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-
1039 Complexity Domains. *Cell* 171: 615-627.e616
1040
1041 Murthy AC, Dignon GL, Kan Y, Zerze GH, Parekh SH, Mittal J, Fawzi NL (2019) Molecular
1042 interactions underlying liquid-liquid phase separation of the FUS low-complexity domain.
1043 *Nature Structural & Molecular Biology* 26: 637-648
1044
1045 Nonhoff U, Ralser M, Welzel F, Piccini I, Balzereit D, Yaspo ML, Lehrach H, Krobisch S
1046 (2007) Ataxin-2 interacts with the DEAD/H-box RNA helicase DDX6 and interferes with P-
1047 bodies and stress granules. *Mol Biol Cell* 18: 1385-1396
1048
1049 Patel A, Lee HO, Jawerth L, Maharana S, Jahnel M, Hein MY, Stoyanov S, Mahamid J, Saha
1050 S, Franzmann TM *et al* (2015) A Liquid-to-Solid Phase Transition of the ALS Protein FUS
1051 Accelerated by Disease Mutation. *Cell* 162: 1066-1077
1052
1053 Preissler S, Ron D (2019) Early Events in the Endoplasmic Reticulum Unfolded Protein
1054 Response. *Cold Spring Harb Perspect Biol* 11
1055
1056 Protter DSW, Parker R (2016) Principles and Properties of Stress Granules. *Trends Cell Biol*
1057 26: 668-679
1058
1059 Protter DSW, Rao BS, Van Treeck B, Lin Y, Mizoue L, Rosen MK, Parker R (2018)
1060 Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule
1061 Assembly. *Cell Rep* 22: 1401-1412
1062
1063 Ramaswami M, Taylor JP, Parker R (2013) Altered ribostasis: RNA-protein granules in
1064 degenerative disorders. *Cell* 154: 727-736
1065
1066 Rayman JB, Karl KA, Kandel ER (2018) TIA-1 Self-Multimerization, Phase Separation, and
1067 Recruitment into Stress Granules Are Dynamically Regulated by Zn²⁺. *Cell Reports* 22: 59-
1068 71
1069
1070 Saito M, Hess D, Eglinger J, Fritsch AW, Kreysing M, Weinert BT, Choudhary C, Matthias P
1071 (2019) Acetylation of intrinsically disordered regions regulates phase separation. *Nat Chem*
1072 *Biol* 15: 51-61

1073
1074 Satterfield TF, Pallanck LJ (2006) Ataxin-2 and its Drosophila homolog, ATX2, physically
1075 assemble with polyribosomes. *Hum Mol Genet* 15: 2523-2532
1076
1077 Schuller AP, Wojtynek M, Mankus D, Tatli M, Kronenberg-Tenga R, Regmi SG, Dip PV,
1078 Lytton-Jean AKR, Brignole EJ, Dasso M *et al* (2021) The cellular environment shapes the
1079 nuclear pore complex architecture. *Nature* 598: 667-671
1080
1081 Scoles DR, Meera P, Schneider MD, Paul S, Dansithong W, Figueroa KP, Hung G, Rigo F,
1082 Bennett CF, Otis TS *et al* (2017) Antisense oligonucleotide therapy for spinocerebellar ataxia
1083 type 2. *Nature* 544: 362-366
1084
1085 Shin Y, Brangwynne CP (2017) Liquid phase condensation in cell physiology and disease.
1086 *Science* 357
1087
1088 Shulman JM, Feany MB (2003) Genetic Modifiers of Tauopathy in Drosophila. *Genetics* 165:
1089 1233-1242
1090
1091 Sidrauski C, Mcgeachy AM, Ingolia NT, Walter P (2015) The small molecule ISRIB reverses
1092 the effects of eIF2 α phosphorylation on translation and stress granule assembly. *eLife* 4
1093
1094 Singh A, Hulsmeier J, Kandi AR, Pothapragada SS, Hillebrand J, Petrauskas A, Agrawal K,
1095 Rt K, Thiagarajan D, Jayaprakashappa D *et al* (2021) Antagonistic roles for Ataxin-2 structured
1096 and disordered domains in RNP condensation. *eLife* 10
1097
1098 Strome S, Wood WB (1982) Immunofluorescence visualization of germ-line-specific
1099 cytoplasmic granules in embryos, larvae, and adults of *Caenorhabditis elegans*. *Proc Natl Acad*
1100 *Sci U S A* 79: 1558-1562
1101
1102 Sudhakaran IP, Hillebrand J, Dervan A, Das S, Holohan EE, Hulsmeier J, Sarov M, Parker R,
1103 VijayRaghavan K, Ramaswami M (2014) FMRP and Ataxin-2 function together in long-term
1104 olfactory habituation and neuronal translational control. *Proc Natl Acad Sci U S A* 111: E99-
1105 E108
1106
1107 Taylor JP, Brown RH, Cleveland DW (2016) Decoding ALS: from genes to mechanism.
1108 *Nature* 539: 197-206
1109
1110 Tillotson R, Selfridge J, Koerner MV, Gadalla KKE, Guy J, De Sousa D, Hector RD, Cobb
1111 SR, Bird A (2017) Radically truncated MeCP2 rescues Rett syndrome-like neurological
1112 defects. *Nature* 550: 398-401
1113
1114 Toretzky JA, Wright PE (2014) Assemblages: functional units formed by cellular phase
1115 separation. *J Cell Biol* 206: 579-588
1116
1117 Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-
1118 Seq. *Bioinformatics* 25: 1105-1111
1119
1120 Van Treeck B, Parker R (2018) Emerging Roles for Intermolecular RNA-RNA Interactions in
1121 RNP Assemblies. *Cell* 174: 791-802
1122

1123 Van Treeck B, Protter DSW, Matheny T, Khong A, Link CD, Parker R (2018) RNA self-
1124 assembly contributes to stress granule formation and defining the stress granule transcriptome.
1125 *Proc Natl Acad Sci U S A* 115: 2734-2739
1126
1127 Vogler TO, Wheeler JR, Nguyen ED, Hughes MP, Britson KA, Lester E, Rao B, Betta ND,
1128 Whitney ON, Ewachiw TE *et al* (2018) TDP-43 and RNA form amyloid-like myo-granules in
1129 regenerating muscle. *Nature* 563: 508-513
1130
1131 Wang F, Li J, Fan S, Jin Z, Huang C (2020) Targeting stress granules: A novel therapeutic
1132 strategy for human diseases. *Pharmacological Research* 161: 105143
1133
1134 Wheeler JR, Matheny T, Jain S, Abrisch R, Parker R (2016) Distinct stages in stress granule
1135 assembly and disassembly. *eLife* 5
1136
1137 Wolozin B, Ivanov P (2019) Stress granules and neurodegeneration. *Nature Reviews*
1138 *Neuroscience* 20: 649-666
1139
1140 Wong YL, Lebon L, Edalji R, Lim HB, Sun C, Sidrauski C (2018) The small molecule ISRIB
1141 rescues the stability and activity of Vanishing White Matter Disease eIF2B mutant complexes.
1142 *eLife* 7
1143
1144 Xie J, Kozlov G, Gehring K (2014) The "tale" of poly(A) binding protein: the MLLE domain
1145 and PAM2-containing proteins. *Biochim Biophys Acta* 1839: 1062-1068
1146
1147 Yang P, Mathieu C, Kolaitis RM, Zhang P, Messing J, Yurtsever U, Yang Z, Wu J, Li Y, Pan
1148 Q *et al* (2020) G3BP1 Is a Tunable Switch that Triggers Phase Separation to Assemble Stress
1149 Granules. *Cell* 181: 325-345.e328
1150
1151 Yi H, Park J, Ha M, Lim J, Chang H, Kim VN (2018) PABP Cooperates with the CCR4-NOT
1152 Complex to Promote mRNA Deadenylation and Block Precocious Decay. *Mol Cell* 70: 1081-
1153 1088.e1085
1154
1155 Yokoshi M, Li Q, Yamamoto M, Okada H, Suzuki Y, Kawahara Y (2014) Direct binding of
1156 Ataxin-2 to distinct elements in 3' UTRs promotes mRNA stability and protein expression. *Mol*
1157 *Cell* 55: 186-198
1158
1159 Yoshida M, Yoshida K, Kozlov G, Lim NS, De Crescenzo G, Pang Z, Berlanga JJ, Kahvejian
1160 A, Gehring K, Wing SS *et al* (2006) Poly(A) binding protein (PABP) homeostasis is mediated
1161 by the stability of its inhibitor, Paip2. *The EMBO Journal* 25: 1934-1944
1162
1163 Youn JY, Dyakov BJA, Zhang J, Knight JDR, Vernon RM, Forman-Kay JD, Gingras AC
1164 (2019) Properties of Stress Granule and P-Body Proteomes. *Mol Cell* 76: 286-294
1165
1166 Zhang K, Daigle JG, Cunningham KM, Coyne AN, Ruan K, Grima JC, Bowen KE, Wadhwa
1167 H, Yang P, Rigo F *et al* (2018) Stress Granule Assembly Disrupts Nucleocytoplasmic
1168 Transport. *Cell* 173: 958-971.e917
1169
1170 Zhang Y, Ling J, Yuan C, Dubruille R, Emery P (2013) A role for Drosophila ATX2 in
1171 activation of PER translation and circadian behavior. *Science* 340: 879-882
1172

1173 Zyryanova AF, Kashiwagi K, Rato C, Harding HP, Crespillo-Casado A, Perera LA, Sakamoto
1174 A, Nishimoto M, Yonemochi M, Shirouzu M *et al* (2021) ISRIB Blunts the Integrated Stress
1175 Response by Allosterically Antagonising the Inhibitory Effect of Phosphorylated eIF2 on
1176 eIF2B. *Mol Cell* 81: 88-103.e106
1177