Purification of the neurodegenerative disease associated protein TDP-43 and development of TDP-43 aggregation inhibitors

PhD Thesis

Toby John Baker Marshall BSc (Hons)

Lancaster University

September 2023

I, Toby Marshall, confirm that the work presented in this thesis is my own and has not been submitted in substantially the same form for the award of a higher degree elsewhere. Where information has been derived from other sources, I confirm this has been indicated in the thesis.

Submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy.

Abstract

Transactive response DNA-binding protein-43 (TDP-43) is a protein that has been implicated in multiple neurodegenerative diseases, including amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD). In these diseases, TDP-43 is found aggregated in the cytoplasm of neurones in the brain and spinal cord and it is hypothesised that this aggregation leads to neuronal degeneration. Despite identification of TDP-43 as a constituent of pathological aggregates in 2006, progress in biochemical characterisation of TDP-43 and its aggregation has been limited by an inability to purify sufficient soluble protein to allow characterisation. In this study, a novel method for the purification of TDP-43 has been developed. The resulting purified protein exists in multiple oligomeric states depending on buffer conditions, displays evidence of secondary structural content by circular dichroism spectroscopy and in preliminary studies demonstrates DNA binding activity. A TDP-43 C-terminal fragment was also purified and a fluorescence-based assay developed to monitor its aggregation, with transmission electron microscopy (TEM) used to image the aggregates produced. In this assay, small molecules were tested as aggregation inhibitors. Following minimal success re-purposing generic aggregation inhibitor molecules, a series of targeted peptide-based inhibitors were designed. The third-generation peptide inhibitors, designed with the aid of the artificial intelligence system AlphaFold, reduced the aggregation of the C-terminal fragment, with TEM identifying changes to the morphology of the aggregates produced. Finally, a “druggable” *Saccharomyces cerevisiae* yeast cell model of TDP-43 proteinopathy was developed, in which molecules with potential as TDP-43 aggregation inhibitors can be tested further.

Acknowledgements

I would like to thank Dr Fiona Benson for her time, guidance and support throughout this PhD project. Thank you also to Dr Edward Parkin for his support and feedback on the project as it progressed. Thank you to the late Professor David Allsop, for his encouragement and support during the initial stages of my PhD and providing me with the opportunity to carry out this work. Thank you also to the late Dr Nigel Fullwood, for his guidance and who along with Dr Elisabeth Shaw took the time to train me in electron microscopy. Thank you also to Dr Mehul Makwana for his advice and help. Thanks also go to the Sir John Fisher Foundation, for funding this work.

Finally, I am very grateful for all the love and support from my family, not just during this PhD, but throughout my time in education. To my mother and father, my sister Emma, Nanny, Pappy, Grandma and Grandad, thank you.

Contents

List of Figures

List of Tables

List of Abbreviations

- AD - Alzheimer’s disease
- AIM4 - [4,5-bis{(N-carboxy methyl imidazolium)methyl}acridine] dibromide
- ALS - amyotrophic lateral sclerosis
- Apo-AII - apolipoprotein A-II
- APS - ammonium persulfate
- AU - Arbitrary units
- Aβ - amyloid beta
- BCA - bicinchoninic acid
- Bid - BH3-interacting domain death agonist
- C9ORF72 - Chromosome 9 Open Reading Frame 72
- CAB - carbonic anhydrase B
- CBP - Creb binding protein
- CD - circular dichroism
- CFTR - cystic fibrosis transmembrane conductance regulator
- *Chop/Ddit3* - CCAAT/enhancer-binding homologous protein
- CHT - ceramic hydroxyapatite
- CNS - central nervous system
- CTF - C-terminal fragment
- DAPI - 4′,6-diamidino-2-phenylindole
- DDR - DNA damage response
- DMSO - dimethylsulfoxide
- DSB - double strand break
- DTT - dithiothreitol
- ECL - enhanced chemiluminescence
- EDTA - Ethylenediaminetetraacetic acid

- EMSA - electrophoretic mobility shift assay
- ER - endoplasmic reticulum
- fALS - familial amyotrophic lateral sclerosis
- FTLD - Frontotemporal Lobar degeneration
- FUS - fused in sarcoma
- G3BP1 - Ras GTPase-activating protein binding-protein 1
- GF - Gel Filtration
- GFP - green fluorescent protein
- hiAPP - human islet amyloid polypeptide
- hRNP - heteronuclear ribonucleotide binding protein
- Hsp110 - heat shock protein 110
- Hsp70 - heat shock protein 70
- HSP8A - heat shock protein 8a
- IDA - Iminodiacetic acid
- IMAC - immobilised metal affinity chromatography
- IP - Immunoprecipitation
- IPTG - isopropylthio-β-galactoside
- LAMP-2A - lysosomal associated membrane protein 2a
- LARKS - Low-complexity Aromatic-Rich Kinked Segments
- LATE - limbic-predominant age-related TDP-43 encephalopathy
- LB - Luria-Bertani
- LCD - low-complexity domain
- LDAO - lauryldimethylamine oxide
- LLPS - liquid-liquid phase separation property
- MBP - Maltose binding protein
- MPD - 2-Methyl-2,4-pentanediol (hexylene glycol)
- mTORC1 - mammalian target of rapamycin complex 1

- NHEJ - non-homologous end joining
- NLS - nuclear localisation signal
- NMR - Nuclear magnetic resonance
- NTD – N-terminal domain
- NusA - N-utilization substance A
- OG - Octyl-β-D-Glucopyranoside
- OPTN - optineurin
- P62/SQSTM1 - Sequestosome 1
- PBS - phosphate buffered saline
- PD - Parkinson’s disease
- PEG 3350 - Polyethylene glycol 3350
- PEG-PE - polyethylene glycol-phosphatidylethanolamine
- PelB - pectate lyase B leader sequence
- PROTACs - Proteolysis targeting chimeras
- PrP - prion protein
- PSF - PTB-associated splicing factor
- PVDF - polyvinylidene fluoride
- QBP1 - Polyglutamine binding peptide 1
- RAN translation - repeat associated non-ATG translation
- RBM14 - RNA binding motif 14
- RFU - Relative fluorescence units
- RNP-1 - ribonucleoprotein motif 1
- RNP-2 - ribonucleoprotein motif 2
- RNS - reactive nitrogen species
- ROS - reactive oxygen species
- RRM - RNA recognition motif
- sALS - sporadic amyotrophic lateral sclerosis

- SC-U - synthetic complete without uracil
- SDS – sodium dodecyl sulphate
- SDS-PAGE - sodium dodecyl sulphate polyacrylamide gel electrophoresis
- SEC - size exclusion chromatography
- SOC - Super Optimal broth with Catabolite repression
- SOD1 - superoxide dismutase 1
- SUMO - Small Ubiquitin-like Modifier
- TDP-25 - 25kDa fragment of TDP-43
- TDP-43 - Transactive response DNA-binding protein-43
- TEM - transmission electron microscopy
- TEMED - (N,N,N′,N′-Tetramethylethylenediamine)
- ThT - Thioflavin T
- UBQLN2 - Ubiquilin-2
- UPS - ubiquitin proteasome system
- UTR - untranslated region
- VCP - Valosin-containing protein
- YFP - yellow fluorescent protein
- YPD - yeast extract peptone dextrose
- Δerg6 - *ERG6* gene deletion mutant
- ΔNLS TDP-43 - TDP-43 with a nuclear localisation signal deletion

1. Literature Review

1.1 Neurodegeneration

A 2019 report by the United Nations revealed that by 2050, 1 in 6 people of the world's population will be over 65, an increase from 1 in 11 in 2019 (Un.org, 2019). However, with increasing life spans, and an ageing population, comes an increasing prevalence of diseases associated with ageing, such as neurodegeneration. Neurodegenerative diseases cause progressive loss of brain functions, and so progressive clinical symptoms. These can be cognitive, as in Alzheimer's disease, or affect motor function as in Parkinson's disease and amyotrophic lateral sclerosis (ALS, commonly known as motor neurone disease), or a combination, depending on the brain region affected by the neurodegenerative pathology. Despite different clinical symptoms, a broad range of neurodegenerative diseases share a common pathology- toxic aggregation of misfolded or mutant proteins. This is seen in Alzheimer's disease as extracellular plaques of amyloid-β and intracellular neurofibrillary tangles (NFTs) of tau protein, in Parkinson's disease as Lewy bodies consisting of α-synuclein, and in ALS and some forms of Frontotemporal Lobar Degeneration (FTLD) as aggregates of TDP-43 or FUS proteins (Taylor *et al*., 2002; Baralle, *et al*., 2013).

This review and the following study will examine and investigate transactive response DNA-binding protein 43 (TDP-43), with a particular focus on its aggregation. Investigation will be made into inhibition of aggregation by small molecules, with a view of this being a therapeutic approach in neurodegenerative disease.

1.2 Pathways to protein aggregation

Proteins may aggregate due to alterations to common neuronal pathways. Fig. 1.1 summarises several neuronal pathways that may become defective or be altered in neurodegenerative diseases. In the protein quality control system, involving the ubiquitin proteasome system (UPS), chains of ubiquitin are conjugated to damaged, misfolded or surplus proteins, targeting them to the proteasome for degradation (Ross *et al*., 2015). Chaperones such as heat shock proteins (e.g. the Hsp70 family) are involved in targeting damaged proteins to the proteasome (Lackie *et al*., 2017), while disaggregase activity of Hsp70, requiring a J-domain protein and Hsp110 family proteins can disaggregate insoluble aggregates and reactivate inactive proteins (Doyle *et al*., 2013). The autophagy-lysosome pathway is another mechanism of proteostasis, which uses autophagy to target cellular contents, organelles and protein aggregates for degradation in lysosomes. There are three major pathways in this system. The first is macroautophagy, where adaptor proteins such as optineurin and p62/SQSTM1 (Sequestosome 1) target specific cargo for engulfment by autophagosomes, which then fuse with the lysosome, depositing their contents in the lysosomal lumen (Finkbeiner, 2019). The second is microautophagy, where invaginations of the lysosomal membrane occur around substrates to be degraded, and form vesicles which pinch off inside the lysosome. Finally, chaperone mediated autophagy targets polypeptides with a KFERQ motif, and utilises chaperones such as LAMP-2A and HSP8A for translocation of the polypeptides into the lysosome (Finkbeiner, 2019). Degradation occurs in the lysosome by way of its acidified lumen containing approximately 60 soluble hydrolases (Finkbeiner, 2019).

If these pathways become altered or defective, it would allow aggregates of misfolded proteins to build up in neurons and not be removed. Protein ubiquitination, the actions of some chaperones and the proteasome all require ATP (Hershko *et al.,* 1983; Ross *et al.,* 2015), and so reduced mitochondrial function caused by the action of toxic

protein aggregates can lead to impairments in the UPS due to lack of ATP. Decreased action of the UPS can in turn affect the mitochondria, in that abnormal mitochondrial proteins, along with misfolded aggregation prone proteins, are not degraded by the proteasome and cause further mitochondrial degradation (Ross *et al*., 2015).

Cellular stresses and proteotoxic stress can lead to the stalling of mRNA translation, and these stalled mRNAs can be incorporated into stress granules as a cellular defence mechanism (Hans *et al*., 2020). Stress granules may contain these mRNAs, mRNA-associated translation initiation complexes and RNA binding proteins, such as aggregation prone proteins involved in neurodegenerative disease (Hans *et al*., 2020). If these proteins are held in stress granules for a continued length of time, it would allow them to aggregate into larger cellular inclusions (Hans *et al*., 2020).

For some proteins involved in neurodegeneration (TDP-43, tau, α-synuclein, amyloid-β and prion protein, PrP) there is considerable evidence for the spreading of protein aggregates between neurons, seeding further downstream aggregation (Clavaguera *et al*., 2009; Jucker and Walker, 2013).

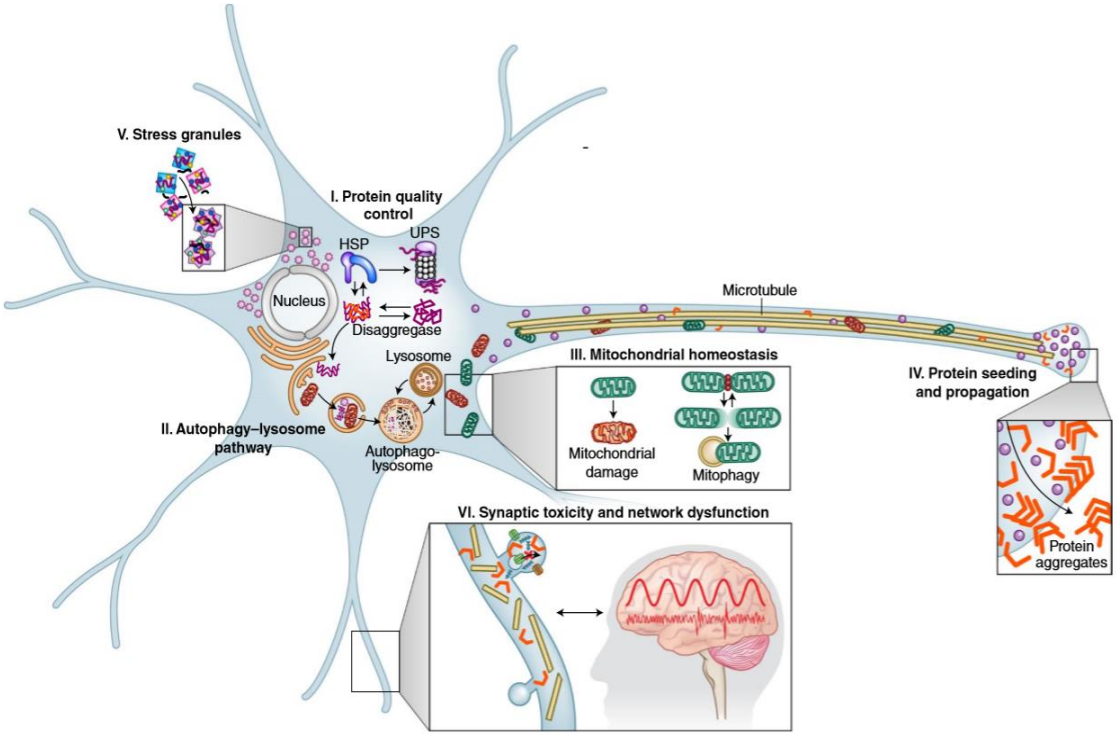


Figure 1.1. Neuronal pathways that can be altered in neurodegenerative disease. I. The protein quality control system, including the ubiquitin-proteasome system (UPS), chaperones such as heat shock proteins (HSPs) and disaggregase. **II**. The autophagy-lysosome system, involved in the compartmentalisation and degradation of cellular components. **III**. Mitochondrial homeostasis and mitochondrial quality control via mitophagy, leading to reduced energy production. **IV**. Protein seeding and propagation, in which misfolded proteins and aggregates may propagate between neurons. **V**. Changes to stress granule dynamics and components may favour aggregation. **VI**. The presence of aggregates at synapses may impair synaptic networks. Figure from Gan et al., (2018).

1.3 Amyotrophic Lateral Sclerosis (ALS), Frontotemporal Lobar Degeneration (FTLD) and Limbic-predominant age-related TDP-43 encephalopathy (LATE)

First described in 1869 by Jean-Martin Charcot, amyotrophic lateral sclerosis (ALS), also known as Lou Gehrig's disease and motor neurone disease, is a degenerative disease of the motor neurons (Corcia and Meininger, 2019). With a prevalence of ~2 per 100,000 individuals, it is the third most common neurodegenerative disease behind Alzheimer's disease and Parkinson's disease. Onset of ALS usually occurs between 50 and 65 years of age, and the disease progresses via degeneration of the upper motor neurons, particularly in the motor cortex, brain stem and spinal cord, and lower motor neurons that extend to the muscles from the spinal cord. Muscle wasting occurs due to lack of stimulation as the motor neurons degenerate in the spinal cord, causing progressive paralysis, which usually becomes fatal between 2 and 5 years after symptom onset (Kiernan *et al*., 2011; Zarei *et al*., 2015). There is currently no cure and there only two approved treatments which moderately extend lifespan (~6 months); riluzole, a glutamatergic neurotransmitter inhibitor, and edaravone, an antioxidant drug (Jaiswal, 2018).

Frontotemporal Lobar Degeneration (FTLD) is the second most prevalent form of neurodegenerative dementia, behind Alzheimer's disease. It includes several behavioural, personality, language and cognitive disorders, associated with degeneration of the frontal and temporal lobes. There are three main variants of FTLD- the behavioural variant (the most common presentation), progressive non-fluent aphasia and semantic dementia (Bahia *et al*., 2013).

Limbic-predominant age-related TDP-43 encephalopathy (LATE) is a recently characterised type of dementia which develops in the very elderly, leading to progressive cognitive impairment. It is similar in its clinical features to Alzheimer's

disease but with a more gradual cognitive decline. Its features are different to the behavioural changes that are characteristic of FTLD, with the limbic structures being affected by the proteinopathy (Nelson *et al*., 2019; Zhang *et al*., 2020).

A common pathology in these three neurodegenerative diseases, and to an extent in Alzheimer's disease and Parkinson's disease, is the presence of aggregates containing Transactive response DNA-binding protein-43 (TDP-43) in the cytoplasm of neurons (Neumann *et al*., 2006; Nakashima-Yasuda *et al.*, 2007; Meneses *et al.*, 2021). TDP-43 was identified as a primary constituent of ubiquitinated inclusions in cortical neurons in FTLD and motor neurons in ALS by Neumann *et al*., (2006), and these inclusions have since been shown to be common to 97% of ALS cases (including both familial and sporadic cases) (Scotter *et al*., 2015).

It should be noted however that TDP-43 is not the sole proteinopathy in these diseases. The first proteinopathy to be associated with ALS was that of superoxide dismutase 1 (SOD1), an enzyme involved in the antioxidant defence mechanism. Pathogenic mutations in the *SOD1* gene appear to be responsible for approximately 15% of familial ALS but only 1.2% of sporadic ALS cases (Zou *et al*., 2017), potentially by a toxic gain of function of the misfolding and aggregating enzyme (Mejzini *et al*., 2019). Another protein, fused in sarcoma (FUS), has been identified as a component of aggregates in both ALS and FTLD cases. FUS is a DNA and RNA binding protein, involved in transcription regulation, RNA splicing and gene expression (Vance *et al*., 2009; Neumann *et al*., 2009) and its aggregation may cause loss of function due to sequestration of the protein or toxic gain of function of aggregates (Nolan *et al*., 2016). FTLD can also be caused by tauopathy- accumulation of tau is found in around 45% of FTLD cases (Boxer *et al*., 2012). FTLD-tau cases can be classified into several subtypes, including Pick's disease, corticobasal degeneration and progressive supranuclear palsy (reviewed in Bahia *et al*., 2013).

1.4 Genetics of ALS and FTLD

The majority of ALS cases are sporadic (sALS) and occur in people without a family history of the disease. Up to 10% of cases of ALS are familial (fALS) and occur in patients who have at least one other affected family member, but these cases are clinically indistinguishable from sALS cases (Kirby *et al*., 2016). Similarly, FTLD-TDP may also be sporadic or familial. Genetic causes are known for approximately 11% of sALS cases and 68% fALS cases (Scotter *et al*., 2015), with the most characterised mutations occurring in *SOD1, FUS, TARDBP* (encoding TDP-43) and Chromosome 9 Open Reading Frame 72 (*C9ORF72*) (Mejzini *et al*., 2019) (Fig. 1.2). Given the overlap in genetic mutations between ALS and FTLD (Fig. 1.2), the shared TDP-43 proteinopathy in these diseases is perhaps not surprising. In some cases, a patient may develop symptoms of both ALS and FTLD, and so these diseases are thought to be on a spectrum (Fig. 1.3) depending on the area of the brain affected.

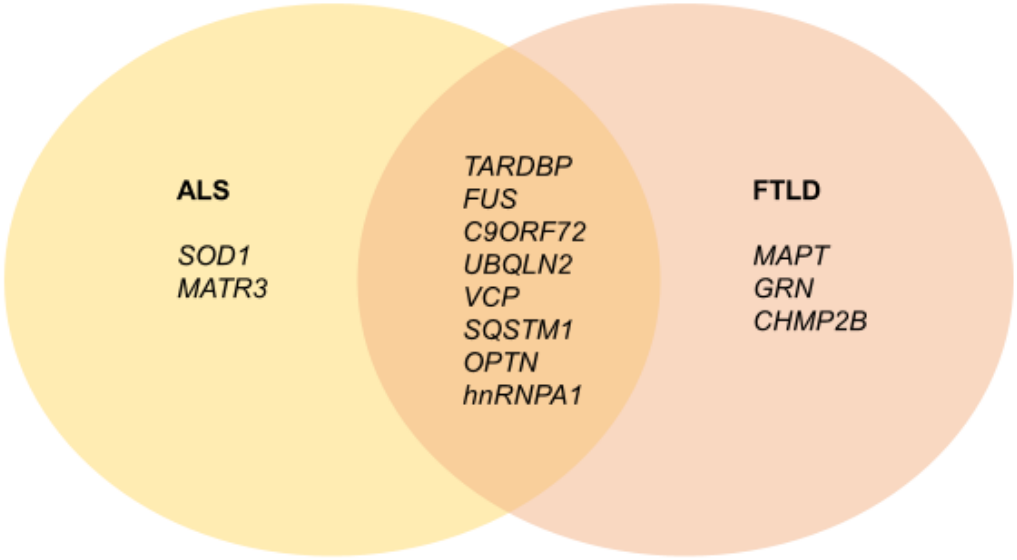


Figure 1.2. Mutated genes identified in ALS and/or FTLD. (Mejzini et al., 2019), (Deng et al., 2011), (Fecto et al., 2011), (Watts et al., 2004), (Maruyama et al., 2010), (Johnson et al., 2014), (Kim et al., 2013), (Baker et al., 2006), (Skibinski et al., 2005), (Rubino et al., 2012) and (Pottier et al., 2015).

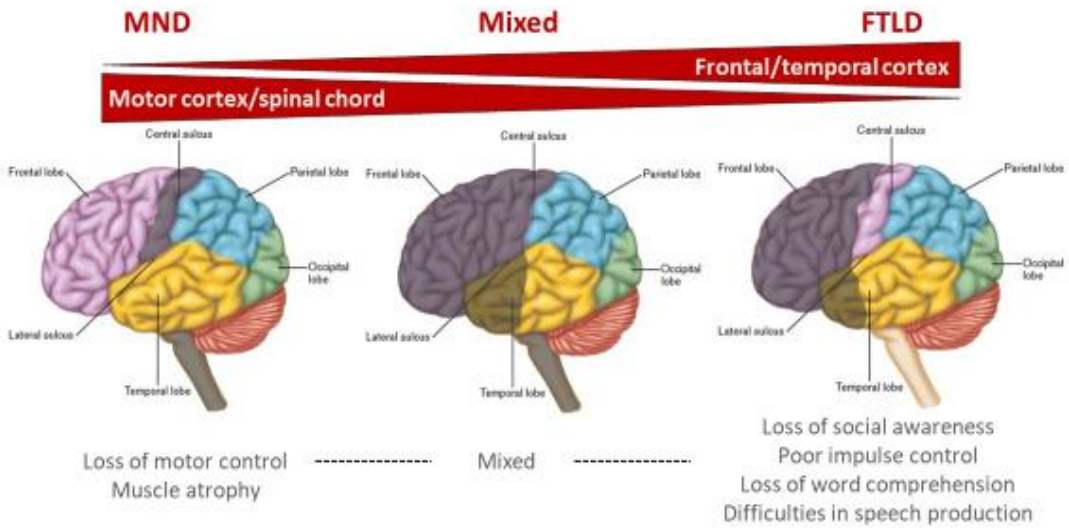


Figure 1.3. The spectrum of disease from ALS (MND-motor neuron disease) to FTLD. Affected areas of the brain are shaded. Provided by Professor David Allsop.

The most common cause of fALS and familial FTLD-TDP is a hexanucleotide repeat expansion (GGGGCC) in the non-coding promoter region of the C9ORF72 gene. Discovered by DeJesus-Hernandez *et al*. (2011), this expansion accounts for ~40% of fALS cases and ~18% of familial FTLD cases (Trageser *et al*., 2019). The gene product of C9ORF72 is a guanine nucleotide exchange factor, implicated in lysosome biogenesis, vesicular trafficking, autophagy and mTORC1 signalling (Iyer *et al*., 2018). Normal individuals carry between 2 and 30 GGGGCC repeats, whereas ALS/FTLD individuals with the C9ORF72 expansion show from 66 to over 4400 repeats (Trageser *et al*., 2019).

By generation of C9ORF72 knockout mice, Koppers *et al*. (2015), showed that loss of C9ORF72 function does not induce motor degeneration, suggesting that the repeat expansion leads to gain of function. The repeat expansion may have several effects on the neurons. The G-rich repeat expansion may lead to the formation of structural polymorphisms such as G-quadruplexes in both the DNA and transcribed RNA, that may lead to abortive transcription. The aborted transcripts and the structural polymorphisms themselves may bind and sequester ribonucleoproteins and RNA binding proteins, leading to pathogenic effects (Fay *et al*., 2017; Trageser *et al*., 2019). Another potentially toxic effect of the repeat expansion is the production of dipeptide repeat proteins: poly-(Gly-Ala), poly-(Gly-Pro) and poly-(Gly-Arg). These are generated from the expanded GGGGCC repeat in all three reading frames via repeat associated non-ATG (RAN) translation, and may themselves be toxic, or may form inclusions which TDP-43 can then accumulate around (Mackenzie *et al*., 2013; Cleary *et al*., 2018). Despite the other potential effects of the C9ORF72 repeat expansion, TDP-43 proteinopathy is still observed in ALS/FTLD C9ORF72 cases.

1.5 Introduction to TDP-43

Transactive response DNA binding protein 43 (TDP-43) was first identified by Ou *et al*. (1995) as a cellular modulator of HIV-1 gene expression. It is a 43 kDa ubiquitously expressed protein, encoded by the *TARDBP* gene, located at locus 1p36.22, and is a member of the heteronuclear ribonucleotide binding protein family (hRNPs) (François-Moutal *et al*., 2019). Highly conserved throughout different species, it appears essential for CNS development from embryonic stages to adulthood. Huang *et al*. (2010) detected the ubiquitous expression of TDP-43 in postnatal rodents, however in adult rodents the protein became undetectable in the periphery but expression remained at high levels in the CNS.

1.5.1 TDP-43 structure

Like other RNA binding proteins, TDP-43 has a beads-on-a-string structure, where distinct functional domains are separated by short linking sequences (Fig 1.4). It is constructed of four distinct regions; the N terminal domain (NTD), two RNA recognition motifs (RRM1 and RRM2) and a highly disordered C terminal domain (CTD) (Fig. 1.4 (a)). Despite 3-dimensional structures of the NTD, the RNA-binding domains and regions of the CTD becoming available in recent years, structural determination of the entire protein in its monomeric form has been a challenge due to its intrinsically disordered nature (François-Moutal *et al*., 2019), and difficulties in purifying recombinant protein. Wright *et al*. (2020), finally produced a potential whole structure of a TDP-43 recombinant protein (Fig. 1.4 (b)), where all tryptophan residues of the protein were mutated to alanine to aid in solubility in the purification and structural determination process.

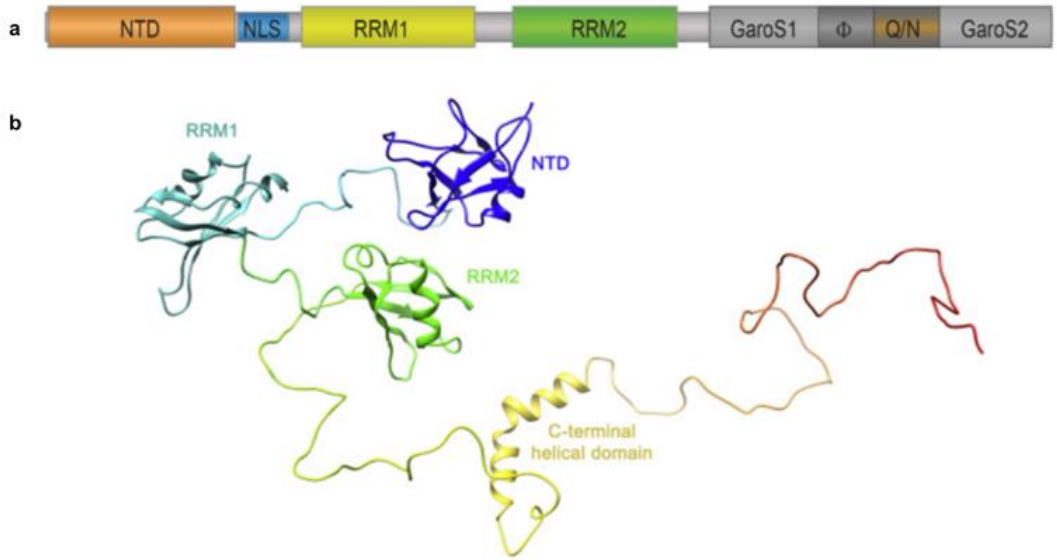


Figure 1.4. TDP-43 structure (a) An overview of the structure of TDP-43. NTD, N-Terminal Domain; NLS, Nuclear Localisation Signal; RRM, RNA Recognition Motifs. GaroS1 (Gly-aromatic-Ser 1), φ (hydrophobic region), Q/N (glutamine/asparagine rich region), and GaroS2 (Gly-aromatic-Ser 2) are the constituents of the CTD (C-Terminal Domain). Adapted from François-Moutal *et al*., (2019). **(b)** Full structure of TDP-43$_{WtoA}$ (Wright *et al*., 2020).

The N terminal domain of TDP-43 consists of residues 1-77, and its monomeric structure was first identified by Mompeán *et al*. (2016) (Fig. 1.5). The structure consists of six β-strands and one α-helix arranged into a ubiquitin-like fold; a fold similar to the reported structure of axin-1. F35, Y43, W68, Y73 and 17 aliphatic side chains make up a hydrophobic core, while the β-strands produce two β-sheets, with strands β1, β2, β3 and β6 producing one sheet and β4 and β5 the other (Mompeán *et al*., 2016). The NTD has been shown to be involved in reversible oligomerisation that is required for the protein's normal roles, and in the formation of nuclear oligomers that are distinct from pathological aggregates (Afroz *et al*., 2017). The NTD may also display structural plasticity affected by its environment. Moretti *et al*. (2022) demonstrated an alternative, dynamic conformation of purified NTD, which was a stable monomer with a lower

aggregation propensity. This conformation, formed in the presence of the detergent Sulfobetaine 3-10, contained a higher alpha helical content at the expense of reduced beta sheet structure, and transition between the native state and the alternative state was reversible. Recently, purified recombinant NTD (1-102) has been shown to exist in an equilibrium between a dimer and a tetramer, which can be shifted with the presence and absence of reducing agents such as dithiothreitol (DTT) at both low (5.5 μM) and high (55 μM) protein concentrations (Herrera *et al*., 2023). The authors also used steady state anisotropy analysis to identify a dissociation constant of 2.4 μM for oligomerisation, and by Stern-Volmer analysis using acrylamide quenching of Trp68 fluorescence that at 1 μM the NTD may behave as a monomer. Taken together, these results suggest that the formation of monomers, dimers, and oligomers may be concentration dependent.

Between the NTD and RRM1 is a linking region containing the Nuclear Localisation Signal (NLS) (residues 78-100). TDP-43 is actively transported into the nucleus to carry out its nuclear roles by the classical nuclear import pathway via importin-α (Nishimura *et al*., 2010). Disruption to this process may lead to accumulation of the protein in the cytoplasm where pathological aggregation could take place. Nuclear export of TDP-43 appears to occur by passive diffusion, and retention of TDP-43 in the nucleus by binding to GU-rich nuclear RNAs (Duan *et al*., 2022). In fact, TDP-43 nuclear efflux to the cytoplasm can be induced in HeLa cells by RNAase treatment and the provision of short GU rich oligomers, suggesting that disruption of TDP-43 RNA binding in the nucleus results in its nuclear efflux to the cytoplasm where it can aggregate (Duan *et al*., 2022). Herrera *et al*. (2023) demonstrated that the tertiary structure of the folded region of the NTD is influenced by the presence of the NLS in a pH dependent manner utilising the fluorescence of Trp68 as an indicator of difference in solvent exposure. At acidic pH (3.9), in the presence of the NLS, the tryptophan residue is more solvent

exposed than in a mildly alkaline (pH 7.5) environment indicating an alteration in structure, however this difference is not observed in the absence of the NLS.

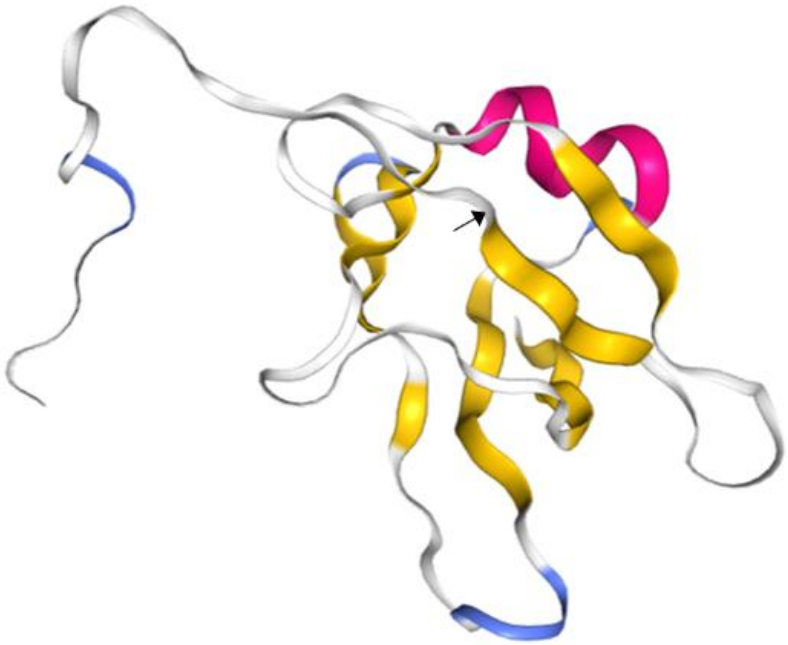


Figure 1.5. The N terminal domain of TDP-43 (His tagged). The arrow indicates the start point of the native protein. The α helix is displayed in pink, the β sheets in yellow (Mompeán *et al.*, 2016), PDB accession number 2N4P.

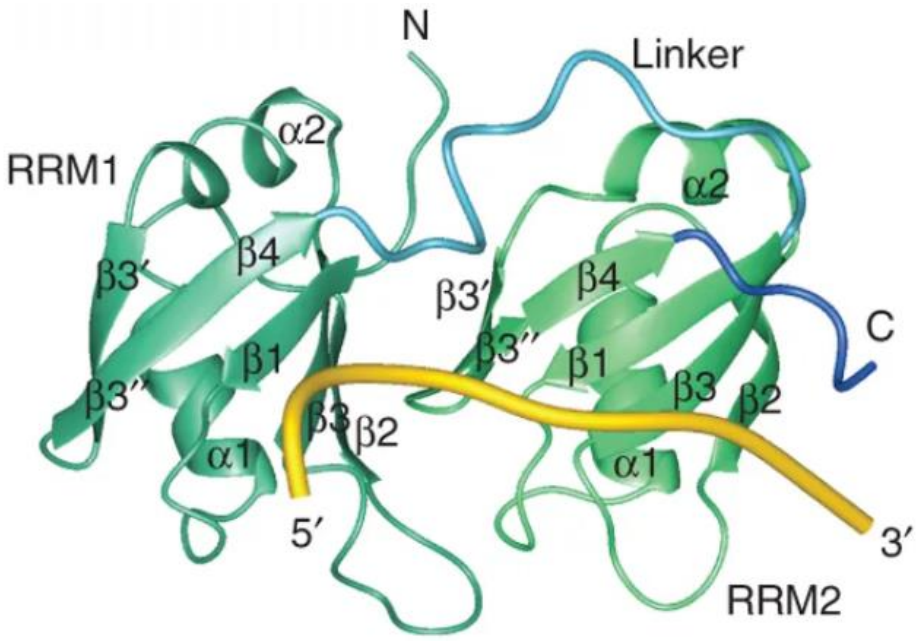


Figure 1.6. Ribbon representation of the TDP-43 RRM1 and RRM2 domains bound to a short RNA strand. The RNA strand is shown in gold, the RRM domains in green and the linker regions in blue (Lukavsky et al., 2013).

RRM1 spans amino acids 106-177, while RRM2 spans residues 192-259 and both of these contain highly conserved sequence motifs required for nucleic acid binding. These motifs, known as RNP-1 and RNP-2, consist of the sequences KGFGFVRF and LIVLGL in RRM1, and RAFAFVTF and VFVGRC in RRM2 (François-Moutal *et al*., 2019). As shown in the structure produced by Lukavsky *et al*. (2013) (Fig. 1.6), both RRM1 and RRM2 consist of a five stranded β-sheet and two α-helices. RRM1 and RRM2 are connected by a flexible linker (amino acids 178-191) and the two domains bind RNA in a 5' to 3' direction. The RRM domains allow TDP-43 to bind nucleic acids, with especially high affinity for TG/UG rich sequences and TG/UG repeats. Kuo *et al*., (2009), demonstrated that TDP-43 could bind both ssDNA and dsDNA, and that both RRM1 and RRM2 can bind DNA. They also demonstrated the ability of both RRM1 and RRM2 to bind UG rich RNA. The authors demonstrated that RRM1 bound with higher affinity to RNAs with higher UG repeat numbers than lower repeat numbers, whereas

RRM2 bound lower UG repeat RNAs with higher affinity. Also, TDP-43 constructs containing both RRMs showed higher affinity to UG repeat RNAs than the RRM constructs individually, suggesting the two RRMs may work in tandem.

The C terminal domain of TDP-43 (CTD) consists of amino acids 260-414, and as shown in Fig. 1.4 contains two Gly-aromatic-Ser regions (GaroS1 and GaroS2), a glutamine/asparagine rich region (Q/N) and a hydrophobic region (φ) (Fig 1.4). All structures of the CTD alone produced so far are of CTD fragments (such as the 311-360 amino acid fragment shown in Fig. 1.7), due to the aggregation prone and naturally disordered nature of the domain (as seen in Fig. 1.4 (b)). However, some structural elements have been identified within the GaroS, Q/N and hydrophobic regions. Jiang *et al*. (2016), demonstrated that the hydrophobic region forms a helix-turn-helix structure (residues 321-343), which if disrupted by mutation may form a hairpin-like β-sheet structure leading to aggregate formation. The Q/N rich region appears to form a β-rich structure, containing β-hairpins (Mompeán *et al*., 2014). It also forms steric-zipper β-sheet structures characteristic of aggregatory proteins, and kinked β-sheets known as LARKS (Low-complexity Aromatic-Rich Kinked Segments) (Guenther *et al*., 2018) which together may lead to aggregation by interaction with these structures in neighbouring TDP-43 molecules (Guenther *et al*., 2018). The CTD is involved in the control of protein-protein interactions due to the ability of the naturally unstructured CTD to form transient β-sheet rich structures. The CTD also allows the protein to oligomerize into stress granules by way of its liquid-liquid phase separation property (LLPS). Three tryptophan residues appear critical for LLPS (residues 334, 385 and 412), and ALS related mutations to these residues alter LLPS properties, which may affect aggregation during stress granule formation (Li *et al*., 2018).

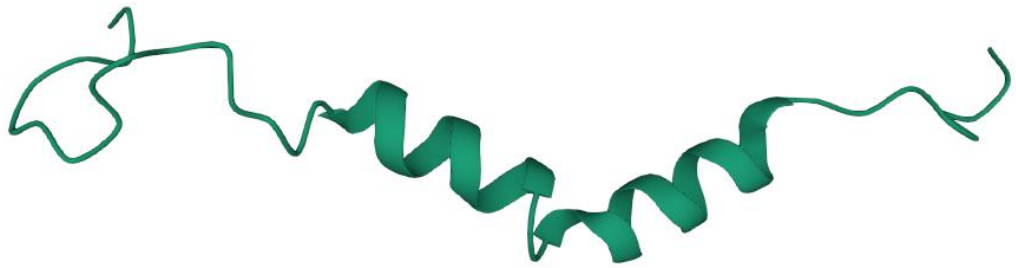


Figure 1.7. A fragment (amino acids 311-360) of the CTD of TDP-43. PDB ID 2N3X, deposited by Jiang et al. (2016).

1.5.2 Normal roles of TDP-43

As previously mentioned, the roles of TDP-43 involve interactions with both nucleic acids and other proteins. Like other hnRNPs, TDP-43 has roles in gene expression (Ou *et al*., 1995; Morera *et al*., 2019), pre-mRNA splicing (Buratti and Baralle, 2010; Ling *et al*., 2015; Donde *et al*., 2019), mRNA autoregulation (Ayala *et al*. 2010); and miRNA regulation and Drosha interactions (Buratti *et al.,* 2010). It also appears to have a role in non-homologous end joining repair of double stranded DNA breaks (Mitra and Hegde, 2019; Mitra *et al*., 2019). In the cytoplasm it appears to co-localise with subcellular compartments (Gao *et al*., 2019).

TDP-43 was first identified as a cellular modulator of HIV-1 gene expression by Ou *et al*., (1995). They hypothesised that the binding of TDP-43 to the TAR DNA element repressed the recruitment of transcription factors, thus repressing gene expression. TDP-43 has also been shown to mediate transcriptional repression of the mouse SP-10 gene, by binding with the SP-10 insulator, preventing enhancer-promoter interactions (Abhyankar *et al*., 2007). Knockdown of TDP-43 by siRNA and mutation of TDP-43 binding sites reduced and abolished the enhancer blocking effects respectively. A genome-wide study by Morera *et al*., (2019), showed that siRNA

knockdown of TDP-43 in HEK cells led to an at least 2-fold reduction in transcription of 2502 protein-coding genes (of 7332 genes studied), and a similar magnitude of increase in transcription of 238 protein coding genes. TDP-43 therefore appears to have roles in both promotion and repression of transcription.

TDP-43 also has roles in pre-mRNA splicing and has been characterised as a negative splicing regulator by recognition of exons in the *CFTR* and *Apo AII* genes via binding of UG-rich elements, and the binding of other hnRNPs via its C-terminal domain (Buratti and Baralle, 2010). In fact, splicing repression has been shown to be a major function of TDP-43 in motor neurons (Donde *et al*., 2019). Expression of a chimeric splicing repressor consisting of the TDP-43 RRM domains fused to an unrelated splicing repressor $RAVER_1$, in Drosophila and mouse TDP-43 knockout models led to rescued splicing repression of TDP-43 targets, slowing the progression of motor deficits and amelioration of motor neuron loss (Donde *et al*., 2019). This repression may be both at normal splice sites and at cryptic sites; TDP-43 knockout leads to cryptic exons being spliced into mRNAs, causing disrupted translation and promoting mRNA decay (Ling *et al*., 2015). TDP-43 is therefore vital for correct splicing and the relevance of this to ALS is demonstrated by the fact that the protein is involved, via its RRM domains and its N-terminus, in the repression of cryptic exon inclusion in *UNC13A*, an identified ALS/FTLD genetic risk factor whose protein product has a role in neurotransmitter release at the synapse (Koike et al. 2023). Another important recently identified role in splicing is in relation to *Zmynd11,* mutations of which have been identified in autism with motor delays, ataxia and intellectual disability (Narayanan *et al*., 2023). Transgenic mice overexpressing A315T mutant TDP-43 were found to have aberrant splicing of *Zmynd11* in the brain and spinal cord, with these changes occurring before motor symptom onset (Narayanan *et al*., 2023).

This splicing activity contributes to autoregulation, both of other proteins and of TDP-43 itself. For example, the SC-35 protein, itself a splicing factor, regulates its own

expression by stimulating unproductive splicing events in the 3' untranslated region (UTR) of its own pre-mRNA. TDP-43 acts to antagonize the action of SC-35, thus contributing to its autoregulation (Dreumont *et al*., 2009). TDP-43 appears to autoregulate its mRNA levels through a similar mechanism. Ayala *et al*. (2010), showed that induced TDP-43 expression in HEK293 cells led to downregulation of endogenous TDP-43 protein levels by binding to 3' UTRs of its pre-mRNA and promoting mRNA instability rather than altering splicing patterns, while Hasegawa-Ogawa and Okano (2021) identified that wild type, but not mutant, TDP-43 could negatively regulate its own promoter activity. Given that over and under expression of TDP-43 appears toxic, this autoregulatory mechanism may be vital to prevent cellular toxicity.

TDP-43 has been identified as a Drosha associated protein and is involved in miRNA regulation with knockdown of the protein in cell culture altering miRNA expression levels, particularly the down-regulation of let-7b and upregulation of miR-663 (Buratti *et al.,* 2010). This occurs by binding of TDP-43 to the RNA in either the miRNA sequence itself (let7-b) or in the hairpin region of the precursor (miR-663) (Buratti *et al*., 2010). TDP-43 has also been shown to associate with and thus limit the availability of miRNAs miR-1 and miR-206 for the RNA-induced silencing complex (RISC) in muscle cells, so regulating the activity of these miRNAs (King *et al*., 2014).

TDP-43 is also thought to be involved in the DNA damage response (DDR) and loss of the protein leads to increased DNA double strand break (DSB) accumulation and inhibition of the non-homologous end joining (NHEJ) pathway in several model systems (Mitra et al., 2019). TDP-43 appears to be rapidly recruited to DSB sites along with other DDR factors such as Ku70/80 and γH2AX, where it acts as a scaffold for the recruitment of the XRCC4-DNA ligase 4 complex for the final ligation steps in the NHEJ pathway (Mitra and Hegde, 2019; Mitra *et al*., 2019).

TDP-43 has been shown to associate with mitochondrial membranes and appears to have roles affecting mitochondrial dynamics and function. Increased TDP-43

expression increases mitochondrial aggregation and fusion, whereas expression of ALS-related mutant TDP-43 has been linked to increased mitochondrial fragmentation; these effects may be due to altered expression of mitochondrial dynamic regulators rather than the direct association between mitochondria and TDP-43 (Gao *et al*., 2019). Overexpression of TDP-43 has also been linked to reduced oxidative phosphorylation, induction of mitophagy (thus having some role in mitochondrial quality control) and impaired mitochondrial anterograde and retrograde transport (Gao *et al*., 2019). Furthermore, in a corticospinal motor neuron model showing TDP-43 proteinopathy, Gautam *et al*. (2019) showed mitochondrial degradation by a pathway different to mitophagy, they termed mitoautophagy. Wang *et al*. (2016) found that mitochondrially localised TDP-43 binds with mitochondria-transcribed mRNAs encoding respiratory complex I subunits ND3 and ND6, impairing their expression and leading to complex disassembly.

Under conditions of cellular stress, mRNA binding proteins consolidate mRNA into cytoplasmic compartments termed stress granules, in order to sequester transcripts not required for a stress response (Hofmann *et al*., 2021). TDP-43 has been reported to localise to these stress granules by binding the RNA components and also by interacting with other stress granule proteins such as TIA-1 (Liu-Yesucevitz *et al*., 2010).

1.6 Factors leading to TDP-43 aggregation

1.6.1 Post-translational modification

The aggregation of TDP-43 can be explained in part by the structural elements identified in the protein, particularly the naturally unstructured CTD that may form transient β-rich, aggregation prone structures (Mompeán *et al*., 2014). However, given the number of mutations identified in ALS/FTLD patients, and the identified changes to

TDP-43 found in aggregates, there are clearly alterations to cellular processes and to the protein itself that lead to aggregation and the inability of the cell to remove misfolded and aggregating protein. These changes include post-translational modifications, mutation to TDP-43, fragmentation of the protein and changes to protein quality control systems (reviewed in Hergesheimer *et al.,* 2019).

TDP-43 appears susceptible to multiple post-translational modifications, some of which may influence its propensity to aggregate. The TDP-43 inclusions identified in ALS and FTLD neurons by Neumann *et al*., (2006), were ubiquitinated and phosphorylated. Lysine 79 was the only ubiquitinated site revealed in brain samples from ALS patients studied by Kametani *et al*. (2016). Given that ubiquitin is added to proteins for recognition by the UPS, these labelled aggregatory proteins could be a species not correctly degraded by the UPS which are allowed to aggregate and potentially become toxic. Some reports suggest that the majority of serine and threonine residues in the CTD may be phosphorylation targets (Kametani *et al*., 2016), with Ser409/410 being well documented in aggregates (Neumann *et al*., 2009). Phosphorylation of Ser409/410 in cell models is enhanced by oxidative stress and may increase aggregation (Goh *et al*., 2017). Contrastingly however, Gruijs da Silva *et al*. (2022), found *in vitro* hyperphosphorylation of TDP-43 purified with a solubilising maltose binding protein (MBP) tag to reduce aggregation. Furthermore C-terminal phosphomimic mutations reduced aggregation both *in vitro* and in cellular experiments, and these mimics maintained normal TDP-43 nuclear transport and RNA regulator functions (Gruijs da Silva *et al*., 2022).

Lysine acetylation is another post-translational modification that has been shown to increase aggregation of TDP-43 and acetylated Lys-145 has been detected in inclusions found in ALS spinal cord (Cohen *et al*., 2015). Expression of TDP-43 acetylation mimics, K145Q and K192Q, as well as co-transfection of the acetyltransferase Creb-binding protein (CBP) with a TDP-43 NLS mutant showed

increased formation of aggregates in a mammalian cell model (Cohen *et al*., 2015). These mimics also showed decreased RNA binding activity, suggesting acetylation may lead to some loss of function which may in turn trigger aggregation (Cohen *et al*., 2015).

Another post-translational modification to be considered is that of disulfide bonds and their aberrant formation when TDP-43 is held in stress granules for prolonged periods, promoting aggregation (Jiang et al., 2017). The aging process and environmental toxins both generate reactive oxygen and nitrogen species. S-nitrosylation of TDP-43 by reactive nitrogen species (RNS) has been shown to facilitate disulfide bond formation and subsequent TDP-43 aggregation in human induced pluripotent stem cell and *in vivo* models (Pirie *et al*., 2021). Furthermore, the aggregation triggered further nitrosative stress, leading to further aggregation, thus representing a feed-forward mechanism of S-nitrosylation induced aggregation (Pirie *et al*., 2021).

Not all post-translational modifications increase the propensity of TDP-43 to aggregate however. Zhao *et al*. (2021) showed that O-GlcNAcylation, the addition of the monosaccharide N-acetylglucosamine to Ser/Thr residues via an O-linked glycosidic bond, suppresses TDP-43 proteinopathies, while promoting TDP-43 normal splicing functions.

1.6.2 Mutations

Mutation to TDP-43 may influence the propensity of the protein to aggregate. The majority of mutations to *TARDBP* affect the CTD of the encoded protein, and some ALS related mutations in the protein, such as A315T and A315E appear to accelerate aggregation compared to wild type (Laos *et al*., 2019). Some, such as G298S, appear to increase cytoplasmic mislocalisation (Sun *et al*., 2018). A recently identified mutation outside of the CTD and adjacent to RRM1, K181E, disrupts RNA binding and increases

intra-nuclear aggregation, and when introduced into a ΔNLS construct led to increased cytoplasmic aggregation (Chen *et al*., 2019).

Given that mutation to *TARDBP* accounts for only ~5% of ALS cases (Kirby *et al*., 2016), the other mutations identified in ALS and FTLD patients point to other cellular changes that may occur in these diseases (Mathis *et al.,* 2019). As mentioned previously, alterations to proteostasis systems would allow aggregates to build up and persist. Mutations to previously mentioned UPS genes including *UBQLN2* and *VCP*, and autophagy genes including *OPTN* and *SQSTM1* implicate changes to mechanisms by which cells would clear misfolded and aggregated protein in these diseases. Furthermore, inhibition of the UPS and autophagy systems in cell models has shown increased protein aggregation with associated cytotoxicity (Wang *et al*., 2010). Defects to the endosomal-lysosomal pathway have also been implicated in TDP-43 pathology, as deletion of genes related to this pathway has been shown to increase TDP-43 aggregation and cytotoxicity in yeast (Leibiger *et al*., 2018).

1.6.3 TDP-43 protein fragments

One further change that contributes to TDP-43 aggregation is its fragmentation into 35 kDa and 25 kDa C-terminal fragments (CTFs). These fragments are often found in aggregates from ALS and FTLD brain tissue, but rarely spinal cord tissue, and are generated by caspase 3,4 and 7 cleavage of the full-length protein, particularly under conditions of cell stress, or possibly by the translation of alternative transcripts that may be upregulated in disease (Berning and Walker, 2019). Expression of CTF transcripts has been shown to produce aggregates and cause cellular toxicity in HEK293 cells (Zhang *et al*., 2009).

In contrast to the N-terminally truncated CTFs described above, research has also identified C-terminally truncated fragments and isoforms of TDP-43 to be aggregation

prone. Weskamp *et al*. (2020) demonstrated in an induced pluripotent stem cell model that neuronal hyperexcitability, a feature of ALS, led to the formation of cytoplasmic aggregates that were lacking the C-terminal region of TDP-43, and that these aggregates could further sequester full length TDP-43 by conserved N-terminal interactions. Furthermore, the authors demonstrated that the shortened TDP-43 isoforms (denoted sTDP43) were formed by an alternative splicing event utilising a splicing acceptor in the *TARDBP* 3' untranslated region, resulting in the inclusion of a unique 18 amino acid C-terminus not found in full length TDP-43. This sequence included a putative nuclear export sequence that resulted in increased cytoplasmic localisation when an sTDP43 encoding construct was expressed in cultured rodent primary neurons (Weskamp *et al*., 2020).

A further C-terminally truncated splice isoform, generated by the use of non-canonical splice sites in exon 6, skipping 1020 base pairs and encoding a 272 amino acid protein, was identified to share the same 18 amino acid C-terminus demonstrated by Weskamp *et al*., (2020) (Shenouda *et al*., 2022). When expressed in neuronal cell lines, this construct resulted in the formation of cytoplasmic ubiquitinated aggregates, and an antibody raised against the 18 amino acid C-terminal sequence could be used in immunofluorescent staining of ALS spinal cord tissues to show elevated levels of C-terminally truncated TDP-43 alongside TDP-43 pathology (Shenouda *et al*., 2022).

1.7 Prion-like properties of aggregates

The ability of the CTD to transiently form various secondary structures including β-sheet rich regions could mean that it can functionally resemble a β-sheet rich prion protein- proteins that misfold and propagate by inducing misfolding in other prion proteins in neurodegenerative diseases such as Creutzfeldt-Jacob disease (Das and Zou, 2016). These proteins can transmit between neurons and so propagate the

disease through the nervous system. Propagation of TDP-43 aggregates between neurons may be an explanation for the documented progression of ALS through upper and lower motor neurons, hypothesised to start at a particular focal point (Ravits *et al*., 2007). Supporting this prion like propagation hypothesis, insoluble TDP-43 aggregates prepared from ALS and FTLD brains have been shown to seed intracellular aggregation in SH-SY5Y cells, and to induce cell death in these cells (Nonaka *et al*., 2013). In the same study, aggregates were shown to propagate between cultured cells in co-culture experiments. In a mouse model, Ding *et al*. (2021) showed that injection of pre-formed TDP-43 fibrils into the primary cortex of mice expressing human TDP-43 induced the anterograde spreading of pathological TDP-43 along pyramidal tract axons, leading to the display of ALS-like neuropathological symptoms. In their study of the role of S-nitrosylation in TDP-43 aggregation, Pirie *et al*., (2021), demonstrated that exposing cultured neuronal cells to extracellular recombinant TDP-43 led to a dramatic increase in neuronal reactive nitrogen species (RNS) production. In turn, these cells exhibited increased TDP-43 aggregation and cell death.

How this spread of aggregates occurs remains to be fully elucidated, however TDP-43 aggregates have been shown to be secreted in exosomes (Iguchi *et al*., 2016), and transmission also may occur across synaptic connections, or by tunnelling nanotubes as is seen with α-synuclein aggregates (Nonaka and Hasegawa, 2020).

Furthermore, a *C9orf72* repeat expansion co-culture model by Khosravi *et al*. (2020) showed that cell to cell transmission of a poly-(Gly-Ala) (poly-GA) dipeptide repeat protein could promote mislocalisation and aggregation of TDP-43 CTFs in neighbouring cells, by inhibition of proteasome function following poly-GA uptake. This suggests that TDP-43 aggregates are not the only species that can lead to the apparent cell to cell transmission of the TDP proteinopathy.

To support a “dying-forward” hypothesis of disease spreading, that is, progression of the pathology and disease into the lower motor neurons from an initiation point in the

motor cortex, a mouse model expressing ΔNLS TDP-43 in the motor cortex demonstrated a primary cortical pathology which was then followed by a secondary degeneration of the lower motor neurons (Reale *et al*., 2023).

1.8 The molecular architecture of TDP-43 aggregation

In several neurodegenerative diseases (e.g. Alzheimer's, Parkinson's and Huntington's disease), the misfolded or mutant proteins aggregate into amyloid fibrils of a width of approximately 10nm and a length of 0.1-10μm (Ross and Poirier, 2004). Within these fibrils, the proteins form β-sheet organised regions. Analysis of amyloid fibrils of amyloid-β (as seen in AD), of α-synuclein (as seen in PD) and of expanded polyglutamine repeat proteins such as Huntingtin in Huntington's disease, shows a common β-sheet plus β-turn structure shown in Fig.1.8 (Ross and Poirier, 2004).

Mature amyloid fibrils were long considered the key species responsible for cytotoxicity in amyloid neurodegenerative diseases, given that they were the pathological hallmarks of the disease. However, this idea has been challenged by evidence for pre-fibrillar species, such as protofibrils and soluble oligomers being the main cytotoxic species, with the larger amyloid fibres and inclusions being produced as a protective "sink" (Lambert *et al*., 1998; Walsh *et al*., 2002). This cytotoxicity may come from the misfolded nature of the oligomers, exposing aggregation prone regions in the protein allowing them to organise into larger structures and interact with cell components such as the mitochondria, microtubules and ER, affecting normal cellular function (Stefani, 2010). The formation of the soluble oligomers may also lead to increased production of ROS. Tabner *et al*. (2005) demonstrated the accumulation of hydrogen peroxide in the early soluble oligomer and protofibril stages of amyloid-β peptide aggregation with further work demonstrating that Aβ (1-42) aggregates and fibrils retain a redox ability and degrade H_2O_2 in the presence of Cu (II) ions (Mayes *et al*., 2014). Hydrogen

peroxide is readily converted to the hydroxyl radical via the Fenton reaction, and this species could be responsible for oxidative damage to the neuron.

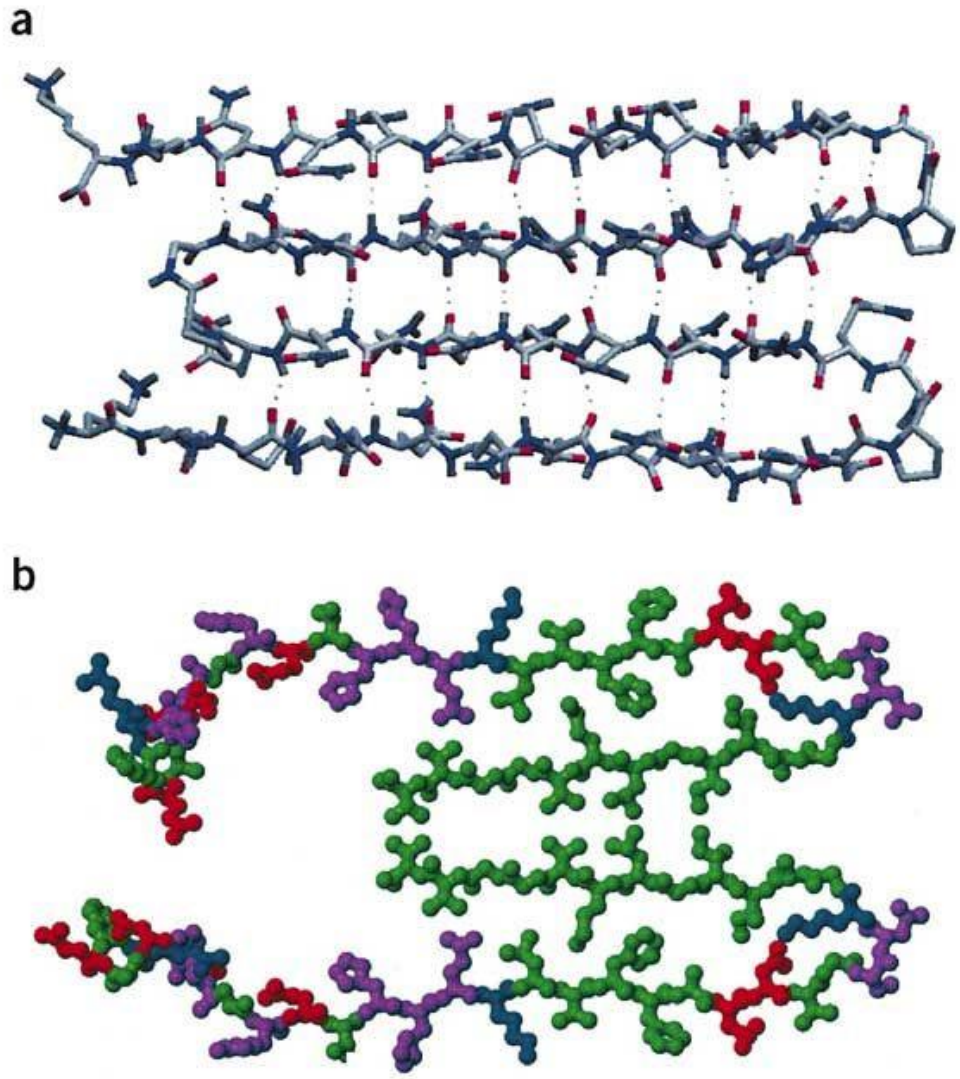


Figure 1.8. β-sheet plus β-turn structures found in amyloid proteins. (a) Expanded polyglutamine proteins and (b) in a model of amyloid-β (1-40). Taken from Ross and Poirier (2004).

Aggregates of TDP-43 appear to be structurally different to other amyloid proteins. Cryo-electron microscopy of aggregates isolated from frontal and motor cortices of two patients diagnosed with ALS with FTLD showed an “amyloid-like filament” with a core

consisting of a double spiral shaped fold of amino acids 282-360, (Fig. 1.9) (Arseni *et al*., 2021). Abundant glycine and neutral polar residues allow for 10 short beta strands, with only two longer than 3 residues. The fold appears to “stack”, with exposed polar residues on the surface and hydrophobic residues buried within the core (Arseni *et al*., 2021).

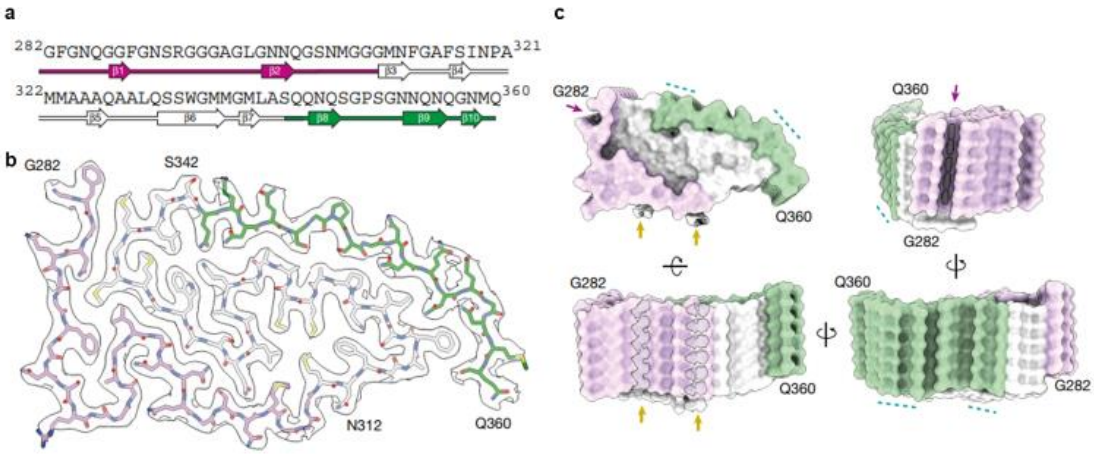


Figure 1.9. The core architecture of TDP-43 aggregates. (a) The amino acid sequence corresponding to the core of the double spiral fold and the corresponding beta sheets formed from it. (b) The structure of the double spiral fold gained from cryo-EM. (c) An atomic model showing the stacking of 6 molecules to form a filament. Magenta indicates the glycine rich region, white the hydrophobic core and green the Q/N rich region. Yellow arrows indicate additional planar densities adjacent to flat strips on the surface of the glycine-rich region, magenta arrows indicate additional densities within the prominent groove formed by the main chain of G282–Q286 and the side chain of Q286, and cyan dashes indicate polar patches on the surface of the Q/N rich region. Adapted from (Arseni et al., 2021).

In fact, multiple regions within TDP-43 have been identified to be aggregation prone and Arseni *et al*. (2021) suggested that filaments formed in different TDP-43 pathologies may have a different structure as indicated by their differing sensitivities to proteolysis. Fig. 1.10 shows an AGGRESCAN prediction for TDP-43; regions above

the “hotspot threshold” (blue) are computationally predicted to be aggregation prone, and many of these regions have been identified to be aggregation prone empirically. TDP-43 may, therefore, have multiple regions that could be centres of aggregation, especially considering the aggregation prone nature of C-terminally truncated fragments that would not contain the double spiral fold core region.

Further supporting that aggregates of TDP-43 are non-amyloid in nature, cytoplasmic aggregates formed by overexpression of TDP-43 in NSC-34 cells do not bind amyloid diagnostic dyes such as thioflavin T (ThT), and lack characteristic β-sheet enriched structure (Cascella *et al*., 2022). Furthermore, they do not demonstrate fibrillar morphology, instead appearing round and granular in structure (Cascella *et al*., 2022). Aggregates formed *in vitro* from recombinant TDP-43 may also show a lack of filamentous amyloid structure with the same absence of ThT fluorescence, with far-UV circular dichroism spectroscopy revealing an α-helix and random coil containing structure with a lack of β-sheet enrichment (Cascella *et al*., 2022). This lack of amyloid nature was also suggested earlier by lack of binding of aggregated TDP-43 constructs to amyloid dyes by Johnson *et al*. (2009). The recombinant protein aggregate structures suggested by Cascella *et al.* (2022) are contrasting to those demonstrated by Kumar *et al*. (2023) however, who produced TEM images of distinctly filamentous recombinant TDP-43 aggregates.

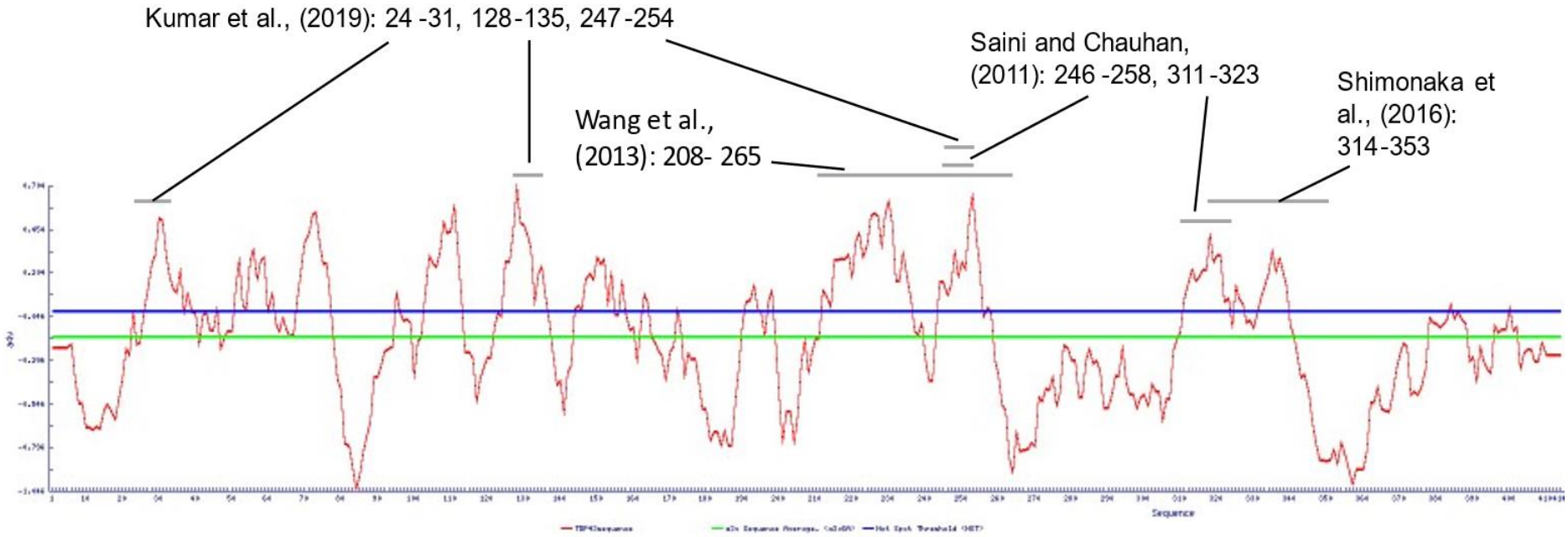


Figure 1.10. AGGRESCAN result showing the average aggregation propensity (a4V) (y-axis) for each amino acid in the TDP-43 sequence and corresponding sequences identified experimentally. Red, Aggregation propensity of each residue in the TDP-43 amino acid sequence; Green, aggregation-propensity values per amino acid sequence average; Blue, Hot spot threshold. Hotspots identified empirically in the literature are shown with grey lines corresponding to the identified region.

1.9 The pathological implications of TDP-43 aggregation

The role of the aggregation of TDP-43 in neurodegenerative disease is often debated, given that in some TDP-43 models cytotoxicity does not correlate with increased aggregation (as reviewed by Hergesheimer *et al*., 2019). Equally however, other models, some of which previously mentioned, do reflect that cytotoxicity and TDP-43 aggregation are linked (Zhang *et al*., 2009; Nonaka *et al*., 2013; Cascella *et al*., 2022). Aggregation may be pathogenic by both a gain and loss of function mechanism. In terms of the former, the sequestration of the protein into cytoplasmic aggregates would result in it being unable to fulfil its normal roles, thus affecting gene expression, mRNA regulation and splicing, autoregulation, DDR, miRNA regulation and its mitochondrial roles, therefore causing widespread cellular dysfunction. Indeed, knockdown and the resulting loss of function of TDP-43 in adult mouse hippocampus resulted in loss of functional mature excitatory synapses and cognitive deficits (Ni *et al*., 2021). TDP-43 knockdown mice demonstrate age-dependent motor dysfunction and neurodegeneration, with increased cytoplasmic dsRNA accumulation in the dorsal and ventral horns of the spinal cord compared to wild type controls (Milstead *et al*., 2022). Neurodegeneration was particularly evident in areas of increased dsRNA deposition, with increased astro- and micro-gliosis also identified (Milstead *et al*., 2022).

There could also be a gain of function with regard to the TDP-43 aggregates, sequestering other proteins such as RBM14 and PSF along with mRNAs to inclusion bodies, and so preventing their normal activity (Dammer *et al*., 2012). Altman *et al*. (2021) identified Ras GTPase-activating protein binding-protein1 (G3BP1) to co-localise with axonal TDP-43 condensates. The authors also found this aggregation to result in reduced local protein synthesis by translation suppression, particularly of nuclear encoded mitochondrial proteins including respiratory chain complex proteins. This led to reduced mitochondrial integrity and function, apparently resulting in reduced function of neuromuscular junctions and increased neurodegeneration.

To support the gain of function effect further, Cascella *et al*., (2022), expressed full length TDP-43 in a motor neuron/neuroblastoma hybrid model and monitored the formation of various sizes of cytoplasmic aggregates with time while concomitantly monitoring levels of oxidative stress and mitochondrial dysfunction. Via this method, it was found that the highest levels of neuronal dysfunction correlated with the formation of the largest aggregates, rather than oligomers and smaller aggregate species (Cascella *et al*., 2022),

Furthermore, seeding of SY5Y cells with aggregates purified from ALS brain tissue leads to formation of aggregates and toxicity in the cells, suggesting the aggregates themselves may have effects, as well as demonstrating the ability of misfolded aggregated TDP-43 to trigger further aggregation as seen in cell-to-cell spreading (Nonaka and Hasegawa, 2012).

To support the role of TDP-43 pathology further, and providing a potential mechanism of cell dysfunction, in a ΔNLS mouse model, which expresses human TDP-43 with an ablated nuclear localisation sequence in brain and spinal cord neurons, the activation of cell stress components was investigated by Luan *et al*., (2023). This model demonstrates TDP-43 pathology with progressive motor degeneration. Activation of the integrated stress response effectors (CCAAT/enhancer-binding homologous protein (*Chop/Ddit3*) and activating transcription factor 4 (*Atf4*)) was identified, with up-regulation of both anti- (Bcl2) and pro-apoptotic genes (BH3-interacting domain death agonist (*Bid*)) in early disease stages, which is replaced with predominantly pro-apoptotic signalling in later stages of disease progression, particularly after motor dysfunction phenotype onset (Luan *et al*., 2023). The integrated stress response is a signalling cascade which maintains protein synthesis, folding and decay, which by prolonged activation can trigger downstream apoptosis (Costa-Mattioli and Walter, 2020).

1.10 Development of therapies targeting TDP-43 aggregation

Preventing TDP-43 aggregation may be a potential therapy target for diseases such as ALS and FTLD. There are currently few drugs in development that specifically target the TDP-43 aggregation process. One is the Polyglutamine binding peptide 1 (QBP1), which inhibits the aggregation of several other amyloidogenic proteins (e.g. Huntingtin), and appears to inhibit aggregation of TDP-43 fragments containing its Q/N rich region (Mompeán *et al*., 2019). AIM4, an acridine derivative, has been shown to significantly reduce aggregation of CTD fragments, and reduce the aggregation of YFP tagged TDP-43 in a yeast model (Prasad *et al*., 2016). Antibodies to the epitopes 178-393 and 256-269, within the RRM2 and CTD regions, have been shown to reduce *in vitro* aggregation of a phosphorylated TDP-43 construct at low antibody to protein ratios (Esposto & Martic, 2021). A mitochondrial localisation inhibitory peptide, PM1, has been shown to reduce mitochondrial mis localisation of TDP-43 with some alleviation of TDP-43 proteinopathy (Gao *et al*., 2020). Recently one of 18 artificial peptide binders designed based on mathematical contact energies between residues targeted to the LCD successfully reduced the aggregation of TDP-43 constructs *in vitro* when used at 50 times molar excess to the protein (Kamagata *et al*., 2023). Interestingly, greater aggregation inhibition was shown by the positive control of an anti-TDP-43 antibody at much lower concentrations in this work.

Liu *et al*. (2013) tested 5 peptide inhibitors based on potential aggregation sequences determined in their own studies. These peptides were successful at reducing aggregation of both a TDP-43 construct lacking residues 187-192 and TDP-43 aggregation triggered by arsenite stress in a HeLa cell overexpression model. This did not however reduce the cytotoxicity caused by the TDP-43 overexpression used in this model. The authors concluded that TDP-43 aggregation was therefore not the cause of cell death in their model. However, as TDP-43 expression is tightly controlled by autoregulation, including association with its own promoter (see section 1.5.2), and as

TDP-43 appears to have many cellular targets that may be affected by overexpression, an overexpression model may not be truly representative of the expression and aggregation conditions found in neurodegenerative disease.

As an alternative to targeting the aggregating protein itself, Ibudilast appears to increase aggregate clearance by autophagy enhancement (Chen *et al*., 2020), and *Bis*(thiosemicarbazonato)-copper complexes appear to reduce aggregate formation by modulation of cellular kinase activity (Parker *et al*., 2012). Specifically targeting the preformed aggregates for clearance could be one useful therapeutic approach. Proteolysis targeting chimeras (PROTACs), consisting of an aggregate binding moiety of BTA (benzothiazole-aniline derivative, produced by removing a methyl group from the amyloid binding dye ThT) coupled to pomalidomide (to bind and recruit E3 ligase) have been produced, which selectively targets protein aggregates for degradation by the proteasome by ubiquitination (Tseng *et al*., 2023). One such PROTAC decreased C-terminal fragment aggregates and rescued cytotoxicity in Neuro-2a cells, and furthermore improved the mobility of a *C. elegans* model by reducing aggregates in the nervous system (Tseng *et al*., 2023).

1.11 Project aims

This project aims to develop a purification method for recombinant full length, untagged TDP-43, with a view to using it in aggregation assays. With minimal literature available on *in vitro* aggregation assays for TDP-43, another aim will be to develop and optimise such an aggregation assay. These assays will be used to investigate molecules identified in the literature to have aggregation inhibitory ability against other proteins for repurposing against TDP-43. Furthermore, the assays will be used to test and develop peptide inhibitors specifically designed against aggregation prone sequences of TDP-43 identified in the literature or predicted points of TDP-43 interaction generated *in silico*. An *S. cerevisiae* TDP-43 aggregation model will also be developed; a cell model in which any potential inhibitor molecules could be tested for anti-aggregation or toxicity-reducing effects.

2. Materials and Methods

2.1 Plasmid DNA purification

Bacterial expression plasmids used in this project (Table 2.1) were kindly provided by Dr Fiona Benson. Yeast expression plasmids encoding fluorescently tagged TDP-43 were acquired from Addgene. pRS416Gal TDP43 WT YFP and pRS426Gal TDP43 WT GFP were a gift from Aaron Gitler (Addgene plasmid # 27447 and # 27467 respectively). The plasmids were stored in *Escherichia coli* bacterial glycerol stocks prior to the project. To purify plasmids from these stocks, the bacteria were cultured overnight at 37ºC with 250 rpm shaking in 50 ml Terrific broth (Melford Biolaboratories) containing kanamycin (30 µg ml^{-1}) for pKMH6 and pKMH9 or ampicillin (50 µg $ml^{-}1$) for pRS416Gal TDP43 WT YFP and pRS426Gal TDP43 WT GFP. Bacteria were harvested by centrifugation using a Beckman Avanti centrifuge with a JA25.50 rotor (7250 rpm, 4ºC, 20 min). Plasmids were purified from the resulting pellets using the Qiagen Midi-prep plasmid purification kit according to the supplier's instructions, using a Beckman Avanti centrifuge with a JA25.50 rotor for centrifugation steps.

Table 1.1. TDP-43 and TDP-43 CTF encoding plasmids.

Plasmid name	Vector	TDP-43 amino acid residues	Encoded Protein	Tag
pKMH6	pET24a	1-414	Full length TDP-43	No tag
pKMH9	pET24a	208-414	TDP-43 CTF	C-His
pRS416Gal TDP43 WT YFP	pRS416	1-414	TDP-43-YFP	C-terminal YFP
pRS426Gal TDP43 WT GFP	pRS426	1-414	TDP-43-GFP	C-terminal GFP

2.2 Ethanol precipitation of DNA

Ethanol precipitation was used to concentrate purified plasmids to a concentration high enough for sequencing. 3 M sodium acetate, pH 5.2, was added to the plasmid sample (1/10 sample volume), followed by 2 volumes of -20ºC, 100% ethanol. The reaction was incubated for 1 hour at -20ºC, then centrifuged in an Eppendorf 5424 (14,680 rpm, 20 min, 4ºC). The supernatant was removed and 300 µl of -20ºC, 80% (v/v) ethanol was added to the pellet, being careful to not resuspend. The sample was again centrifuged using the Eppendorf 5424 (14,680 rpm, 5 min, 4ºC). The supernatant was discarded, and the pellet resuspended in a volume of MilliQ water appropriate for the desired level of concentration. Concentrated plasmids were then stored at -20ºC.

2.3 Bacterial transformation

SHuffle™ T7 Express *E. coli* (New England BioLabs) (Genotype: *fhuA2 lacZ::T7 gene1 [lon] ompT ahpC gal λatt::pNEB3-r1-cDsbC* (Spec^R, *lacI^q*) *ΔtrxB sulA11 R(mcr-73::miniTn10*--Tet^S*)2 [dcm] R(zgb-210::Tn10* --Tet^S*) endA1 Δgor Δ(mcrC-mrr)114::IS10)* were used for recombinant protein expression. To transform these cells with the plasmid constructs, a tube of frozen SHuffle™ stock was thawed on ice. Approximately 3 ng of purified and concentrated plasmid DNA was added to the cells, and carefully mixed. The mixture was incubated on ice for 30 min, then heat shocked in a 42ºC water bath for 30 seconds, followed by a further 5 min incubation on ice. Room temperature SOC outgrowth media (New England BioLabs) (950 µl) was added to the sample, followed by incubation at 30ºC with 250 rpm shaking for 1 hour. The cells were then mixed thoroughly, and diluted 1 in 10 in SOC media. Dilutions were spread onto LB plates containing kanamycin (30 µg ml^{-1}) and incubated overnight at 37ºC.

2.4 Small scale protein expression

For small scale protein expression, typically a colony of transformed Shuffle™ *E. coli* was selected and added to 3 ml of LB media containing kanamycin (30 µg ml^{-1}), and incubated overnight at 37ºC with 250 rpm shaking to produce a starter culture. This was then added to 100 ml of LB media containing kanamycin (30 µg ml^{-1}), and the larger culture was incubated at the desired temperature until mid-log phase (OD600 of 0.6) with 250 rpm shaking, then protein expression was induced by the addition of 1 M IPTG to a 1 mM final concentration. Cultures were typically grown for 3 hours at 30ºC or 25ºC, or overnight at 16ºC with 250 rpm shaking before cells were harvested by centrifugation (2500 rpm, 8 min, Eppendorf 5810R).

For analysis of expression over time, 1 ml samples were taken at hourly intervals, and cells were harvested by centrifugation (5000 rpm, 3 min, Eppendorf 5424). Cells were resuspended in 40 µl of prepared 4x Laemmli Sample Buffer (BioRad), heated at 98ºC for 3 min, then stored at -20ºC.

2.5 Small scale solubility analysis

Cultures from small scale protein expression were centrifuged (2500 rpm, 8 min, Eppendorf 5810R) and the supernatant discarded. The cells were resuspended in 1/20th of the culture volume of BugBuster™ (Novagen) and incubated at room temperature on a rocker for 20 min. The sample was sonicated in a room temperature water bath sonicator (Grant SUB6) for 6 min to complete lysis, then centrifuged (13,000 rpm, 30 min, 4ºC, Eppendorf 5424). The supernatant was retained as the soluble cell fraction, and the pellet was resuspended in an equal volume of a denaturing resuspension buffer (50 mM HEPES, pH 8.0, 300 mM NaCl, 10 mM imidazole, 10 mM DTT, 8 M urea) as the insoluble cell fraction. Samples of each were combined 1:1 with prepared Laemmli Sample Buffer (BioRad) and subjected to SDS-PAGE.

2.6 Large scale protein expression

For large scale protein expression, a starter culture was set up by addition of a colony of transformed Shuffle™ *E. coli* picked from a selection plate to 30 ml of LB media containing kanamycin (30 µg ml^{-1}), followed by overnight incubation at 37ºC with 250 rpm shaking. This starter culture was then added to 970 ml of LB media containing kanamycin (30 µg ml^{-1}), to produce a 1 litre culture, which was incubated at 30ºC with 250 rpm shaking until mid-log phase (OD600 of 0.6). Protein expression was induced by the addition of 1 M IPTG to a final 1 mM concentration, and the cultures were incubated for 3 hours at 37ºC or overnight at 20ºC with 250 rpm shaking. Cells were

harvested by centrifugation using a Beckman Avanti centrifuge with a JLA8.1000 rotor (3000 rpm, 25 mins, 16ºC).

2.7 SDS-PAGE

SDS-PAGE gels were self-cast using the Bio-Rad mini protean system (8.3 x 7.3 x 0.1 cm gel size) or Cambridge Electrophoresis EV200 equipment (165 x 165 x 2 mm or 165 x 165 x 1 mm gel size). Resolving gels routinely consisted of 10%, 12% or 15% acrylamide/bis acrylamide (37.5:1) in 375 mM Tris-HCl, pH 8.8, 0.1% (w/v) SDS, 0.05% (w/v) APS and 0.05% (v/v) TEMED. Stacking gels contained 4% acrylamide/bis acrylamide (37.5:1), in 125 mM Tris-HCl, pH 6.8, 0.1% (w/v) SDS, 0.1% (w/v) APS and 0.1% (v/v) TEMED. Gels were run in SDS-PAGE running buffer (25 mM Tris, 192 mM glycine, 0.1% (v/v) SDS, pH 8.3)

2.7.1 Analytical SDS-PAGE

Protein samples were combined 3:1 with 4x Laemmli sample buffer (BioRad), or 1:1 with 2x SDS sample buffer (Invitrogen™) and heated for 3 min at 98ºC. Precision Plus Protein™ All Blue pre-stained protein standards (Bio-Rad) were routinely used as molecular weight markers. Gels were run in SDS-PAGE running buffer for 1-1.5 hours at 160 V, then stained for 2-4 hours in Coomassie Blue stain (20% (v/v) methanol, 10% (v/v) glacial acetic acid, 0.125% (w/v) Coomassie Blue R250) and de-stained overnight in de-staining solution (20% (v/v) methanol, 10% (v/v) glacial acetic acid). Alternatively, gels were stained with Instant blue (Abcam), for 1 hour to overnight. Gels were then imaged using a ChemiDoc™ XRS system (BioRad).

2.7.2 Preparative SDS-PAGE

Samples were combined 1:1 with 2x SDS-PAGE sample buffer (Thermo Scientific) then loaded onto the gel without heating. For small scale preparative SDS-PAGE, gels were run at 130 V until the dye front reached the bottom of the gel, approximately 1.5-2 hours. For large scale, the gel was run at 100-150 V, maintaining wattage at 5 W, for 6-8 hours, until the ladder band corresponding to 37 kDa reached the end of the gel.

Gels were then negatively stained with the Pierce Zinc Reversible stain kit (Thermo Scientific), according to the supplier's instructions to visualise the band of interest. The band was then cut from the gel and de-stained, again according to the supplier's instructions.

2.8 Immunoblotting

Following SDS-PAGE, proteins were transferred to an Immobilon-P PVDF membrane (Millipore) pre-equilibrated in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol) using a BioRad Criterion blotting system, according to the supplier's instructions, with the application of 100 V for 1 hour. Following transfer, the membrane was rinsed in phosphate buffered saline (PBS) (Sigma) for 5 min, then submerged in PBS containing 0.1% (v/v) Tween20 and 5% (w/v) Marvel milk powder, and incubated on a rotamixer for 1 hour at room temperature. The membrane was then rinsed in PBS for 5 minutes, and the primary antibody, diluted appropriately (1:2000 for rabbit polyclonal anti-TDP-43 antibody (Invitrogen^{TM})) in 5 ml of PBS containing 0.1% (v/v) Tween20 and 5% (w/v) Marvel milk powder, was added, and mixed on the rotamixer at 4ºC overnight. The membrane was then washed once for 1 min and twice for 15 min in PBS containing 0.1% (v/v) Tween20. The secondary antibody diluted appropriately in 20 ml of PBS containing 0.1% (v/v) Tween20 and 5% (w/v) Marvel was then added (1:8000 for Anti Rabbit IgG F(ab')2Peroxidase antibody produced in goat (Sigma)), and

the membrane was mixed on the rotamixer at room temperature for 1 hour. The membrane was then washed once for 1 min and twice for 15 min in PBS, before the blot was developed using Clarity ECL reagent (BioRad) and imaged using the ChemiDoc^{TM} XRS system (BioRad).

2.9 Isolation and solubilisation of inclusion bodies following lysozyme/sonication lysis

Cells harvested from large scale protein expression were resuspended in 50 mM Tris-HCl, pH 8.0, 500 mM NaCl (40 ml). 50 mg ml^{-1} lysozyme (1 ml) (Thermo Scientific), 1 U µl^{-1} DNase (100 µl) (Thermo Scientific) and 1 tablet of Pierce Protease Inhibitor (Thermo Scientific) was added and the sample was incubated on ice for 30 min with stirring. The sample was sonicated on ice using a probe sonicator (50 kHz, 6 cycles of 30 seconds), then centrifuged using a Beckman Avanti centrifuge with a JA25.50 rotor (24,000 g, 4ºC, 30 min). The supernatant was discarded, and the pellet was washed sequentially with 50 mM HEPES, pH 8.0, 1% Triton X-100, (40 ml), followed by 50 mM HEPES, pH 8.0, 2 M NaCl (40 ml), then 50 mM HEPES, pH 8.0 (40 ml). Insoluble material was pelleted between each wash by centrifugation using a Beckman Avanti centrifuge with a JA25.50 rotor (24,000 g, 4ºC, 30 min).

The supernatant was discarded and the pellet was then dissolved overnight at room temperature in denaturing buffer (50 mM HEPES, pH 8.0, 300 mM NaCl, 10 mM imidazole, 10 mM DTT, 8 M urea) (100 ml), with stirring. The resulting solution was centrifuged using a Beckman Avanti centrifuge with a JA25.50 rotor (24,000 g, 4ºC, 30 min) to remove any remaining debris. The supernatant was retained and stored at -80ºC.

2.10 Inclusion body preparation and solubilisation using BugBuster™

Cells harvested from large scale protein expression were resuspended in 5 ml/g of pellet in BugBuster™ (Novagen). 25 units of benzonase per ml of BugBuster™ was added and the mixture was incubated for 40 minutes at room temperature on a rotamixer. The suspension was centrifuged using a Beckman Avanti centrifuge with a JA25.50 rotor (16,000 g, 4ºC, 20 min), and the pellet was resuspended in the same volume of BugBuster™ as was the initial cell pellet. Lysozyme was added to a final concentration of 1000 U/ml, the sample was mixed by vortexing and incubated on a rotamixer for 5 minutes at room temperature. An equal volume of 1 in 10 BugBuster™ diluted in sterile water was added, the sample was mixed well, then centrifuged using a Beckman Avanti centrifuge with a JA25.50 rotor (5000g, 4ºC, 15 min). The pellet was resuspended in 10 ml/g of original pellet weight in 1 in 10 BugBuster™ diluted in sterile water, then centrifuged using a Beckman Avanti centrifuge with a JA25.50 rotor (5000g, 4ºC, 15 min). This step was repeated twice (a total of 3 washes), then the final pellet was resuspended in 10 ml denaturing buffer (50 mM HEPES, pH 8.0, 300 mM NaCl, 10 mM imidazole, 10 mM DTT, 8 M urea). The sample was then incubated overnight at room temperature on the rotamixer to completely dissolve the inclusion body pellet.

2.11 Concentration of protein samples

Up to 200 ml volume of protein samples were concentrated using an Amicon® Stirred Cell, 200 ml, with a 10 kDa cellulose cut off membrane (Merck-Millipore). The membrane was prepared as outlined in the membrane instructions, and the cell pressurised at up to 3 bar.

For concentration of CTF samples, Vivaspin™ 20 centrifugal concentrators with a 5 kDa molecular weight cut off were used. Samples were centrifuged at 4000 rpm at

12ºC (Rotanta 460R, Hettich Zentrifugen), until the sample volume was sufficiently reduced to provide the necessary concentration factor.

2.12 Buffer exchange by dialysis

Samples were transferred into SnakeSkin™ pleated dialysis tubing (Thermo Scientific) (3 kDa cut off), or Slide-a-lyzer™ cassettes (Thermo Scientific) and dialysed against a total of 4 L of buffer, with two changes, at a minimum of 2 hours each (2x 2 hours in 1 L, overnight in 2 L).

2.13 Hydroxyapatite chromatography

Protein samples were dialysed into hydroxyapatite binding buffer (5 mM KPO_4, pH 6.8, 10% (v/v) glycerol, 50 mM KCl, 500 nM DTT, 0.2% PEG 3350). Hydroxyapatite chromatography was carried out on an Akta Prime Plus chromatography system, using a BioRad BioScale mini CHT column, 5ml, at 4ºC. The column was washed with 5 column volumes of 1 M NaOH, followed by equilibration with 10 column volumes of hydroxyapatite binding buffer. Protein sample was applied to the column and the column washed with 5 column volumes of hydroxyapatite binding buffer. Proteins retained by the column were eluted with 10 column volumes linear gradient up to 100% hydroxyapatite elution buffer (500 mM KPO_4, pH 6.8, 10% (v/v) glycerol, 50 mM KCl, 500 nM DTT, 0.2% PEG 3350), followed by a wash with 5 column volumes of hydroxyapatite elution buffer.

2.14 Anion exchange chromatography under denaturing conditions

For exploration of anion exchange chromatography as a primary purification step, the inclusion body pellet obtained from BugBuster™ inclusion body preparation was denatured directly into 50 mM Tris-HCl pH 8, 500 nM DTT, 8 M urea.

Anion exchange chromatography under denaturing conditions was performed at room temperature using an Akta Prime Plus chromatography system. 5 ml HiTrap Q XL columns (Cytiva) were equilibrated in denaturing anion exchange binding buffer (50 mM Tris-HCl pH 8, 500 nM DTT, 8 M urea) before sample application. The column was washed with 5-10 column volumes of denaturing anion exchange binding buffer, then the protein was eluted using denaturing anion exchange elution buffer (50 mM Tris-HCl pH 8, 500 nM DTT, 8 M urea, 1 M NaCl).

For gradient elution, elution was carried out with 10 column volumes linear gradient up to 100% elution buffer, after which the column was washed with 5 column volumes of 100% elution buffer.

2.15 Preparation of denatured bacterial whole cell lysate for IMAC

Bacterial cells were harvested from large scale culture, then the pellet resuspended in denaturing IMAC binding buffer for His-tagged protein (20 mM Tris-HCl pH 8, 8 M urea, 500 mM NaCl, 0.02% Triton X-100, 20 mM imidazole). The resulting suspension was subjected to 3 freeze thaw cycles at -20ºC, then the sample was sonicated using a probe sonicator on ice (50 kHz, 6 cycles of 30 seconds, or until the sample viscosity was reduced). The sample was centrifuged using a Beckman Avanti centrifuge with a

JA25.50 rotor (25 000g, 16ºC, 30 min), and the supernatant retained for chromatography.

2.16 Immobilised Metal Affinity Chromatography (IMAC)

For the preparation of Cu^{2+}, Ni^{2+} or Zn^{2+} IDA columns, 0.5 ml of 100 mM $CuSO_4$, $NiCl_2$ or $ZnCl_2$ respectively, were applied to 1 ml HiTrap Chelating columns (Cytiva) using an Akta Prime Plus chromatography system, according to the manufacturer's instructions. Columns were washed with 20 column volumes of MilliQ water, before equilibration in binding buffers for use. For the preparation of zinc Sepharose columns, a 1 ml HisTrap FF column (Cytiva) was stripped using 15 column volumes of 20 mM $NaPO_4$, 500 mM NaCl, 50 mM EDTA, pH 7.4, then washed with 20 column volumes of MilliQ water. $ZnCl_2$ (0.5 ml, 100 mM) was applied, then the column was washed with 20 column volumes of MilliQ water, before equilibration in binding buffers for use.

2.16.1 IMAC of full length, untagged TDP-43 under native conditions

For purification of full length untagged TDP-43 under native conditions, protein samples were dialysed into native IMAC binding buffer: 20 mM Tris-HCl pH 8, 200 mM KCl, 100 mM sucrose, 0.2% PEG 3350, 0.025% Triton X-100. Sample was applied to the appropriate pre-equilibrated IMAC column using a superloop, and the column was washed with 10 column volumes of binding buffer. Gradient elution was carried out with 10 column volumes linear gradient up to 100% elution buffer (20 mM Tris-HCl pH 8, 200 mM KCl, 100 mM sucrose, 0.2% PEG 3350, 0.025% Triton X-100, 500 mM imidazole), after which the column was washed with 5 column volumes of 100% elution buffer.

2.16.2 IMAC of full length, untagged TDP-43 under denaturing conditions

For purification of full length untagged TDP-43 under denaturing conditions as a first purification step, inclusion body pellets were resuspended in denaturing IMAC binding buffer (20 mM $NaPO_4$, pH 8, 8 M urea, 500 mM NaCl), centrifuged to remove any insoluble material using a Beckman Avanti centrifuge with a JA25.50 rotor (25 000g, 16ºC, 30 min) and applied to the appropriate pre-equilibrated IMAC column using a superloop. The column was washed with 10 column volumes of binding buffer and gradient elution was carried out with 10 column volumes linear gradient up to 100% elution buffer (20 mM $NaPO_4$, pH8, 8 M urea, 500 mM NaCl, 500 mM imidazole) after which the column was washed with 5 column volumes of 100% elution buffer.

2.16.3 IMAC of His-tagged TDP-43 CTF under denaturing conditions

For the purification of His-tagged protein TDP-43 CTF under denaturing conditions the sample was prepared as in **2.15** and loaded onto a 1 ml HisTrap FF column (Cytiva), washed with 10 column volumes of binding buffer and gradient elution was carried out with 10 column volumes linear gradient up to 100% elution buffer (20 mM Tris-HCl, pH8, 8 M urea, 500 mM NaCl, 0.02% Triton X-100, 500 mM imidazole) after which the column was washed with 5 column volumes of 100% elution buffer. Alternatively, for step elution, after the post-loading wash, protein was eluted with 10 column volumes of 20 mM Tris-HCl, pH8, 8 M urea, 500 mM NaCl, 0.02% Triton X-100, 300 mM imidazole.

2.17 Heparin chromatography

Pooled peaks from IMAC chromatography under native conditions were pooled and dialysed into heparin binding buffer (20 mM Tris-HCl, pH 8, 0.2% PEG 3350, 0.025% Triton X-100 and 100 mM sucrose), then applied by superloop to a 5 ml HiTrap heparin

HP column (Cytiva) pre-equilibrated in heparin binding buffer. The column was washed with 10 column volumes of binding buffer, before protein was eluted with 10 column volumes linear gradient up to 100% heparin elution buffer (20 mM Tris-HCl, pH 8, 0.2% PEG 3350, 0.025% Triton X-100, 100 mM sucrose and 1.5 M KCl) after which the column was washed with 5 column volumes of 100% elution buffer.

2.18 Size exclusion chromatography (SEC)

2.18.1 Preparative SEC of TDP-43 CTF

For preparative SEC of TDP-43 CTF following IMAC, 3.5 ml of sample was filtered using a 0.2 µm syringe filter and applied to a HiPrep 26/60 Sephacryl S-300HR column and eluted in TDP-43 CTF preparative SEC buffer; 20mM Tris-HCl pH 8, 4 M urea, 250 mM NaCl, 0.02% Triton X-100. Chromatography was carried out on the Akta Prime Plus system.

2.18.2 Preparative SEC of full length TDP-43

For investigation of SEC under denaturing conditions as a first purification step for full length TDP-43, denatured inclusion bodies in 50 mM HEPES, pH 8.0, 300 mM NaCl, 10 mM imidazole, 10 mM DTT, 8 M urea were centrifuged using a Beckman Avanti centrifuge with a JA25.50 rotor (25 000g, 16ºC, 30 min), then filtered using a 0.2 µm syringe filter. 3.5 ml of sample was applied to a HiPrep 26/60 Sephacryl S-300HR column on an Akta Prime Plus system and eluted in inclusion body denaturing buffer; 50 mM HEPES, pH 8.0, 300 mM NaCl, 10 mM imidazole, 10 mM DTT, 8 M urea.

For both preparative and analytical SEC of sample refolded by the dilution method described by Vega *et al*. (2019), at various stages of the purification and refolding process, sample was loaded onto a HiPrep 26/60 Sephacryl S-300HR column on an

Akta Prime Plus system, and eluted in 50 mM HEPES pH 8, 200 mM KCl, 100 mM sucrose, 0.025% Triton X-100, 100 mM imidazole, 0.1% PEG 3350.

2.18.3 Analytical SEC of refolded TDP-43

For analytical SEC of sample refolded by the dilution method described by Doke and Jha, (2022), sample was loaded onto a HiPrep 26/60 Sephacryl S-300HR column on an Akta Prime Plus system and eluted in 50 mM Tris-HCl pH 8, 10% sucrose, 10% glycerol, 500 mM NaCl, 2 mM DTT.

For analytical SEC of sample refolded by the SDS/MPD refolding system, both by urea dilution and direct refolding from the gel, sample was loaded onto a Superdex 200 10/300 gl column on an Akta Prime Plus system or a BioRad NGC chromatography system and eluted with one of the following buffers as the mobile phase depending on conditions being tested:

- 50 mM Tris-HCl, pH 8, 300 mM NaCl, 10% sucrose, 100 mM imidazole, 0.2% PEG-3350, 5 mM SDS, 1.5 M MPD, 5 mM DTT and 0.25% Triton X-100 (refolding into Tris-HCl buffer with Triton X-100).
- 50 mM Tris-HCl, pH 8, 300 mM NaCl, 10% sucrose, 100 mM imidazole, 0.2% PEG-3350, 5 mM SDS, 1.5 M MPD and 5 mM DTT (refolding into Tris-HCl buffer in the absence of Triton X-100).
- 20 mM HEPES, pH 8, 300 mM NaCl, 10% sucrose, 100 mM imidazole, 0.2% PEG-3350, 5 mM SDS, 1.5 M MPD, 5 mM DTT and 0.25% Triton X-100 (refolding into HEPES buffer for glutaraldehyde crosslinking compatibility).
- 50 mM $NaPO_4$, pH 8, 300 mM NaCl, 10% sucrose, 0.2% PEG-3350, 5 mM SDS, 1.5 M MPD and 0.25% Triton X-100 (refolding into CD spectroscopy compatible buffer).

- 50 mM Tris-HCL, pH 8, 300mM NaCl, 10% sucrose, 0.2% PEG-3350, 5 mM SDS, 1.5 M MPD and 0.25% Triton X-100 (for SEC following glutaraldehyde crosslinking of sample in CD spectroscopy buffer).

2.19 Elution of proteins from SDS-PAGE gels

Gel slices were crushed in a suitable volume of elution buffer to submerge the slices (50 mM Tris, pH 7.9, 100 nM EDTA, 1 mM DTT, 150 mM NaCl and 0.1% SDS), and incubated for at least 8 hours at room temperature on a rotator to allow for passive diffusion out of the crushed gel. The sample was then centrifuged (13,000 rpm, 2 min, Eppendorf 5424) to pellet the crushed gel and the supernatant was retrieved.

2.20 Acetone precipitation

Four volumes of -20ºC acetone was added to the protein sample, mixed well, then incubated at -20ºC for 1 hour. The sample was then centrifuged (15,000g, 4ºC, 10 min, Eppendorf 5424) and the supernatant disposed of. The pellet was air dried for 10 min, then resuspended in denaturing buffer (50 mM HEPES, pH 8.0, 300 mM NaCl, 300 mM imidazole, 10 mM DTT, 8 M urea).

2.21 Protein refolding by shock dilution

Denatured protein solution was added dropwise into refolding buffer (50 mM HEPES, pH 8.0, 9.6 mM NaCl, 10 mM KCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 1 M sucrose, 500 mM $_L$-arginine, 750 mM guanidine-HCl, 0.2% (w/v) PEG 3350, 0.25% (v/v) Triton X-100, 1 mM reduced glutathione, 100 nM glutathione disulfide and 1 tablet of EDTA-free Pierce protease inhibitor) (Vega *et al*. 2019) at 4ºC with gentle stirring. The volume of the

refolding buffer was such as to dilute the 8 M urea of the denaturing buffer to 600 mM. The resulting solution was left at 4ºC with gentle stirring for 48 hours.

An alternative dilution refolding solution was tested from Doke and Jha, (2022). This involved dilution from 8M urea down to 0.8 M urea into 50 mM Tris-HCl pH 8, 10% sucrose, 10% glycerol, 500 mM NaCl, 2 mM DTT at room temperature.

2.22 Artificial chaperone assisted protein refolding

Following preparative SDS-PAGE, elution, acetone precipitation and resuspension in denaturing buffer (50 mM HEPES, pH 8.0, 300 mM NaCl, 300 mM imidazole, 10 mM DTT, 8 M urea), sample was diluted into a urea dilution/SDS capturing buffer (50 mM Tris-HCl, pH 8, 300 mM KCl, 150 mM imidazole, 0.15% PEG-3350, 150 mM sucrose, 0.035% Triton X-100, 2 mM DTT and 3.09 mM SDS, with volumes prepared such that the final urea concentration was 600 mM and the SDS concentration 2.85 mM- this represents a 1 in 13.3 dilution of the sample in the SDS capturing buffer. Sample was incubated at room temperature on a slow speed rotamixer for 2 hours. Methyl-β-cyclodextrin (Sigma Aldrich) was then added to a final concentration of 10 mM, with volumes such that the concentration of SDS was reduced to 2 mM. The sample was incubated for a further 2 hours at room temperature on a slow speed rotamixer. An example of the volumes used is as follows: 385 µl of sample in denaturing buffer was diluted with 4.615 ml of urea dilution/SDS capturing buffer. For this 5 ml volume, 2.125 ml of 33.5 mM stock of methyl-β-cyclodextrin in MilliQ water was added.

2.23 SDS/MPD (2-Methyl-2,4-pentanediol) system protein refolding

Following preparative SDS-PAGE, elution, acetone precipitation and resuspension in denaturing buffer (50 mM HEPES, pH 8.0, 300 mM NaCl, 300 mM imidazole, 10 mM DTT, 8 M urea), sample was diluted into a urea dilution/SDS capturing buffer (50 mM Tris-HCl, pH 8, 300 mM NaCl, 100 mM imidazole, 0.1% PEG-3350, 10% sucrose, 0.25% Triton X-100, 2 mM DTT and 10-120 mM SDS) to reduce the urea to 800 mM and the sample was incubated at room temperature for 2 hours. An equal volume of MPD refolding buffer (50 mM Tris-HCl, pH 8, 300 mM NaCl, 100 mM imidazole, 0.1% PEG-3350, 10% sucrose, 0.25% Triton X-100, 2 mM DTT and 4 M MPD) was then added, so halving the concentrations of both the SDS and MPD in the buffers, and the sample was incubated at room temperature without mixing overnight.

2.24 Direct refolding from preparative SDS-PAGE gels

For refolding during elution from preparative SDS-PAGE gels, after gel staining, band excision and de-staining, the gel slices were crushed directly into an SDS/MPD based refolding buffer. The initial Tris-HCl based buffer contained 50 mM Tris-HCl, pH 8, 300 mM NaCl, 10% sucrose, 100 mM imidazole, 0.2% PEG-3350, 5 mM DTT, 100 nM EDTA, 0.25% Triton X-100, 5 mM SDS, 1.5 M MPD. For glutaraldehyde crosslinking a HEPES based buffer was used, containing 20 mM HEPES, pH 8, 300 mM NaCl, 10% sucrose, 100 mM imidazole, 0.2% PEG-3350, 5 mM DTT, 100 nM EDTA, 0.25% Triton X-100, 5 mM SDS, 1.5 M MPD. For direct refolding into a circular dichroism spectroscopy compatible buffer, a sodium phosphate-based buffer was used, containing 50 mM $NaPO_4$, pH 8, 150 mM NaF, 10% sucrose, 0.2% PEG-3350, 0.25% Triton X-100, 5 mM SDS, 1.5 M MPD. The crushed gel in buffer was left overnight on

a rotamix, and the sample was then retrieved and crushed gel removed by filtration with a 0.2 µm syringe filter.

2.25 Silver staining of SDS-PAGE gels

Silver staining of SDS-PAGE gels was carried out using the Pierce™ Silver Stain kit (ThermoScientific), following the manufacturer's instructions. Following electrophoresis, SDS-PAGE gels were washed twice in MilliQ water for 5 minutes, followed by fixation in 30% ethanol, 10% acetic acid for 2 washes of 15 minutes each. Gels were washed twice in 10% ethanol for 5 minutes, then twice for 5 minutes in MilliQ water. The gels were then sensitized for 1 minute in the supplied sensitizer solution (prepared as in the product manual), then washed twice for 1 minute each in MilliQ water. The gel was stained in prepared stain working solution for 30 minutes, washed twice for 20 seconds each in MilliQ, before developing of the gel in prepared developer working solution for 2-3 minutes, until desired band intensity was achieved. Staining was stopped by washing with 5% acetic acid for 10 minutes. Gels were imaged using a ChemiDoc™ XRS system (BioRad).

2.26 Glutaraldehyde crosslinking for determination of oligomeric states

For crosslinking, protein was refolded into a HEPES based refolding buffer (2.24) direct refolding) or a CD spectroscopy buffer (2.24). Glutaraldehyde crosslinking was carried out using a method modified from Fadouloglou *et al*., (2008). Per reaction, 40 µl 50% glutaraldehyde was diluted in 40 µl MilliQ water acidified with 2 µl 5 M HCl. The resulting 82 µl was dispensed into a well of a Corning 24 well cell culture plate. 15-20 µl of protein sample was pipetted onto an inverted glass cover slip (20x20 mm) producing a "hanging drop" and the coverslip was placed on top of the well containing

acidified 25% glutaraldehyde. The plate was incubated at room temperature for an appropriate length of time (10 minutes to 3 hours, depending on experiment), then the cover slip was retrieved from the top of the well, and the drop of sample collected by pipette. The crosslinking reaction was quenched by the addition of 4x SDS-PAGE sample buffer (5 µl) for preparation of samples for SDS-PAGE, or by the addition of 5 µl of 1 M Tris-HCl, pH 8, per 20 µl of sample retrieved in preparation for SEC. All steps were carried out in a fume hood.

2.27 PD-10 column desalting/buffer exchange

The top cap and sealed end of a PD-10 desalting column (Cytiva) were removed, the storage buffer poured off and the top filter removed with forceps. The column was placed in a 50 ml collection tube, filled with buffer into which the protein sample was to be exchanged and the solution allowed to completely enter the packed bed of the column. This was repeated three times to equilibrate the column. The column was filled a fifth time, and centrifuged at 1000g for 2 minutes (Rotanta 460R, Hettich Zentrifugen). The column was placed into a clean collection tube and 2.5 ml of sample was added to the centre of the packed bed, followed by centrifugation at 1000g for 2 minutes. The eluted sample was collected.

2.28 Circular dichroism (CD) spectroscopy

Protein was directly refolded from preparative SDS-PAGE gel extracts into a CD spectroscopy buffer (50 mM $NaPO_4$, pH 8, 150 mM NaF, 10% sucrose, 0.2% PEG-3350, 0.25% Triton X-100, 5 mM SDS, 1.5 M MPD) as in **2.24**. To remove glycine or Tris that may elute from the gel with the protein, sample was buffer exchanged into the same CD buffer with a PD-10 desalting column, using the spin protocol as in **2.27** to avoid sample dilution.

Far-UV CD spectroscopy was carried out using an AppliedPhotophysics Chirascan system. Sample was diluted as appropriate to reduce buffer noise. Scans were taken at 20 ºC from 260-180 nm with a 1 nm step size, with a 1 nm bandwidth, using a 1 mm pathlength cuvette with a sample volume of 300 µl. Triplicate scans were buffer blank corrected, and the BeStSel server was used to convert to Δε (per residue molar absorption units of circular dichroism measured in $M^{-1}cm^{-1}$) and fit and smooth the curve.

2.29 DNA binding assay by co-immunoprecipitation

12xTG or 12xAC repeat probes labelled with Cy3 were purchased from Integrated DNA Technologies. Probes were resuspended in MilliQ water. For binding of DNA to TDP-43, probes were added to protein samples in CD buffer at a final concentration of 0.1 µM in 100 µl final volume. Samples were incubated at 37ºC for 1 hour, then 10 µg of anti-TDP-43 monoclonal antibody (clone DB9) (Sigma-Aldrich) was added. The sample was incubated with rotation overnight at 4ºC. 50 µl of Protein G tagged Dynabeads (ThermoScientific) were added and the sample incubated for 2 hours at 4ºC. The Dynabeads were pelleted using a magnet and the supernatant removed and retained as the “supernatant fraction”. The beads were washed with 150 µl CD buffer, the beads pelleted again, and the supernatant removed. 150 µl of 8 M urea buffer was added (50 mM HEPES, pH 8.0, 300 mM NaCl, 10 mM imidazole, 10 mM DTT, 8 M urea) and the sample boiled at 100ºC for 8 minutes. The Dynabeads were pelleted using a magnet and the supernatant retained as the “pellet fraction”. Nylon membrane (Hybond XL) was washed with MilliQ water, then placed onto the membrane of a bottle top filter. Vacuum was applied to the filter, then 5 µl of supernatant or pellet fraction was spotted onto the filter. The spot was left for the liquid to draw through the nylon membrane,

then the spots were washed with 20 µl CD buffer. Cy3 fluorescence of the membranes was then imaged using the ChemiDoc™ XRS system (BioRad).

2.30 Determination of protein concentration by BCA assay

Protein concentration was determined by BCA assay, using the Pierce™ BCA assay kit. Routinely, the microplate procedure was used. For each assay, a standard curve was produced using the supplied albumin standard serially diluted as specified in the kit. Briefly, 25 µl of unknown sample and each dilution of the standard was added to wells of a Nunc™ 96 well clear flat bottom plate (269620) (ThermoScientific) in triplicate. Working reagent (200 µl, prepared as specified in the kit instructions), was added to each well, and the plate covered and incubated at 37ºC for 30 minutes. The plate was cooled to room temperature and absorbance was measured at 562 nm in a Tecan Infinite M200 Pro multimodal plate reader, or at 560 nm in a Biotek Synergy2 plate reader. The average absorbance measurement of the blank standard was subtracted from each individual standard and unknown replicate, and a standard curve plotted of the blank-corrected measurement against its concentration. The concentration of the unknown was determined via the standard curve.

2.31 Nanodrop concentration determination

A Nanodrop 2000C (Thermo Scientific) was used for estimating the concentrations of DNA or protein samples, using 260 nm absorbance for DNA (with 260/230 and 260/280 ratios for purity), and 280 nm absorbance for proteins (260/280 nm for nucleic acid contamination).

2.32 Custom peptides

Custom peptides were purchased from PeptideSynthetics (Peptide Protein Research Ltd, Fareham). Peptides were supplied lyophilised, and were resuspended in 50 mM HEPES, pH 8 (first and second -generation peptide aggregation inhibitors) or in DMSO (third-generation peptide aggregation inhibitors) to a final concentration of 1 mM.

2.33 Thioflavin T aggregation assays

For investigation of peptide self-aggregation, an aggregation mixture of 20 μM peptide, 15 μM Thioflavin T, 40 mM HEPES, pH 8.0, was prepared in a black, clear bottom 384-well plate (Corning), sealed with SealPlate® film (Sigma). For assays involving heparin, heparin sodium salt (from porcine intestinal mucosa) (Sigma) was added to a concentration of 5 μM. A Biotek Synergy2 plate reader was used, reading every 5 min for 24 hours (20 seconds agitation before reading), at a 37ºC incubation temperature, using an excitation filter of 440/20 nm, and an emission filter of 485/20 nm, to measure Thioflavin T fluorescence levels.

For TDP-43 C-terminal fragment aggregation assays, the purified TDP-43 CTF stock in 4 M urea SEC buffer (20 mM Tris-HCl pH 8, 4 M urea, 250 mM NaCl, 0.02% Triton X-100) was concentrated and adjusted to 54 μM. TDP-43 CTF sample (18.5 μl of 54 μM concentration) was used per 50 μl aggregation reaction, producing a 2.67 times dilution, giving 20 μM TDP-43 CTF in 7.49 mM Tris-HCl, 94 mM NaCl, 0.0075% Triton X-100 with the addition of 20 μM Thioflavin T. HEPES, (50 mM, pH 8), was used to dilute the sample to a 50 μl volume. Aggregation reactions were carried out in a black, clear bottom 384-well plate (Corning) sealed with SealPlate® film (Sigma) using a Tecan Infinite M200 Pro multimodal plate reader, reading every 10 minutes for varying time periods (shaking for 10 seconds with 3 mm amplitude before reading), at 25ºC

incubation temperature using an excitation wavelength of 440 nm, and an emission wavelength of 485 nm, to measure Thioflavin T fluorescence levels.

2.34 395 nm light absorbance assay

An aggregation mixture of 100 µM peptide, 40 mM HEPES, pH 8.0 and, for appropriate samples, 10 µM heparin sodium salt, were incubated for 24 hours at 30ºC with 1000 rpm shaking in an Eppendorf Thermomixer Comfort. Samples were pipetted into 1.5 ml plastic cuvettes, and the absorbance of the sample at 395 nm was determined using a Jenway 6300 spectrophotometer.

2.35 ProteoStat® Protein aggregation assay

ProteoStat® detection reagent (Enzo Life Sciences) was diluted as instructed by the product manual using the supplied assay buffer and MilliQ water. For each 50 µl aggregation reaction, 1 µl of reagent was added.

After optimisation, aggregation reactions consisted of 20 µM TDP-43 CTF protein in 1.5 M urea final concentration. For this, the purified TDP-43 CTF stock in 4 M urea SEC buffer (20 mM Tris-HCl, pH 8, 4 M urea, 250 mM NaCl, 0.02% Triton X-100) was concentrated and adjusted to 54 µM. CTF sample (18.5 µl of 54 µM concentration) was used for each 50 µl aggregation reaction, producing a 2.67 times dilution, giving 20 µM TDP-43 CTF in 7.49 mM Tris-HCl, 94 mM NaCl, 0.0075% Triton X-100. HEPES (50 mM, pH 8) was used to dilute the sample to a 50 µl volume. The volume used for dilution was adjusted for the presence or absence of inhibitor molecules at varying concentrations.

Aggregation reactions were carried out in a black, clear bottom 384 well plate (Corning) sealed with SealPlate® film (Sigma) using a Tecan Infinite M200 Pro multimodal plate

reader, reading every 10 minutes for varying time periods (shaking for 10 seconds with 3 mm amplitude before reading) at 25 ºC incubation temperature, using an excitation wavelength of 550 nm, and an emission wavelength of 600 nm, to measure ProteoStat® fluorescence levels.

2.36 Transmission electron microscopy (TEM)

Protein sample (10-15 µl) was applied to a formvar/carbon supported copper mesh grid (300 mesh) (Agar Scientific) and incubated for 2 minutes. Excess sample was wicked from the grid using Whatman filter paper, then 10-15 µl of 2% phosphotungstic acid (PTA) was applied. The grid was again incubated for 2 minutes, and the excess removed by wicking with filter paper.

Grids were imaged using the JEOL JEM 1010 transmission electron microscope using 80kV voltage and varying magnification as required.

2.37 AlphaFold

AlphaFold2 was used to predict protein structures and multimer formations using the ColabFold google notebooks as follows:

https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb

https://colab.research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb

Protein and peptide amino acid sequences were input into the ColabFold notebooks (Mirdita *et al*., 2022). Output structures were obtained and molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for

Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311 (Pettersen *et al*., 2004).

2.38 AGGRESCAN identification of aggregation hotspots

The web-based software AGGRESCAN (http://bioinf.uab.es/aggrescan/) (Conchillo-Solé *et al*., 2007), was used to identify aggregation hotspots in TDP-43 and in peptide inhibitor sequences. The TDP-43 amino acid sequence was input in FASTA format. The software carries out calculations based on aggregation-propensity values per amino acid, derived experimentally. The average aggregation propensity of a sliding-window is applied to the central amino acid of the window, producing an aggregation propensity for each amino acid in the particular sequence, which can be displayed graphically for the entire sequence. Regions with continual sequence with an aggregation propensity above a "hot spot threshold" can be considered a potential aggregation region. The "hot spot threshold" is defined as the average aggregation-propensity values per amino acid weighted by their frequencies in the SwissProt database (Conchillo-Solé *et al*., 2007).

2.39 Yeast transformation

Two *S. cerevisiae* strains were used in this project:

- *INVSc1: MATa his3D1 leu2 trp1-289 ura3-52 MAT his3D1 leu2 trp1-289 ura3-52*
- *Δerg6: MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*

Yeast cells were spread onto yeast extract peptone dextrose (YPD) (Sigma) agar plates from glycerol stocks using a sterile loop, and incubated at 30ºC for 48 hours. A colony was selected and used to inoculate 10 ml YPD liquid medium, which was incubated for

24 hours at 30ºC with 200 rpm shaking, or until the cell density was high enough to dilute to an OD600 of 0.4 in 50 ml YPD. After dilution, the culture was grown for a further 3 hours at 30ºC with 200 rpm shaking, then cells were pelleted (2500 rpm, 8 min, Eppendorf 5810R). Pelleted cells were resuspended in 10 mM Tris, pH 7.5, 1 mM EDTA (1xTE) (40 ml), pelleted again (2500 rpm, 8 min, Eppendorf 5810R) and resuspended in 2 ml of 100 mM lithium acetate, 5 mM Tris, pH 7.5, 0.5 mM EDTA. Cells were then incubated at room temperature for 10 mins, then 1 µg plasmid DNA and 100 µg denatured salmon sperm DNA were added to 100 µl of the cell suspension. 700 µl of 100 mM lithium acetate, 40% PEG-3350, 10 mM Tris, pH 7.5, 1 mM EDTA was added, the transformation was mixed well, then incubated at 30ºC for 30 mins. DMSO (88 µl) was added and the transformation was mixed well, before being heat shocked at 42ºC for 7 mins. Cells were then pelleted (13,000 rpm, 20 sec, Eppendorf 5424) and the supernatant removed. Pellets were resuspended in 1 ml 1 x TE, pelleted again (13,000 rpm, 20 sec, Eppendorf 5424), resuspended in 100 µl 1 x TE, and plated on a selective Synthetic Complete media without uracil (SC-U) (consisting of synthetic minimal media (Sigma) supplemented with yest synthetic drop out medium without uracil (Sigma)) agar plate. Plates were incubated for 48 hours, and transformant colonies were selected and used to inoculate 3 ml SC-U mini cultures, which were incubated for 24 hours at 30ºC with 200 rpm shaking. Glycerol stocks of transformed cultures were produced by the addition of sterile 50% glycerol (300 µl) to 700 µl of mini culture.

2.40 Yeast starter culture generation

For yeast spotting assays (**2.41**) protein expression (**2.43**), fluorescence microscopy (**2.44**), generation of growth curves (**2.45**) and disc growth assays (**2.46**), starter cultures were produced by a routine method. Transformed colonies picked from SC-U 2% glucose plates were inoculated into 5 ml SC-U media containing 2% glucose (for

spotting assays and disc growth assays) or 2% raffinose (for expression testing, fluorescence microscopy and generation of growth curves) and incubated overnight at 30ºC with shaking at 250 rpm.

2.41 Yeast spotting assay

Starter cultures were normalised to an OD600 of 0.7, then 1 in 10 serially diluted in filter sterilised water. 5 µl of each dilution was then spotted on an SC-U agar plate containing either 2% glucose (to repress protein expression) or 2% galactose (to induce protein expression). Plates were incubated for 2-3 days at 30ºC and imaged using the ChemiDoc™ XRS system (BioRad).

2.42 Yeast colony fluorescence

Colonies of yeast on agar plates expressing fluorescently labelled TDP-43 were imaged using the AxioZoom.V16 microscope. Colonies expressing GFP tagged protein were imaged using 395 nm excitation and 509 nm emission, while colonies expressing YFP tagged protein were imaged using 513 nm excitation and 527 nm emission. Images were taken using Zen imaging software (Zeiss) and prepared using Fiji (ImageJ).

2.43 Yeast protein expression

Starter cultures were diluted to an OD600 of 0.7 in 10 ml SC-U media containing 2% galactose, and incubated at 30ºC with 250 rpm shaking for 8 hours. Cells were harvested by centrifugation (2500 rpm, 8 min, Eppendorf 5810R) and lysed using YeastBuster™ (Novagen), according to the supplier's instructions. The soluble/insoluble fractions were separated by centrifugation (16,000 g, 20 min, 4ºC,

Eppendorf 5424), and the insoluble pellet fraction resuspended directly in 2x SDS-PAGE sample buffer (Thermo Scientific). Fractions were analysed for protein expression by SDS-PAGE and immunoblot.

2.44 Yeast fluorescence microscopy

Starter cultures were diluted to an OD600 of 0.7 in 10 ml SC-U media containing 2% galactose, and incubated at 30ºC with 250 rpm shaking for 6 hours. 333 µl of the resulting cultures were added to 666 µl of 100% ethanol, and incubated at room temperature for 1 hour. Cells were pelleted (2500 rpm, 2 min, Eppendorf 5415D), then resuspended in filter-sterilised PBS (1 ml). Cells were pelleted again (6000 rpm, 1 min, Eppendorf 5415D) and resuspended in 15 µl Vectashield with DAPI (Vectorlabs). The resulting suspension was incubated for 2 minutes at room temperature, then 10 µl was dropped onto clean microscope slides. Coverslips were applied, sealed with clear nail polish, and the slides were stored at 4ºC, protected from light until imaging. Images were taken using a Leica Stellaris 5 confocal microscope using the 100x oil immersion objective (1000x total magnification). For imaging of DAPI labelled DNA, excitation/emission wavelengths of 350/465 nm were used. For YFP tagged TDP-43, excitation/emission wavelengths of 513/530 nm were used. For GFP tagged TDP-43, excitation/emission wavelengths of 488/510 nm were used. Composite images were prepared from z-stacks using Fiji (ImageJ).

2.45 Yeast growth curves

Starter cultures were normalised to an OD600 of 0.1 in SC-U media containing no sugar source. 96 well, clear, flat-bottomed plates (ThermoScientific Nunc 269620) were sterilised by soaking in 100% ethanol for 5 minutes, then dried. The plate lid was treated for condensation reduction by filling with 4 ml of 0.05% Triton X-100 in 20%

ethanol and soaking for 15 seconds. Appropriate volume of carbohydrate source (raffinose for non-induction or galactose for induction) was added to each well, followed by inhibitor where appropriate, then normalised culture was added to bring the final volume in each well to 200 µl.

The plate was incubated for 72 hours at 30ºC in a Tecan Infinite M200 Pro multimodal plate reader, with linear shaking of 6mm amplitude and 250 rpm, taking readings of absorbance at 600 nm every 15 minutes.

2.46 Yeast disc growth assay

Starter cultures were diluted to an OD 0.7 in SC-U 2% glucose, then diluted 1 in 10, a dilution identified by the spotting assay for TDP-43 expressing strains to grow to good density on a glucose plate (TDP-43 expression), but to show minimal growth on a galactose (TDP-43 induction) plate, while there is no difference in the growth between the plates for a non-expression yeast strain. 750µl of this dilution was spread across glucose and galactose plates and allowed to dry. Sterile Whatman 6mm antibiotic assay discs were saturated with 20 µl of drug/carrier control, then pressed firmly onto the agar surface. Plates were incubated for 2-3 days at 30ºC.

3. Expression and Purification of untagged TDP-43

3.1 Introduction

Despite its first identification in 1995 (Ou *et al*. 1995), and subsequent identification in neuronal inclusions from ALS and FTLD patients in 2006 (Neumann *et al*., 2006), details of the structure and biochemical functions of TDP-43 have remained elusive. This is in part due to the difficulty in purifying recombinantly expressed TDP-43 in a soluble and folded state. Purification and characterisation of the N-terminal and RRM regions of TDP-43 has been successfully carried out, however potential structures of the aggregation prone full-length protein and C-terminal region have only recently been resolved (Wright et al., 2020; Jiang et al., 2016).

One of the reasons for this, as will be demonstrated in this chapter, is the formation of insoluble inclusion bodies of TDP-43 in the *E. coli* expression host. Purification of histidine-tagged protein from these inclusions by a denaturing and refolding method has been demonstrated by Vega *et al*. (2019) and Doke and Jha (2022). Some TDP-43 can be recovered from the cell soluble fraction and purified in a soluble manner, as demonstrated by Wright *et al*., (2020), however the addition of the detergent sodium lauroyl sarcosinate (sarkosyl), was still required to keep the protein soluble during purification steps. In all of these methods, TDP-43 was purified using a histidine affinity tag. Purification of TDP-43 utilising large affinity tags such as the maltose binding protein (MBP) has also been achieved, which produces soluble protein in the *E. coli* expression host, but cleavage of the tag results in instant aggregation, indicating the TDP-43 protein produced may be inherently unstable (Wang *et al*., 2018). TDP-43 with a cleavable SUMO tag also produces soluble protein in *E. coli*, and following tag cleavage soluble TDP-43 has been successfully purified (Kumar *et al*., 2023). This chapter reports the results of investigations of different methodologies used to purify full length untagged TDP-43 expressed in *E. coli*.

3.2 TDP-43 is expressed in *E. coli* as a predominantly insoluble protein

A plasmid (pKMH6) designed to express untagged full length TDP-43 has previously been generated, (Humphreys, 2013), and prior to initiating expression studies the plasmid was purified and the sequence of the cloned insert confirmed as wild type. Following purification of plasmids from bacterial glycerol stocks and sequencing, protein expression could be carried out. Previous work by Dr Benson's group (Humphreys, 2013) showed that untagged TDP-43 was insoluble in conventional *E. coli* expression strains such as BL21 cells, and so SHuffle T7 Express was chosen as an alternate expression strain, which is designed to aid in the correct folding of proteins (Lobstein *et al*., 2012). SHuffle T7 Express *E. coli* (New England BioLabs) were transformed with the pKMH6 plasmid, encoding recombinant full length, untagged TDP-43. Small scale protein expression using the pKMH6 transformed SHuffle was carried out at 30ºC and 25ºC, to confirm expression of the protein, with timepoint samples taken to monitor expression levels. As shown in Fig. 3.1, a band of approximately 43 kDa increased in prominence with time from induction, suggesting that expression of TDP-43 was successful at both 30ºC (lanes b-e) and 25ºC (lanes f-i). As expected, and as shown by the prominence of the bands on the gel, the concentration of TDP-43 was higher in the 30ºC expression timepoint samples than the 25ºC samples, due to the higher growth rate of the cells at the higher temperature. Immunoblot confirmed the expression of TDP-43 (Fig. 3.1 (b)).

For direct purification from cell extracts, soluble protein is required. The solubility of the TDP-43 expressed at these two temperatures was tested by small scale solubility analysis using BugBuster™ as a lysing agent followed by separation of the soluble and insoluble cell fractions by centrifugation. Fig. 3.2 shows the results of this solubility analysis. At both 30ºC and 25ºC, the majority of the TDP-43 expressed was isolated in the insoluble fraction of the cell lysate (comparison made between lanes c and d and

between lanes f and g). The reduction in temperature from 30ºC to 25ºC did not appear to have any significant effect on the proportion of soluble protein isolated.

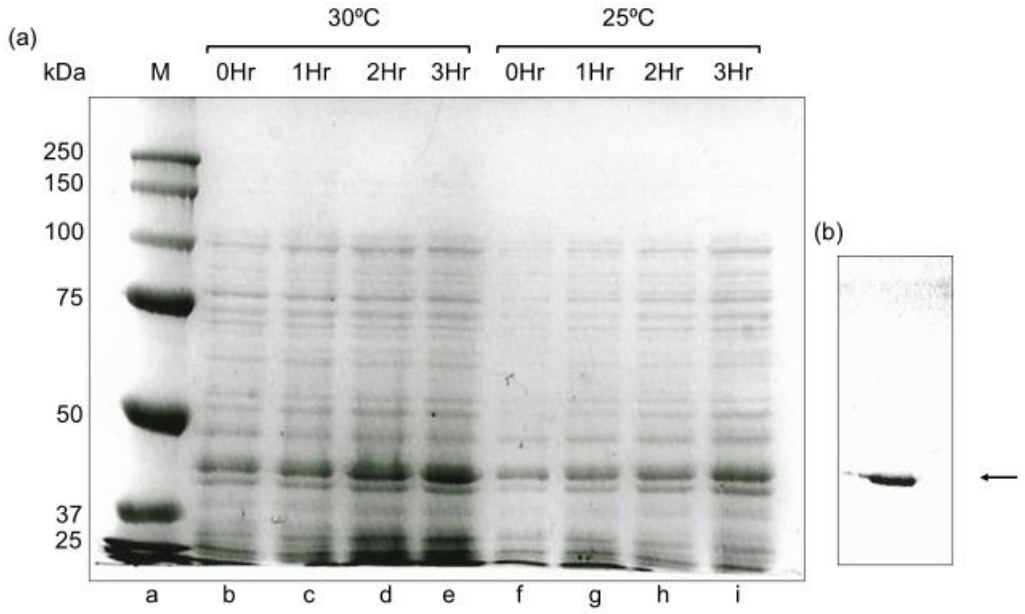


Figure 3.1. Time course expression of TDP-43 (**a**) SDS-PAGE gel showing whole cell extracts of timepoint expression samples at 30ºC and 25ºC. Timepoints show time from expression induction by IPTG addition. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); 0Hr, 1Hr, 2 Hr, 3Hr, whole cell extracts from expression culture taken at 0 hours, 1 hour, 2 hours or 3 hours after IPTG addition, for cultures at 30ºC (lanes b-e) or 25ºC (lanes f-i) (**b**) Immunoblot of a 30ºC 3 hr expression sample, probed with rabbit polyclonal anti-TDP-43 antibody (Invitrogen). Arrow indicates TDP-43 (~43 kDa).

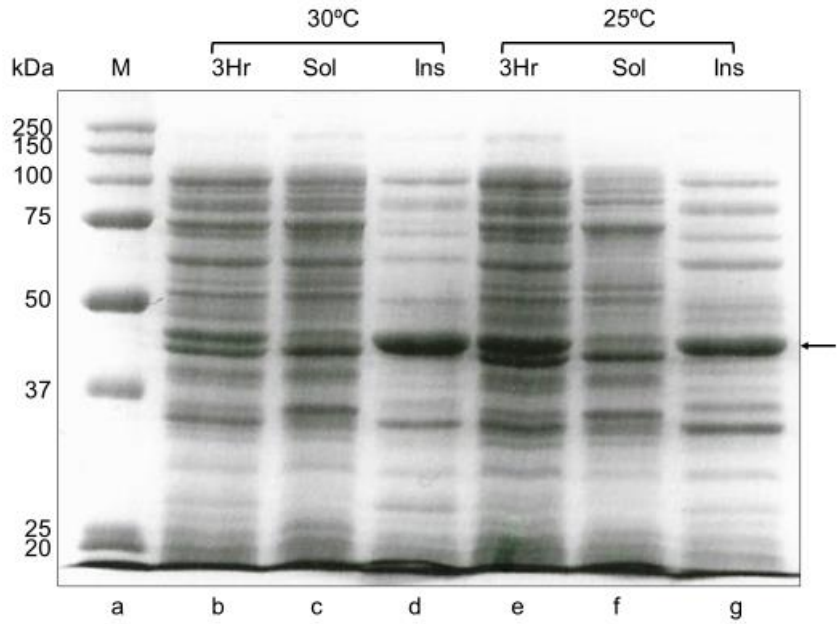


Figure 3.2. Solubility analysis of TDP-43 expressed at 30ºC and 25ºC. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); 3Hr, whole cell extract from a three-hour expression timepoint sample; Sol, cell soluble fraction; Ins, cell insoluble fraction. Samples from expression at 30ºC (lanes b-d) and 25ºC (lanes e-g). Arrow indicates TDP-43 (~43 kDa).

A further small-scale expression was then carried out at 16ºC to try to optimise the production of soluble TDP-43. Reducing temperature is a method that has been successfully used to reduce the formation of bacterial inclusion bodies for some insoluble recombinant proteins (Schein, 1989). The low temperature slows the growth rate and metabolic activity of the bacteria, and so an overnight (~15 hour) expression was tested. Small scale solubility analysis using BugBuster™ to lyse cells produced the results shown in Fig. 3.3. The majority of TDP-43 produced was again found in the insoluble fraction of the cells, shown by the more prominant 43 kDa band indicated in Fig. 3.3 lane d compared to lane c and so the reduced temperature again did not appear to have an effect on the proportion of insoluble protein produced. An alternative

approach was required to produce soluble protein with relevant biological activity that could be purified.

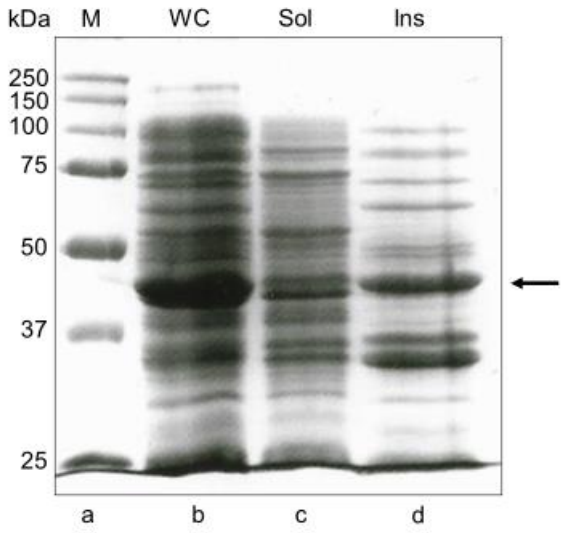


Figure 3.3. Solubility analysis of TDP-43 expressed overnight at 16ºC. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); WC (lane b), whole cell extract following ~15h expression; Sol (lane c), cell soluble fraction; Ins (lane d), cell insoluble fraction. Arrow indicates TDP-43 (~43kDa).

3.3 Purification of untagged TDP-43 from inclusion bodies following refolding of the inclusion body extract

Given that the expression of TDP-43 in SHuffle *E. coli* continued to produce insoluble inclusion bodies, a method was required to purify the protein from these inclusion bodies. Following large scale TDP-43 protein expression in pKMH6 transformed SHuffle *E. coli*, the bacterial cells were harvested by centrifugation and lysed by lysozyme and sonication. The inclusion bodies were isolated via washing and centrifugation steps, then denatured and solubilised. The sample was then concentrated using an Amicon® stirred cell, producing the sample in Fig. 3.4 (lane b), and "refolded" using the shock dilution method from Vega *et al*, (2019). Following this, the sample was again concentrated using an Amicon® stirred cell, producing a stock of soluble protein under native conditions that could be used to test purification methods. Fig. 3.4 (lane c) demonstrates the solubilised and refolded inclusion bodies prepared by this method. All purification methods tested under native conditions used sample prepared by this method as a starting material.

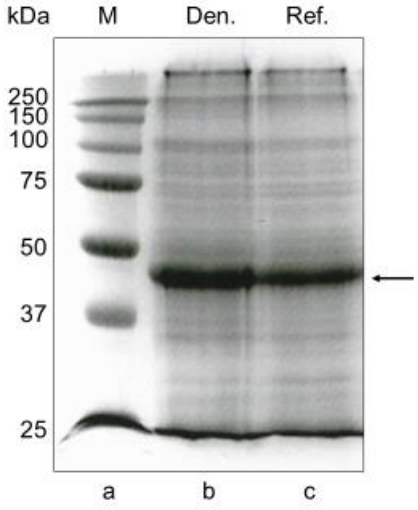


Figure 3.4. Denatured and "refolded" inclusion body preparations. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); Den, (lane b), solubilised, concentrated denatured inclusion bodies; Ref. (lane c), soluble sample following refolding by shock dilution and concentration using the Amicon® stirred cell. Arrow indicates TDP-43.

3.4 Hydroxyapatite chromatography for the purification of TDP-43

Once concentrated, the protein samples were prepared for hydroxyapatite chromatography. Humphreys (2013) demonstrated some success with TDP-43 purification by hydroxyapatite chromatography and so it was hypothesised that this method may be useful for purification from the inclusion body preparation. However, this protein was taken from the soluble cell fraction and so minimal protein was available for purification and furthermore, the resulting chromatography fractions were not able to be concentrated successfully without sample loss for use in further chromatography steps (Humphreys, 2013). Refolded protein sample as in Fig. 3.4 was dialysed into a hydroxyapatite chromatography binding buffer (5 mM KPO_4, pH 6.8, 10% (v/v) glycerol, 50 mM KCl, 500 nM DTT, 0.2% PEG 3350). Hydroxyapatite chromatography was then tested on an Akta Prime Plus chromatography system, using a Bio-Rad Bio-Scale Mini CHT ceramic hydroxyapatite column. This chromatography attempt was unsuccessful- all components of the sample injected onto the column eluted during injection or in the very first wash step, suggesting no binding of proteins to the column. There may be several explanations for this: previous freeze thaw steps affecting the protein sample in a way which removes its ability to bind hydroxyapatite, the inclusion of PEG-3350 in the hydroxyapatite buffer could have prevented hydroxyapatite binding, or the sample simply does not have affinity for the CHT hydroxyapatite column. Dialysis of protein into a hydroxyapatite buffer in the absence of PEG-3350 resulted in sample precipitation. This chromatography method did therefore not appear suitable for TDP-43 purification.

3.5 Purification of refolded TDP-43 using Immobilised Metal Affinity Chromatography (IMAC)

As an alternative purification method, immobilised metal affinity chromatography (IMAC) was tested. The two RRM domains of TDP-43 have been identified to bind to zinc ions (Garnier *et al*., 2017), and so IMAC was hypothesised to be a potentially useful purification step. Furthermore, Voráčková *et al*., (2011), demonstrated the purification of untagged zinc-finger proteins using IMAC, and found the success of the purification varied not only with the metal ion used, but also the type of chelating material used in the column.

Refolded inclusion body preparation was dialysed into a suitable IMAC binding buffer, and copper, nickel and zinc IDA, and nickel and zinc Sepharose columns were tested, with varying levels of success. The copper IDA column, as demonstrated by Fig. 3.5 resulted in a highly concentrated elution peak, containing TDP-43 along with many contaminants. This column appeared to demonstrate the highest level of TDP-43 binding and elution, but also the high level of contamination. Furthermore, there appeared to be some leaching of the copper ions, shown by a loss of blue colour from the column, and the most concentrated fractions (fractions 6 and 7, Fig. 3.5 lanes d and e) precipitated on storage at 4ºC within approximately 3 hours. Cu^{2+} IDA therefore appeared to not be a suitable IMAC medium.

Ni^{2+} IDA resulted in a small elution peak, and did not appear to demonstrate significant binding of TDP-43, with the majority clearly being present in the flow through, as shown by Fig. 3.6 (lane b). The resulting fractions did appear “cleaner” than those produced with the copper column, but were however very dilute and so a true comparison to Cu^{2+} IDA IMAC could not be made from these gels. The fractions could also not be successfully concentrated to better observe purity- after concentrating fractions 6-8

with an Amicon® Ultra centrifugal filter unit, no band could be observed on an SDS-PAGE gel, indicating the protein had been lost.

Chromatography using the HiTrap chelating IDA column chelated with Zn^{2+} ions was unsuccessful; there was no elution peak and protein bands could not be identified in any of the collected fractions, indicating there was no binding to the Zn^{2+} IDA column.

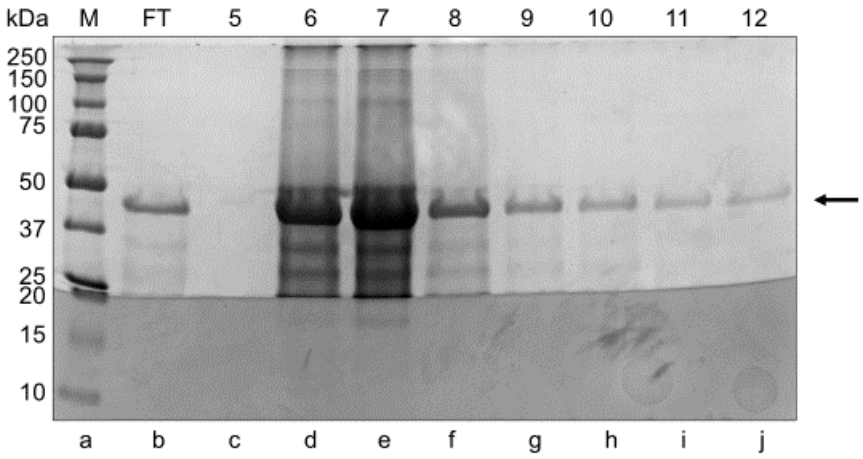


Figure 3.5. Cu^{2+} IDA IMAC. Gradient elution IMAC fractions of "refolded" inclusion body extract using a 1ml HiTrap chelating (IDA) column chelated with Cu^{2+} ions. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); FT (lane b), column flow through; 5-12 (lanes c-j), 1 ml elution fractions corresponding to the protein elution peak at 5-12ml. Arrow indicates TDP-43.

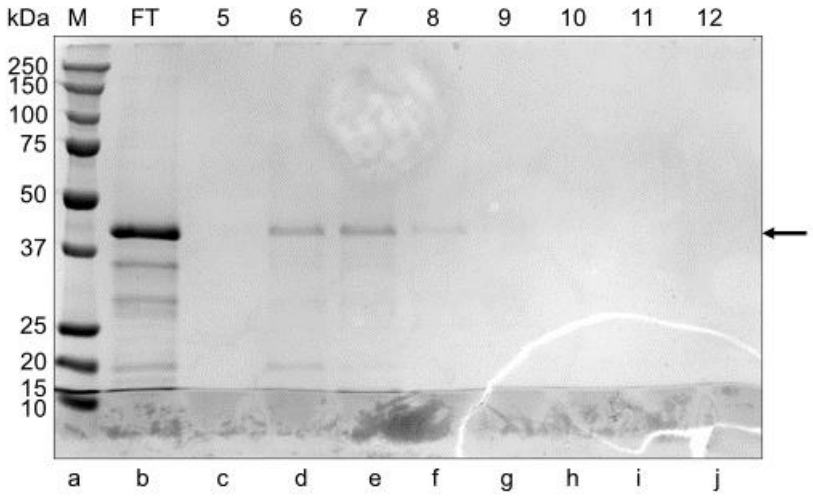


Figure 3.6. Ni^{2+} IDA IMAC. Gradient elution IMAC fractions of "refolded" inclusion body extract using a 1ml HiTrap chelating (IDA) column chelated with Ni^{2+} ions. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); FT (lane b), column flow through; 5-12 (lanes c-j), 1 ml elution fractions corresponding to the protein elution peak at 5-12ml. Arrow indicates TDP-43.

Sepharose fast flow from Cytiva was then investigated as the chelating media, using either Ni^{2+} ions as supplied in the HisTrap FF columns, or Zn^{2+} by stripping the HisTrap FF column and chelating it using $ZnCl_2$. The results of this are shown in Fig. 3.7 and 3.8. TDP-43 appeared to have a higher affinity for the metal ions when using the Sepharose chelating media compared to the IDA media, shown by the larger observed elution peak for an equal amount of protein sample applied. The elution peak from the Zn^{2+} column occurred earlier in the elution gradient than the peak from the Ni^{2+} column indicating potentially weaker binding (Fig. 3.8 lane e (fraction 6) for Zn^{2+} and Fig. 3.7 lane f (fraction 7) for Ni^{2+}). There did not appear to be a significant difference in the levels of remaining contaminants in the eluted fractions between the two ions.

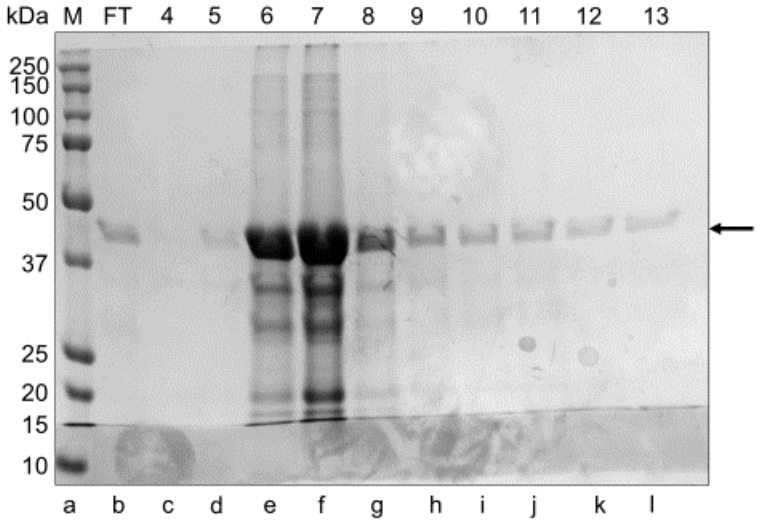


Figure 3.7. Ni^{2+} Sepharose IMAC. Gradient elution IMAC fractions of "refolded" inclusion body extract using a 1ml HisTrap FF (Sepharose fast flow) column chelated with Ni^{2+} ions. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); FT (lane b), column flow through; 4-13 (lanes c-l), 1 ml elution fractions corresponding to the protein elution peak at 4-13ml. Arrow indicates TDP-43.

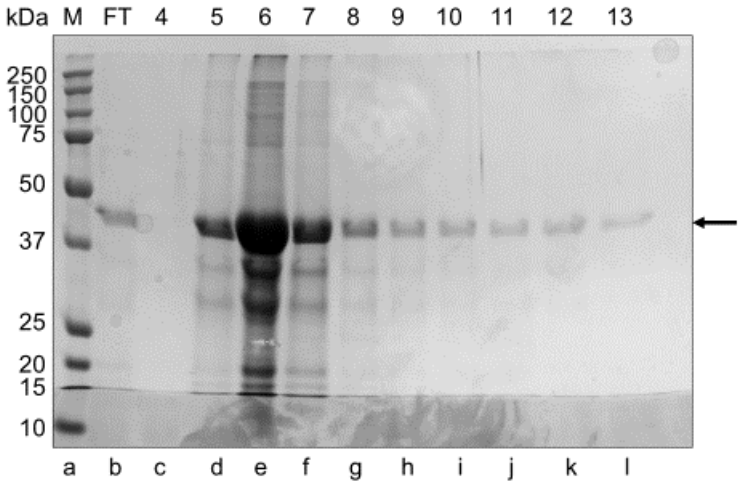


Figure 3.8. Zn^{2+} Sepharose IMAC. Gradient elution IMAC fractions of "refolded" inclusion body extract using a 1ml HisTrap FF (Sepharose fast flow) column chelated with Zn^{2+} ions. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); FT (lane b), column flow through; 4-13 (lanes c-l), 1 ml elution fractions corresponding to the protein elution peak at 4-13 ml. Arrow indicates TDP-43.

The refolded inclusion body preparation sample in the previous IMAC runs was loaded onto the column with no imidazole in the sample or equilibration buffers, before proceeding to gradient elution. A wash with low levels of imidazole before the gradient was applied was tested for the Ni^{2+} and Zn^{2+} Sepharose columns, however this did not result in any improvement in purity, with the TDP-43 and contaminating proteins eluting together at the same points in the wash and gradient. From these IMAC experiments, use of the Zn^{2+} Sepharose chromatography column with a gradient elution appeared to be most successful and so for further chromatographic steps, fractions produced from this method were taken forward as the protein pool.

3.6 Purification of refolded TDP-43 using heparin chromatography as a second chromatographic step

Regardless of improvements that can be made by altering IMAC conditions the majority of chromatographic purifications will likely need to utilise more than one column type. Heparin chromatography was tested as a second chromatographic purification step. The collected fractions from an IMAC run using a Zn^{2+} Sepharose column were pooled, dialysed into heparin binding buffer, and applied via superloop onto a 5 ml HiTrap Heparin HP column (Cytiva). Flow through material was collected and protein eluted via the application of a KCl gradient. The results of this chromatographic step are shown in Fig. 3.9. No flow through peak was observed, and no bands were observed in the flow through (Fig. 3.9 lane c), indicating apparent total binding of the protein in the sample to the heparin column. The protein eluted across seven 1 ml fractions upon application of the salt gradient, however there did not appear to be any increase in purity of the sample in any of these fractions compared to the pooled IMAC peak applied to the column.

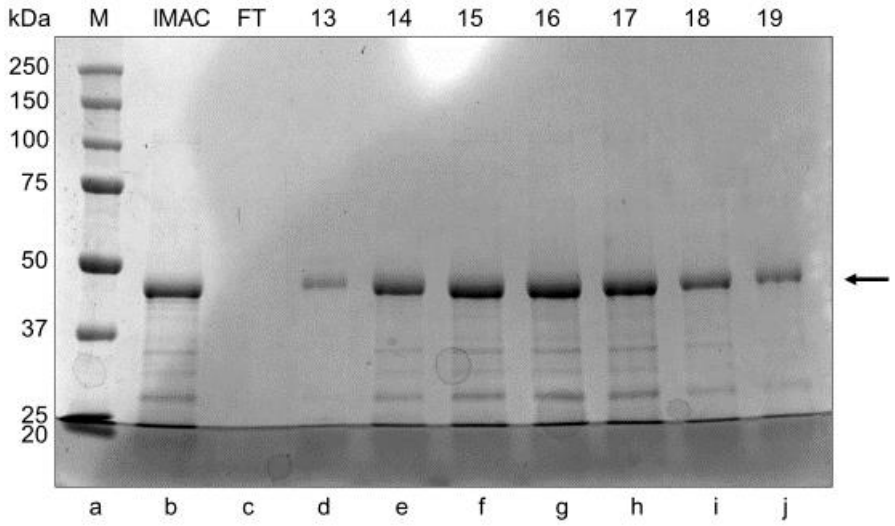


Figure 3.9. Heparin chromatography. Fractions obtained following heparin chromatography using a HiTrap Heparin HP column, after loading of pooled IMAC fractions. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); IMAC (lane b), pooled IMAC fractions loaded onto the column; FT, Flow through; 13-19 (c-j), 1 ml elution fractions corresponding to the protein elution peak at 13-19 ml. Arrow indicates TDP-43.

3.7 Preparative size exclusion chromatography of TDP-43 under native conditions

Given that heparin chromatography did not appear to improve purity, SEC using a HiPrep 26/60 Sephacryl S-300HR column was used, both as a potential purification method, and to identify if other protein species were interacting with TDP-43, resulting in the minimal effect of the chromatographic methods. Pooled elution fractions from heparin chromatography were applied to a HiPrep 26/60 Sephacryl S-300HR column and the protein eluted using the gel filtration buffer as described in Vega et al. (2019). The protein eluted in a single, 45 ml wide peak, centred at approximately 150 ml. 5 ml fractions collected showed this peak to contain TDP-43 along with other protein

contaminants present across the peak (Fig. 3.10). Given the centre of the elution volume of this peak, and the column calibration (Appendix G), it appears that the proteins in the sample were indeed interacting to form soluble oligomers approximately 235 kDa in size. These must have formed at some point in the upstream purification processes, which could explain the lack of effectiveness of any of the chromatographic techniques in separating TDP-43 from contaminating proteins.

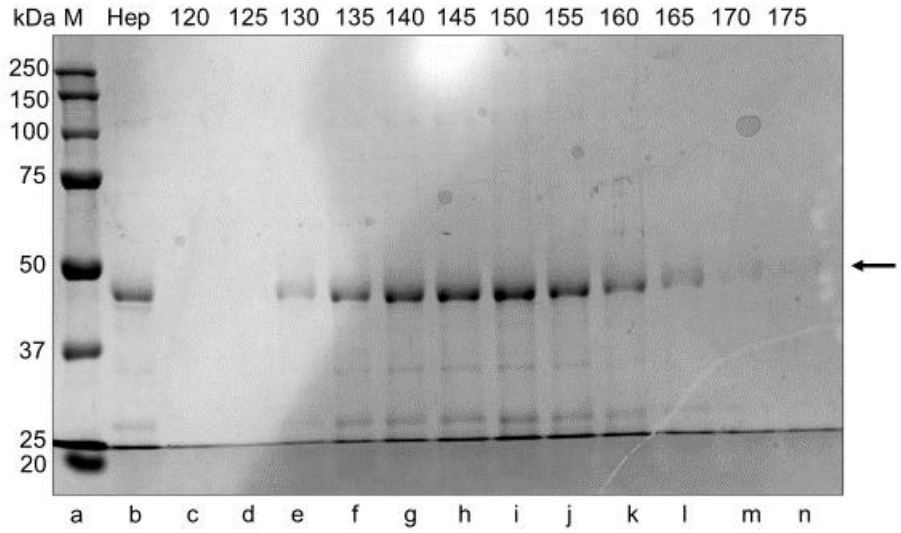


Figure 3.10. SEC of heparin chromatography elution fractions. Eluted fractions from a HiPrep 26/60 Sephacryl S-300HR column loaded with the pooled heparin chromatography fractions from Fig. 3.9. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); Hep (lane b), the pooled heparin chromatography fractions loaded onto the column; 120-175 (lanes c-n), 5 ml fractions corresponding to 120-175 ml elution volume from the column. Arrow indicates TDP-43.

3.8 Size exclusion chromatography of denatured inclusion bodies

Given that the refolding of an inclusion body preparation had resulted in the interaction of TDP-43 with contaminating proteins and the formation of soluble oligomers, it was hypothesised that using preparative SEC under denaturing conditions as an initial step to remove much larger and smaller contaminants could result in better folding without the soluble oligomer formation. Isolated inclusion bodies were solubilised in denaturing buffer (10 ml for 500 ml of bacterial culture), then applied to a HiPrep 26/60 Sephacryl S-300HR column (3-3.5 ml of sample applied per run) and eluted in the same denaturing buffer. TDP-43 eluted across the fractions shown in Fig. 3.11 with the bulk of the TDP-43 eluting across the central 15 ml (5ml fractions 135-145, Fig. 3.11 lanes e-g). Due to the denaturing conditions, the protein eluted from the column far earlier than a folded protein would, due to the increased hydrodynamic radius of a denatured protein compared to a refolded protein.

These three fractions (Fig. 3.11 lanes e-g) were taken, pooled and refolded using the refolding procedure from Vega *et al*. (2019). The resulting solution was concentrated down to 10 ml using an Amicon® stirred cell and was again applied to the HiPrep 26/60 Sephacryl S-300HR column using the Vega *et al*. (2019) gel filtration (GF) buffer as the mobile phase. The protein eluted in a single peak, centred on approximately 165 ml of elution volume. From column calibration in the GF buffer, this would correspond to a species with a molecular weight of approximately 150 kDa, and given the apparent contaminants being present across the peak, it appeared that once again, some form of soluble, contaminated oligomers had been produced. The contaminants shown in Fig. 3.12 appear different to those apparent in the 135-145 (Fig. 3.11 lanes e-g) fractions, and so these may also be degradation products occurring during refolding and concentration.

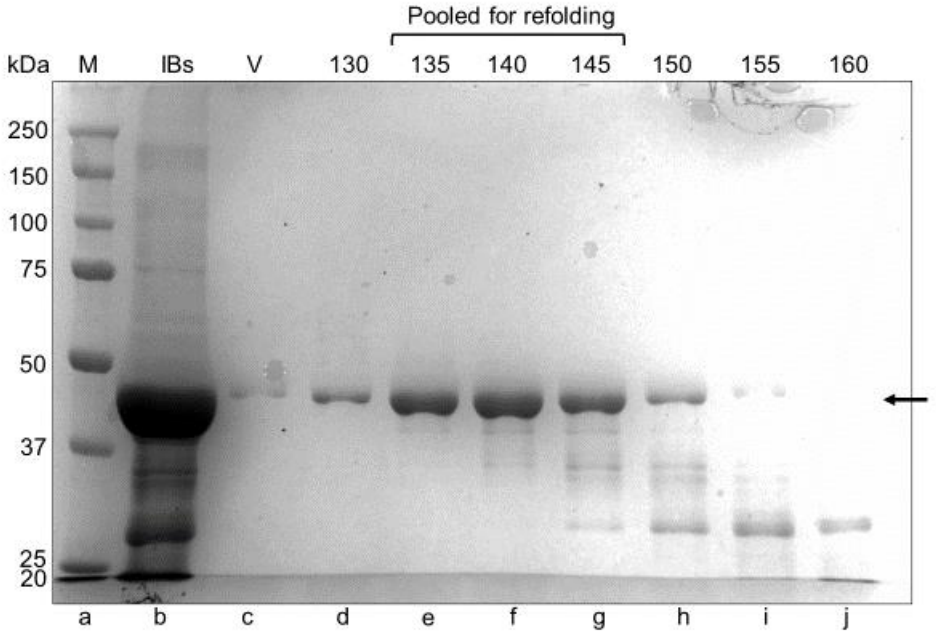


Figure 3.11. SEC of denatured inclusion bodies. TDP-43 containing fractions produced by preparative SEC of inclusion bodies using the HiPrep 26/60 Sephacryl S-300HR column under denaturing conditions. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); IBs (lane b), solubilised denatured inclusion bodies loaded onto the column; V (lane c), protein eluted in the void volume; 130-160 (lanes d-j), 5 ml fractions eluted between 130 and 160ml elution volume. Arrow indicates TDP-43.

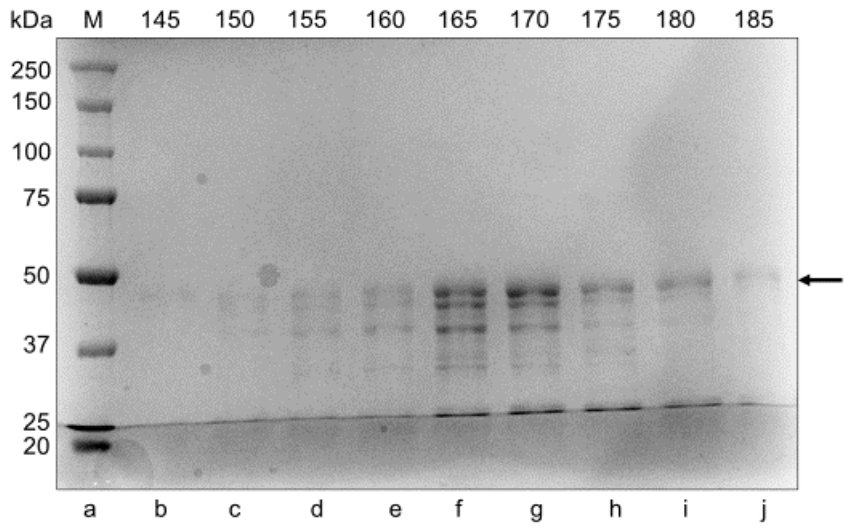


Figure 3.12. SEC of "refolded" TDP-43 fractions. TDP-43 containing fractions produced by HiPrep 26/60 Sephacryl S-300HR SEC of fractions 135-145 from Fig. 3.11 refolded by the Vega *et al*. (2019) refolding method. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); 145-185 (lanes b-j), 5 ml fractions eluted between 145 and 185ml elution volume. Arrow indicates TDP-43.

3.9 Purification of TDP-43 using anion exchange chromatography under denaturing conditions

Given the difficulty in maintaining solubility of TDP-43 samples after the refolding step, and the apparent formation of soluble aggregates of TDP-43 and contaminating species, options to purify the protein in its denatured state were considered.

Anion exchange chromatography under denaturing conditions was tested as a first purification step. This method relies on the charge of the protein in solution at a given pH, and is not reliant on the protein's structure, and so can be carried out under denaturing conditions. The inclusion body preparation was denatured in an 8 M urea anion exchange binding buffer, and applied to a HiTrap Q anion exchange column via

a superloop. The fractions eluted by salt gradient elution are shown in Fig. 3.13. TDP-43 bound to the anion exchange column, with minimal protein detected in the flow through compared to the sample loaded and the eluted fractions. The protein eluted across the entire salt gradient, and continued to elute in the final 1 M NaCl wash, so being diluted across 25 fractions of 1 ml each. The SDS-PAGE gels of the fractions show the chromatography had minimal effect on the purity of the protein sample, with contaminants eluting with TDP-43 across all fractions. Therefore, anion exchange chromatography under these conditions appeared unsuccessful as a purification method.

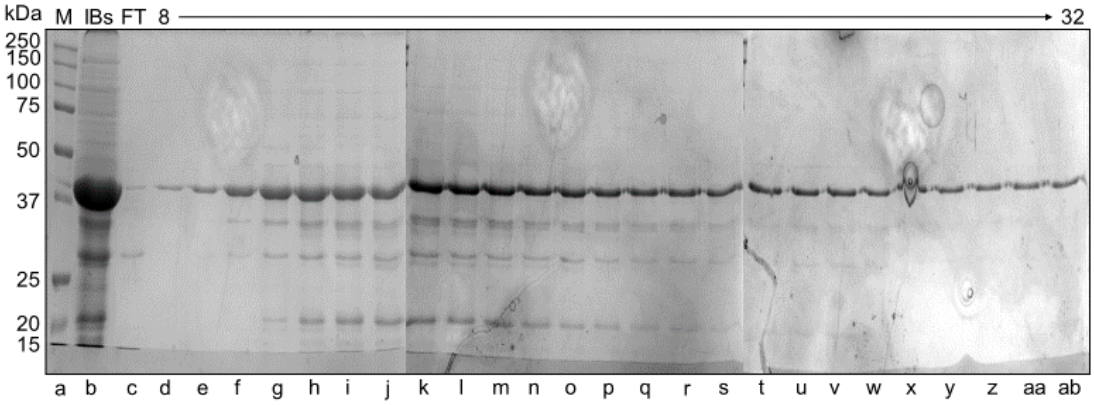


Figure 3.13. Anion exchange chromatography of denatured inclusion body extracts. Gradient elution peak produced by anion exchange of denatured inclusion bodies using a HiTrap Q anion exchange column. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); IBs (lane b), denatured inclusion body preparation; FT (lane c), column flow through; 8-32 (lanes d-ab), elution fractions 8-32 ml.

3.10 Purification of TDP-43 using IMAC under denaturing conditions

IMAC under denaturing conditions was trialled as a first purification step. Metal chelation does not necessarily require native conditions (unless the chelating region is produced by secondary structure) and so it was hypothesised that due to the chelating regions identified in TDP-43, and the prevalence of histidine, tryptophan and cysteine residues in the TDP-43 sequence, TDP-43 could be purified using its metal binding ability. As under native conditions, columns chelated with different metal ions were tested. Fig. 3.14 shows an SDS-PAGE gel of the fractions acquired when a Cu^{2+} IDA column was used for IMAC under 8 M urea conditions. As under native conditions, TDP-43 clearly bound to the column along with many contaminating proteins, and eluted in a sharp peak during the elution gradient. There was no visible improvement in the purity of the fractions compared to the denatured inclusion body sample loaded. A HisTrap FF column chelated with Ni^{2+} (Sepharose chelating media) was then tested (Fig. 3.15). As with the copper ion column, TDP-43 and many contaminants bound to the column and eluted together in a sharp peak, with no apparent increase in the purity of the sample. A Zn^{2+} Sepharose column was also tested, however a strong flow through peak and no elution peak were observed, with no protein detectable in any collected fractions of the elution gradient, indicating that under denaturing conditions the proteins did not appear to bind to the Zn^{2+} ion column.

The results presented here demonstrate that although TDP-43 binds to Cu^{2+} IDA and Ni^{2+} Sepharose columns under denaturing conditions, denaturing IMAC could not be used as a purification step as no separation of TDP-43 from contaminants was achieved by imidazole gradient elution.

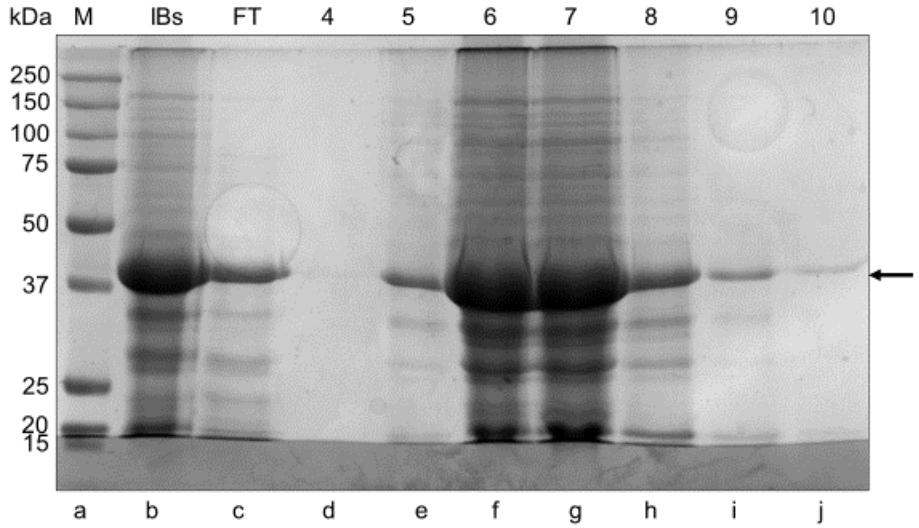


Figure 3.14. Cu^{2+} IMAC under denaturing conditions. IMAC fractions of denatured inclusion body extract using a 1ml HiTrap chelating (IDA) column chelated with Cu^{2+} ions. M, Precision Plus Protein™ All Blue protein standards (Bio-Rad); IBs, denatured inclusion body preparation; FT, column flow through; 4-10, elution fractions corresponding to the protein elution peak at 4-10 ml. Arrow indicates TDP-43.

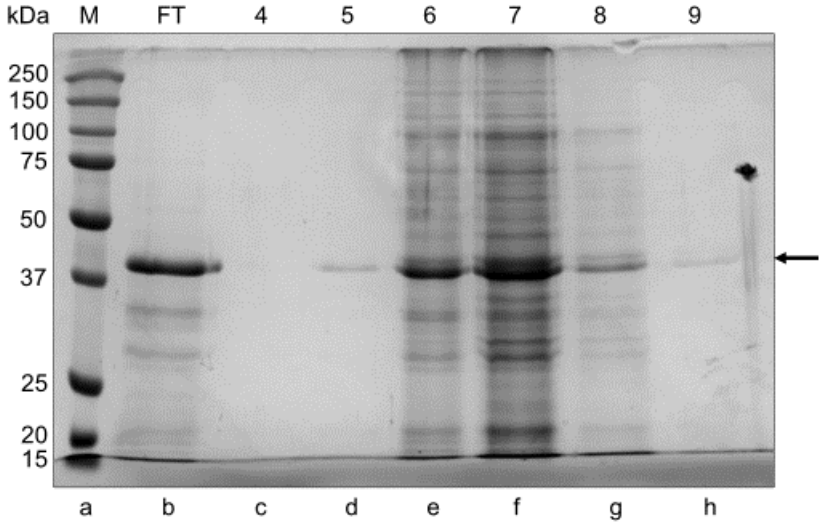


Figure 3.15. Ni^{2+} Sepharose IMAC under denaturing conditions. IMAC fractions of denatured inclusion body extract using a 1ml HisTrap column chelated with Ni^{2+} ions. M, Precision Plus Protein™ All Blue protein standards (Bio-Rad); FT, column flow through; 4-9, elution fractions corresponding to the protein elution peak at 4-9ml. Arrow indicates TDP-43.

3.11 Purification of TDP-43 using preparative SDS-PAGE

Since chromatographic approaches to purify TDP-43 either under native conditions from solubilised and refolded inclusion bodies or under denaturing conditions from denatured inclusion bodies had not been successful, an alternative approach was required. Preparative SDS-PAGE was investigated, first as a small-scale pilot, then scaled up to improve gel loading capacity and resolution for selection of TDP-43 from the gel. This method could potentially select TDP-43 based on its size with much higher resolution than SEC. To do this, inclusion bodies were isolated from large scale (1L) culture, denatured and concentrated. The Bio-Rad Mini-Protean system was used to carry out a pilot preparative SDS-PAGE. Nine wells of the standard size 10 well comb were taped over, and 350 µl of a 1:1 mix of concentrated denatured inclusion bodies and SDS-PAGE loading buffer were loaded into the large well created in the stacking gel. Following electrophoresis, the gel was negatively stained to visualise the protein bands. The prominent, 43 kDa band was cut from the gel and de-stained, before the gel slice was crushed in gel elution buffer and left overnight. The mixture was centrifuged to pellet the gel, and the supernatant retrieved. The supernatant was subjected to acetone precipitation, in order to remove SDS, and the precipitated protein was resuspended in 8 M urea denaturing buffer. The results of this pilot purification are shown in Fig. 3.16 (a), with a corresponding immunoblot shown in Fig. 3.16 (b), confirming the main band as TDP-43.

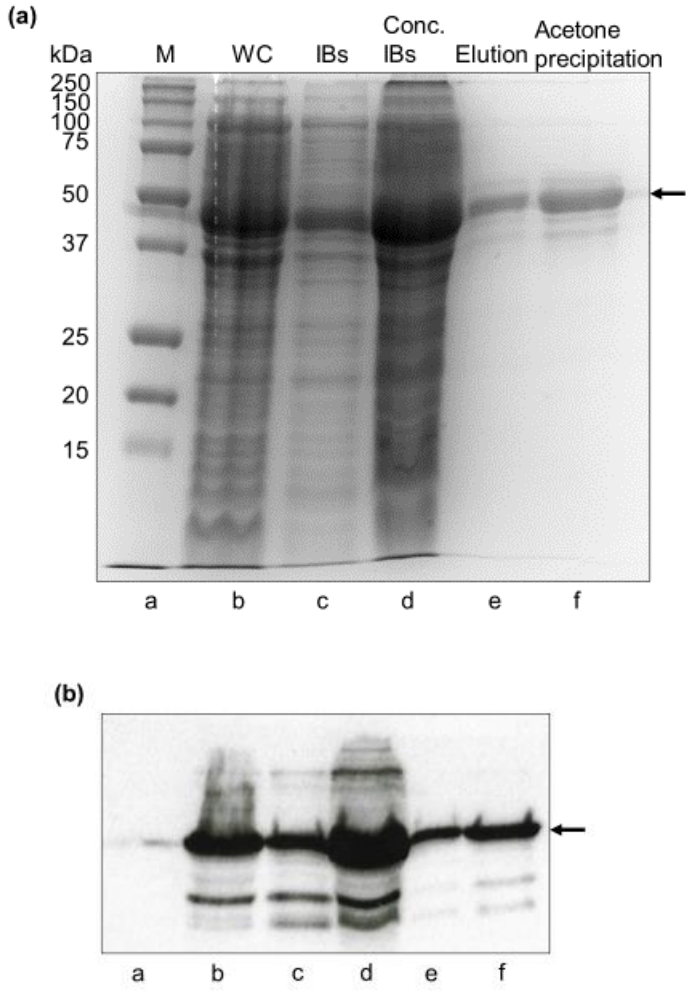


Figure 3.16. Small scale pilot preparative SDS-PAGE. (**a**) SDS-PAGE analysis and (**b**) corresponding immunoblot of the pilot preparative SDS-PAGE purification of TDP-43 using the Bio-Rad Mini-Protean system. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); WC (lane b), whole cell extract from a 1 ml aliquot of a 1L pKMH6 SHuffle culture following 3 hours of induction; IBs (lane c), diluted denatured and solubilised inclusion bodies; Conc. IBs (lane d), concentrated solubilised inclusion bodies; Elution (lane e), elution from gel slice; Acetone precipitation (lane f), sample following acetone precipitation and resuspension in denaturing buffer. Arrow indicates TDP-43. Immunoblot probed with rabbit polyclonal anti-TDP-43 antibody (Invitrogen).

The SDS-PAGE gel in Fig. 3.16 (a) and corresponding immunoblot in Fig. 3.16 (b) indicate that preparative SDS-PAGE could be a useful method to purify TDP-43 in its denatured state. As shown, TDP-43 was successfully eluted from the gel slice and resuspended after acetone precipitation. In this small-scale pilot experiment, approximately 200 µl of concentrated denatured inclusion bodies were loaded onto a Bio-Rad Mini-Protean gel and following gel extraction and acetone precipitation resulted in a 300 µl sample of approximately 2 mg/ml. As shown in Fig. 3.16 (a), some contaminating proteins still remain in the sample, which due to the small size of the mini gel and overloading (equivalent to lane d in Fig. 3.16(a)), were cut from the gel along with TDP-43 due to a lack of resolution. The immunoblot in Fig. 3.16 (b) indicates the presence of two smaller species that could be TDP-43 fragments formed either in the bacterial expression system or in the purification process to this point.

Given the limited loading capacity and resolution of the mini gel, it was clear that scale up to a larger SDS-PAGE system would be required to allow purification of a larger amount of TDP-43 with fewer contaminating proteins. To scale up, a Cambridge electrophoresis model EV200 electrophoresis tank was used. Resolving gels of 12% and 10% were tested for resolution of the TDP-43 band from surrounding proteins. As before, bands were excised from a negatively stained gel, and the proteins were eluted by passive elution. This was followed by acetone precipitation and resuspension in 8 M urea denaturing buffer. The results from these preliminary scale up experiments are shown in Fig. 3.17 (a) with a corresponding immunoblot in Fig. 3.17 (b).

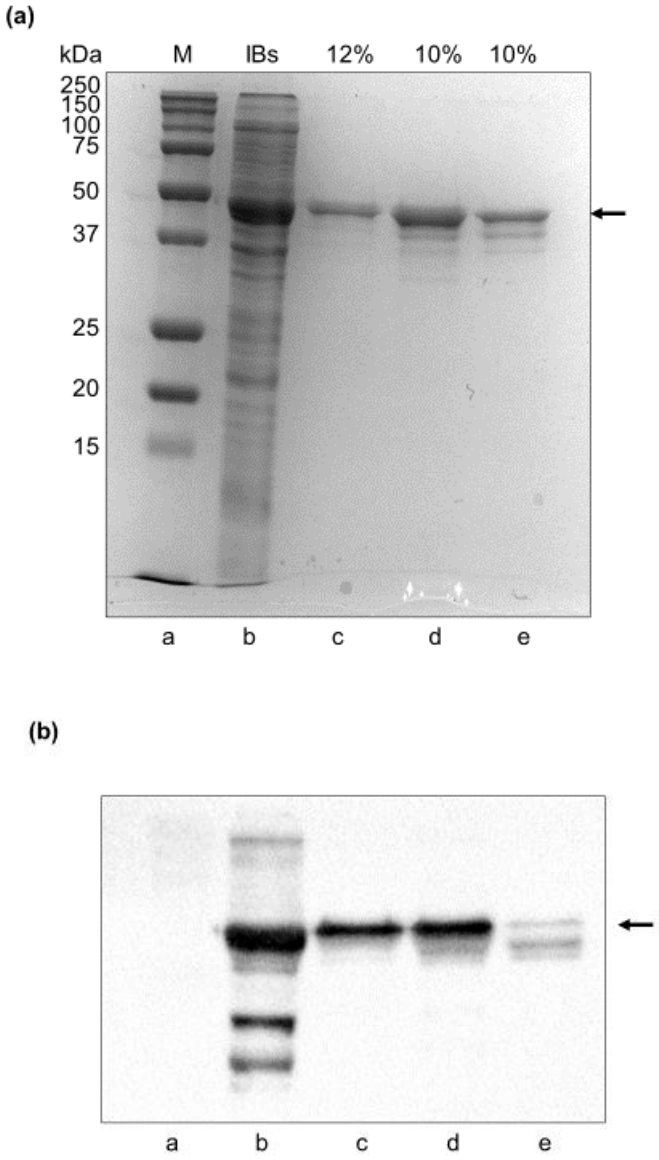


Figure 3.17. Analysis of large-scale preparative SDS-PAGE pilot purifications. (**a**) SDS-PAGE analysis and (**b**) corresponding immunoblot of three initial pilot experiments of large scale preparative SDS-PAGE. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); IBs (lane b), concentrated solubilised inclusion bodies; 12% (lane c), purification from a 12% gel; 10%, 2 trials (lanes d and e) of purification from 10% gels. Arrow indicates TDP-43. Immunoblot probed with rabbit polyclonal anti-TDP-43 antibody (Invitrogen).

These initial large-scale pilots of preparative SDS-PAGE produced slightly improved results compared to the small-scale pilot. In all three cases, no bands larger than TDP-43 were selected, and due to the increased loading capacity of the gel, 2 ml of sample (combined 1:1 with SDS-PAGE loading buffer) could be loaded on the gel without any significant overloading. However, further optimisation was clearly required; bands smaller than TDP-43 remained in the sample in all cases, and in the case of the second attempt using a 10% acrylamide concentration (Fig. 3.17 lane e), TDP-43 does not appear to be the main band selected, as demonstrated by the lack of a prominent band in the immunoblot (Fig. 3.19). A 10% gel did appear to improve separation of the 37-50 kDa region in the preparative gel, however following staining, the bands appeared more diffuse than in the 12% gel, leading to slightly increased difficulty in band selection. Furthermore, the thicker 2 mm gel did not stain as clearly with the negative stain as the smaller 1 mm thick pilot gel, again adding to difficulty in band selection. Further optimisation was therefore carried out.

To improve staining, 1 mm thick purification gels were run in the same way, allowing improved selection of bands, as shown by Fig. 3.18. A 10% acrylamide concentration was first used (Fig. 3.18 lane c), and as shown produced a cleaner final product after elution and acetone precipitation than the results shown in Fig. 3.17. Fig. 3.18 appears to show a single band corresponding to TDP-43, however, during optimisation of loading volumes for the 1 mm large scale gels, it was noticed that the band corresponding to TDP-43 in the inclusion body preparation was in fact a doublet, consisting of TDP-43 and another similarly sized protein. This doublet can just be discerned as a narrow band on the leading edge of the TDP-43 band in a concentrated preparative SDS-PAGE gel extract shown in Fig. 3.19 (lane b). In the stained preparative gel, resolution between the two bands of this doublet was not visible, therefore it is likely that the resulting sample shown in Fig. 3.18 (lane c) in fact contains two unresolved bands, rather than a single band of TDP-43. To try to improve resolution

to allow selection of the TDP-43 band only, 8%, 8-10% gradient, 12% and 15% gels were tested. These too did not show resolution of the doublet in the preparative gel, and appeared to produce the same result as the 10% gel. The only way found to resolve these bands was to reduce the loading volume applied to the gel, however this greatly limited the quantity of TDP-43 that could be purified.

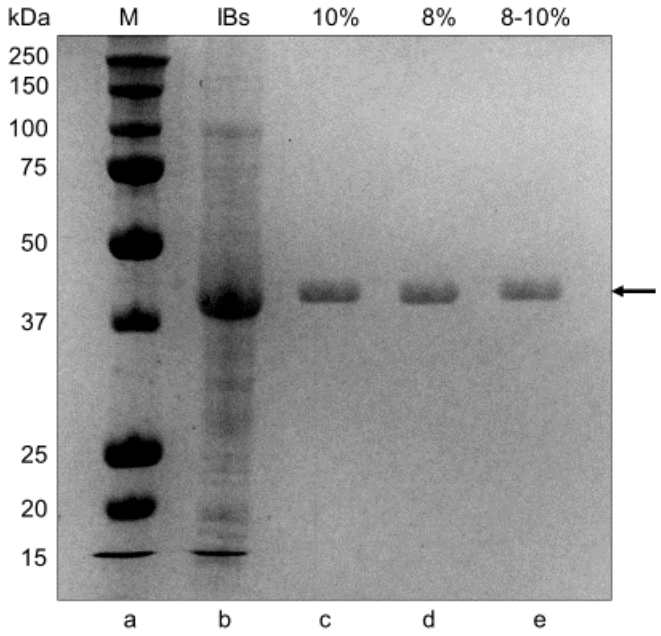


Figure 3.18. Analysis of purification of TDP-43 using 1 mm thick gel preparative SDS-PAGE gels. Three tests of preparative SDS-PAGE using 1 mm thick gels. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); IBs (lane b), concentrated solubilised inclusion bodies; 10% (lane c), purification from a 10% gel; 8% (lane d), purification from an 8% gel; 8-10% (lane e), purification from an 8-10% gradient gel. Arrow indicates TDP-43.

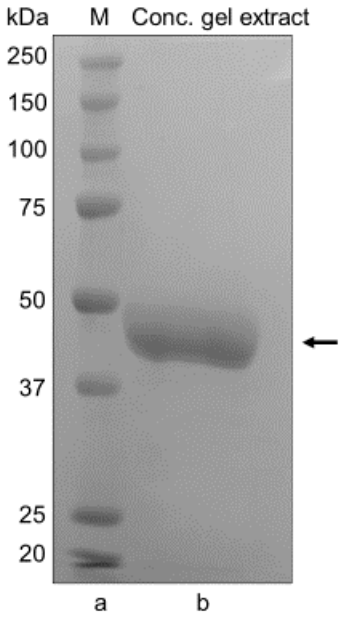


Figure 3.19. Concentrated contaminated preparative SDS-PAGE extract. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); Conc. Gel extract (lane b), concentrated preparative SDS-PAGE gel extract from a 10% preparative gel. Arrow indicates TDP-43.

Given the difficulty in removing the contaminating band by preparative SDS-PAGE, re-investigation of the material applied to the preparative gels was considered. Comparison was made between the concentrated denatured inclusion bodies prepared by the lysozyme/sonication method, and the insoluble cell fraction produced by solubility analysis using BugBuster™, as shown in Fig. 3.2. The separation of the soluble and insoluble fractions in these figures appears to show the doublet being split between the fractions, with TDP-43 being predominantly found in the insoluble fraction. Large scale 1 litre culture expression was carried out and the cells were harvested. Instead of lysis by lysozyme and sonication (as used to provide starting material for preparative SDS-PAGE purification described above), the cells were lysed using BugBuster™, and the inclusion bodies prepared using the Novagen® BugBuster™ inclusion body preparation protocol as supplied by the manufacturer. Samples of the

soluble cell fractions and concentrated denatured inclusion bodies produced by each method were compared and the results are shown in Fig. 3.20. The BugBuster™ protocol gave a cleaner inclusion body preparation (Fig. 3.20. (b) lane c) with a potential reduction in the contaminating band, and the soluble fraction appeared to have a higher total protein concentration, suggesting the lysis was better than the lysozyme/sonication method. However, when loading volumes were tested for preparative SDS-PAGE using the BugBuster™ inclusion body preparation, the contaminating doublet was seen to still be present in the sample.

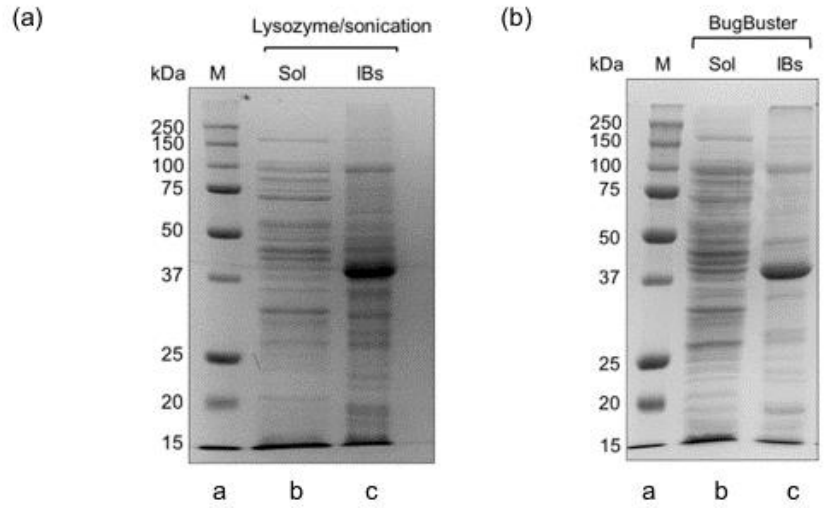


Figure 3.20. Comparison of inclusion body extract preparation methods. A comparison between the soluble and concentrated inclusion body fractions of bacterial cells lysed by lysozyme/sonication (**a**) and by BugBuster™ (**b**). M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); Sol (lane b), soluble fraction; IBs (lane c), concentrated denatured inclusion bodies.

Refolding of the sample containing TDP-43 and the single contaminant, using identical buffers and process to Vega *et al*. (2019), was then tested. If the TDP-43 protein was to refold to a dimer, as in Vega *et al*., (2019), then it could likely be separated from the contaminant by SEC. The sample was diluted in refolding buffer, to reduce the urea concentration to 600 mM, then the sample was concentrated using an Amicon® stirred cell. 3 ml of the resulting sample was then applied to a HiPrep 26/60 Sephacryl S-300HR SEC column. A single elution peak was observed as shown in Fig. 3.21, with peak elution occurring at approximately 165 ml as shown by the gel of fractions collected in Fig 3.22. This corresponded to a large molecular weight species of approximately 140 kDa as determined by the column calibration curve (Appendix G). Concentration of the peak fractions showed this peak to contain the two bands present from the preparative SDS-PAGE extract (Fig. 3.23), suggesting that some form of soluble, but oligomeric species consisting of TDP-43 and the contaminant had been formed by the refolding process. It therefore appeared that the presence of a single contaminating protein could result in the formation of an oligomeric species during the refolding process, and so pure TDP-43 was required to prevent interactions with the contaminant during refolding.

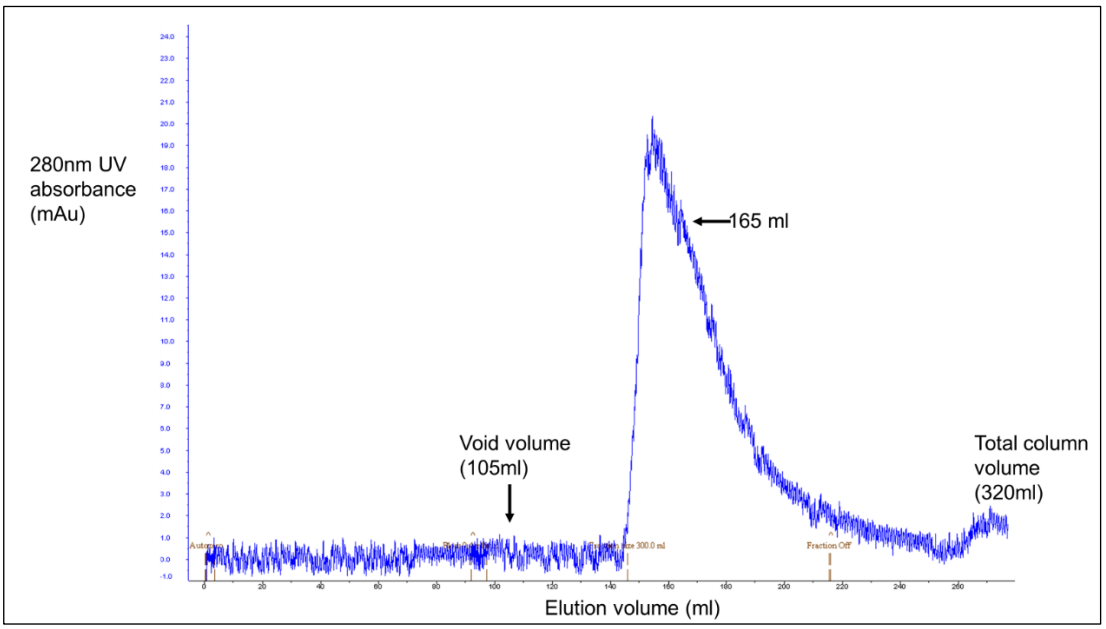


Figure 3.21. Analytical SEC of refolded contaminated TDP-43. UV trace of HiPrep 26/60 S-300HR SEC run, loaded with 3 ml of concentrated refolded TDP-43 purified by preparative SDS-PAGE.

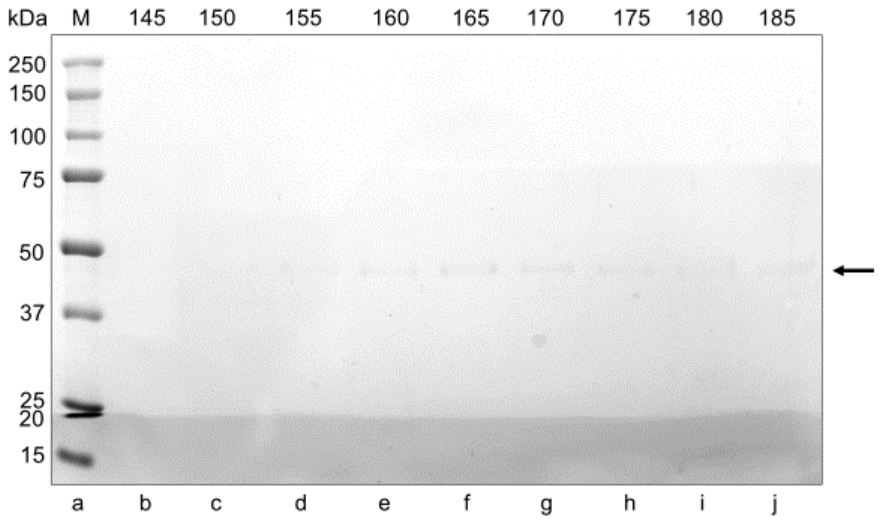


Figure 3.22. Fractions from the analytical SEC of refolded contaminated TDP-43. Peak elution fractions of the HiPrep 26/60 S-300HR SEC run, loaded with 3 ml of concentrated refolded TDP-43 purified by preparative SDS-PAGE. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); 145-185 (lanes b-j), fractions 145-185 ml of elution volume. Arrow indicates TDP-43.

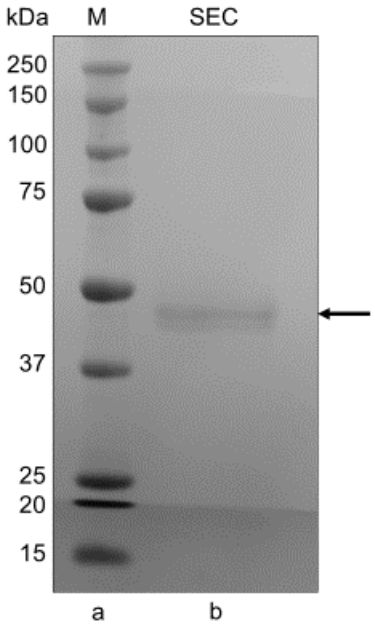


Figure 3.23. Concentrated peak fractions from analytical SEC of refolded contaminated TDP-43. The pooled and concentrated peak fractions from Fig. 3.22. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); SEC (lane b), pooled and concentrated SEC fractions. Arrow indicates TDP-43.

When carrying out preparative SDS-PAGE of inclusion body sample prepared using the BugBuster™ method, it was found that selection of the very centre of the TDP-43 band during band excision in preparative SDS-PAGE could produce pure protein in a single step, as shown in Fig. 3.24. There was no visible contaminant when these extract samples were repeatedly run on analytical SDS-PAGE gels with varying amounts of protein loaded. The concentration of the protein produced after elution, acetone precipitation and resuspension in denaturing buffer was estimated between 0.8 and 1 mg ml^{-1}. Therefore, sufficient pure protein for refolding could be produced with this method.

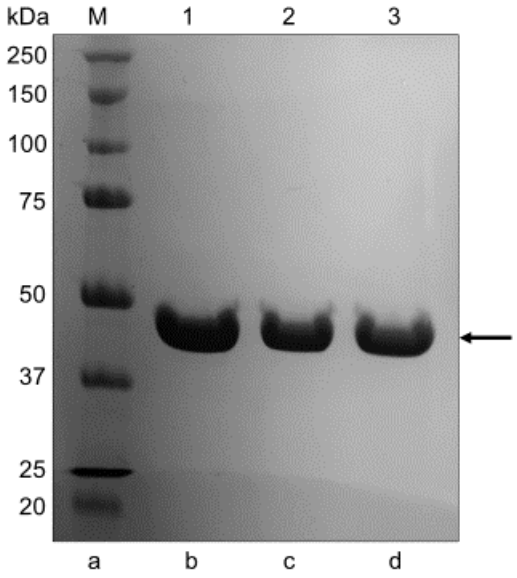


Figure 3.24. Preparative SDS-PAGE purified TDP-43. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); 1, 2 and 3 (lanes b-d), results of three separate preparative SDS-PAGE purifications of TDP-43. Arrow indicates TDP-43.

3.12 Conclusion

This chapter reports the development of a purification method for untagged, full-length TDP-43 produced in *E. coli*. As expected from the literature and previous work by Humphreys, (2013), TDP-43 was expressed as an insoluble inclusion body protein under the expression conditions used. TDP-43 was therefore purified from the inclusion bodies.

Two approaches were taken- denaturation of the inclusion bodies followed by refolding of the protein before purification, or purification under denaturing conditions, producing pure protein for subsequent refolding. The initial approach of refolding the inclusion body mixture produced contaminated oligomers, indicating that TDP-43 interacts with other inclusion body proteins if refolded in their presence. This appears to have led to the lack of success of the multiple chromatographic techniques utilised downstream of this step to purify TDP-43.

Purification was therefore carried out under denaturing conditions. Following unsuccessful utilisation of chromatographic methods, preparative SDS-PAGE was found to be a method that could be used to reliably purify TDP-43 in a single step under denaturing conditions.

4. Refolding of Purified TDP-43

4.1 Introduction

The previous chapter demonstrated that pure, untagged full length TDP-43 could be produced by preparative SDS-PAGE. This method however produces protein in a denatured state, denatured by either the SDS whilst undergoing electrophoresis, or by urea following elution and acetone precipitation. For the purified protein to be most useful in characterisation of its oligomeric state, function and aggregation it needs to be refolded into a native structure.

To do this, techniques can be used that remove, or modulate the effect of, the denaturant. For chemical denaturants such as urea and guanidine hydrochloride, this can be done by simply diluting the denaturant to a point it is no longer effective, or removing it from the protein environment by buffer exchange by dialysis or chromatographic techniques in which the protein is exchanged into a different mobile phase on column. As well as these techniques, denaturing detergents such as SDS can be removed from the protein environment by molecules such as cyclodextrins, modulated by cosolvents, or out competed from binding to the protein by non-denaturing detergents at higher concentrations.

For TDP-43 specifically, methods in the literature have utilised chemical denaturing by urea or guanidine to allow purification of tagged TDP-43 recombinant proteins. The proteins have then been refolded by dilution of urea into buffers containing stabilising salts and crowding agents, producing soluble dimeric or monomeric/oligomeric states respectively, depending on buffer conditions, which may have DNA binding activity (Vega *et al*., 2019; Doke and Jha, 2022). Dilution of a guanidine-HCl denatured construct into a non-denaturing buffer by Furukawa *et al*. (2011) produced a small amount of soluble protein that regains DNA binding activity (though its oligomeric state was uncharacterised), but this method resulted in the majority of the purified protein

precipitating on dilution. This chapter will present experiments designed to optimise refolding of the pure, denatured, untagged TDP-43 produced by preparative SDS-PAGE as described in the previous chapter, with an aim to produce material in a monomeric/dimeric or low order oligomeric state that could in future be used for aggregation assays.

4.2 Refolding TDP-43 by shock dilution produces higher order oligomers

Refolding by shock dilution of the preparative SDS-PAGE purified protein was first investigated using the buffers and method described in Vega *et al*. (2019). After refolding, the protein was applied to the HiPrep 26/60 Sephacryl S-300HR SEC column using the gel filtration buffer described by Vega *et al.* (2019) as the running buffer. The protein eluted as a high molecular weight species, estimated at 230 kDa from the calibration curve, as shown in Fig. 4.1. Pooling and concentration of the peak fractions (elution volume 144 ml to 208 ml) showed this peak to consist of pure TDP-43 (Fig, 4.2), suggesting that a form of soluble, higher order oligomers had been produced by this refolding process. The resulting sample did not appear to precipitate, suggesting that these oligomers were stable in the gel filtration buffer. However, lower order (monomer/dimer) oligomeric states were desired, as these may be more useful in aggregation assays than a higher order species that could already be in the aggregation process.

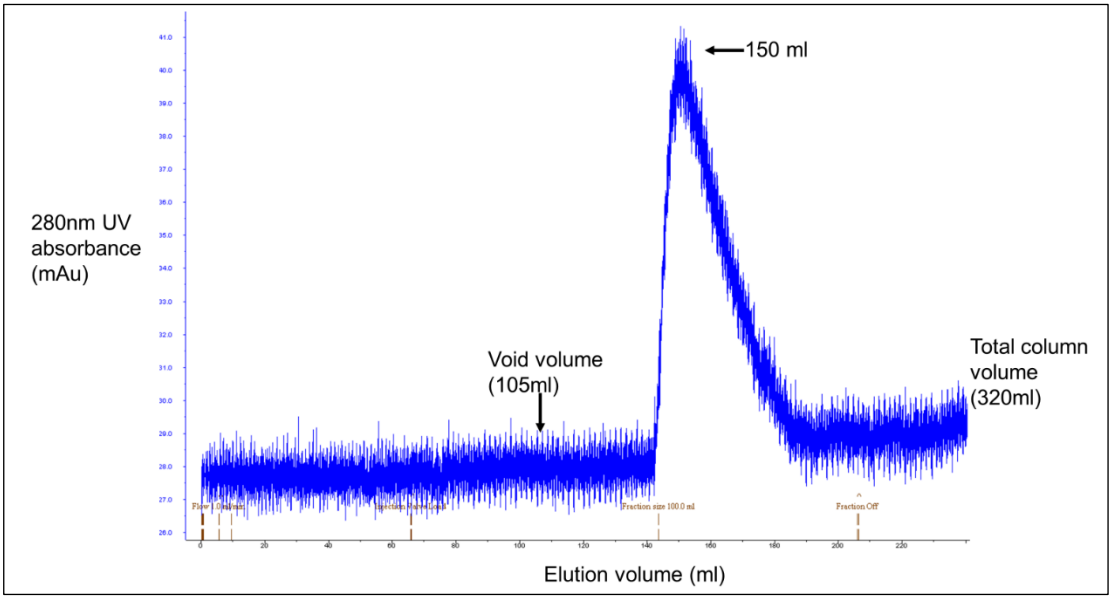


Figure 4.1. Analytical SEC of TDP-43 refolded by shock dilution method. Chromatography trace of absorbance at 280 nm with elution volume of TDP-43 refolded by shock-dilution method and applied to a HiPrep 26/60 Sephacryl S-300HR SEC column.

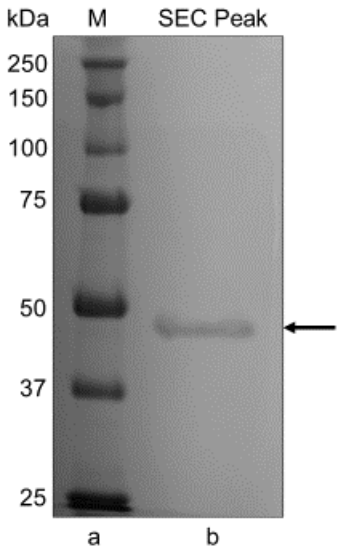


Figure 4.2. Pooled peak fractions from Analytical SEC of TDP-43 refolded by Vega *et al*. (2019) shock dilution method The pooled and concentrated fractions collected for 144-208 ml elution volume in Fig. 4.1. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); SEC Peak (lane b), pooled and concentrated eluted fractions. Arrow indicates TDP-43.

More recently, Doke and Jha (2022) demonstrated refolding of a His-tagged TDP-43 recombinant protein that had been purified under denaturing conditions to a monomeric state, using a buffer containing high levels of crowding agents and NaCl. A monomeric state, if achievable, would be preferable than the higher order oligomers previously obtained. Therefore, refolding by shock dilution with their described method was tested. On application of the resulting sample to the HiPrep 26/60 Sephacryl S-300HR SEC column, using the refolding buffer as the mobile phase, no clear elution peak was observed, as shown in Fig. 4.3. This reoccurred on application of higher sample volume and concentration to the column. Application of these samples resulted in an increase in column backpressure, suggesting on column precipitation of the protein. Cleaning of the column with 0.5 M NaOH resulted in a large elution peak (Fig. 4.4), and resulted in a reduction of the column pressure. This indicated that the protein applied had precipitated inside the column, and so the sample purified by preparative SDS-PAGE did not appear to be stable in this refolding buffer.

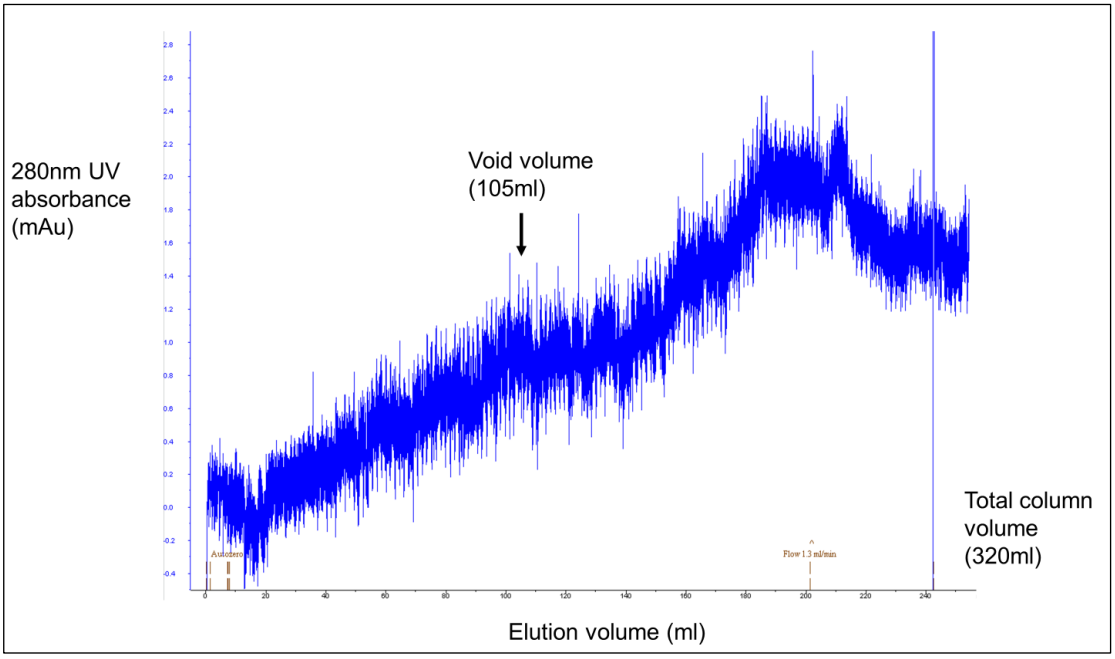


Figure 4.3. Analytical SEC of TDP-43 refolded by Doke and Jha (2022) shock dilution method. 280 nm UV absorbance with elution volume of Doke and Jha (2022) shock-dilution method refolded TDP-43 sample applied to a HiPrep 26/60 Sephacryl S-300HR SEC column.

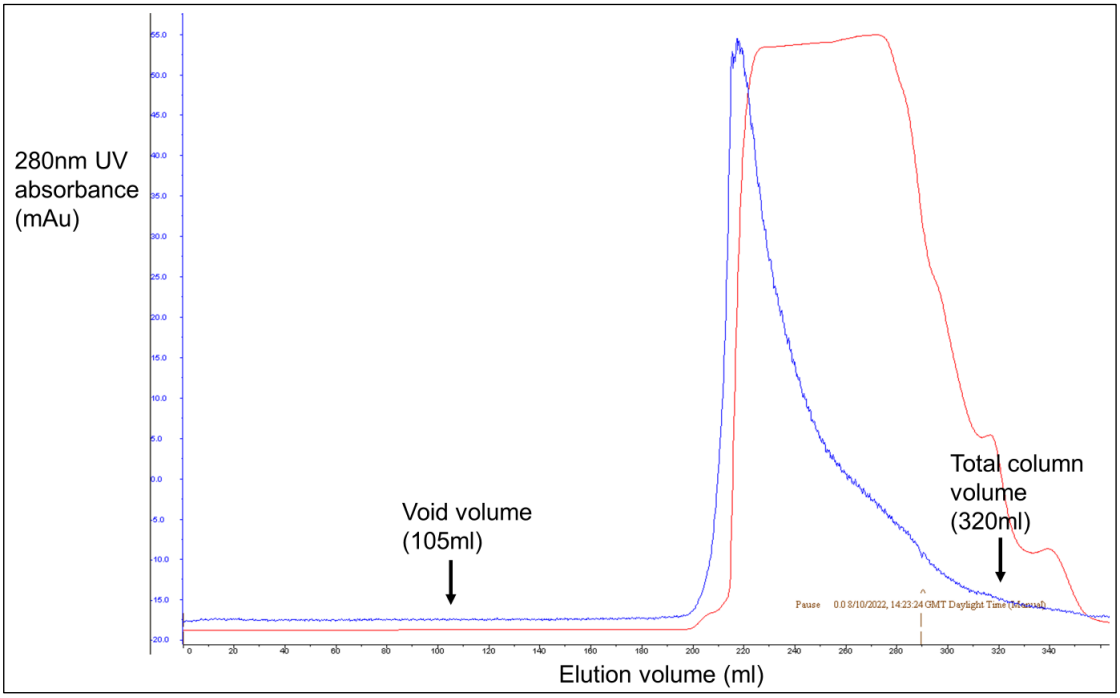


Figure 4.4. Column cleaning following application of TDP-43 refolded by Doke and Jha (2022) shock dilution method. Chromatography trace of the cleaning of the HiPrep 26/60 Sephacryl S-300HR SEC column using NaOH after analytical SEC of Doke and Jha (2022) shock-dilution method refolded TDP-43 sample. Note the sharp 280 nm UV absorbance peak (blue) corresponding with the buffer change to NaOH shown by the conductivity change (red). This UV absorbance peak suggests sudden elution of protein bound inside the column following exposure to NaOH.

4.3 Artificial chaperone assisted refolding of TDP-43

Since the shock dilution methods did not lead to the refolding of TDP-43 to a monomeric/dimeric state, alternative methods for protein refolding were therefore tested. Rozema and Gellman (1995; 1996; 1996), and Hanson and Gellman (1998), describe a method by which a two-step refolding method, known as artificial chaperone assisted refolding, was carried out for proteins including carbonic anhydrase B (CAB) and lysozyme. This method proceeds by dilution of the urea in which the protein is solubilised into a buffer containing detergent, which “captures” the protein in a

denatured state to retain its solubility. An “artificial chaperone”, the cyclic sugar molecule cyclodextrin, is then applied, which by its high affinity for detergent can “strip” the detergent from the protein allowing it to fold. This method was therefore tested, using SDS and methyl-β-cyclodextrin as the detergent and cyclodextrin respectively as described in Hanson and Gellman, (1998) for the refolding of CAB. The pure protein sample was first diluted into the SDS-containing dilution buffer, then gently mixed on a rotamixer for two hours. The solution at this point remained clear and free of any precipitate or aggregate. Methyl-β-cyclodextrin was then added, and the sample returned to the rotamixer. After one hour of the two described in the method (Hanson and Gellman, 1998), visible large aggregates and precipitation had occurred. Continuation to the full two hours resulted in an increase in this precipitation. It therefore appeared that the stripping of the SDS resulted in sample precipitation. To determine whether any proportion of the sample remained soluble and could have potentially folded, SEC was carried out. The sample was filtered using a 0.2 μm syringe filter and applied to the HiPrep 26/60 Sephacryl S-300HR SEC column, using the refolding buffer without SDS or methyl-β-cyclodextrin as the mobile phase. No elution peak was observed, and again, a slight increase in column pressure occurred, suggesting further on column precipitation. Once again, upon NaOH cleaning of the column, an elution peak was observed. Therefore, this method did not appear suitable for TDP-43 refolding.

4.4 SDS/MPD mediated refolding of TDP-43

Given that the “stripping” of detergent resulted in the precipitation of TDP-43, it appeared that detergent was necessary for the protein to remain soluble. An alternative detergent based refolding method is the SDS/MPD refolding system described by Michaux *et al*., (2008) and Roussel *et al*., (2012). In this method, 2-Methyl-2,4-pentanediol (MPD/hexylene glycol) is used to “modulate” SDS from a denaturing detergent to a non-denaturing detergent by affecting its micelle formation (Roussel *et al*., 2014).

This method was first tested by dilution of pure TDP-43 solubilised in 8 M urea (following elution from the preparative SDS-PAGE gel and acetone precipitation) into a buffer containing 120 mM SDS (as in Roussel *et al*., (2012)), to capture the protein in an SDS-denatured state. An equal volume of buffer containing 4 M MPD was then added, resulting in final SDS and MPD concentrations of 60 mM and 2 M respectively, and incubated overnight. The sample remained soluble and was then applied to a Superdex 200 10/300 gl chromatography column for analytical SEC. The resulting chromatogram is shown in Fig. 4.5, and an SDS-PAGE gel of the corresponding fractions shown in Fig. 4.6. As shown, a large elution peak was observed, centred at 11.2 ml, with a second peak at approximately 19.5 ml. From calibration of the column, these peaks would correspond to apparent molecular weights of approximately 380 kDa and 15 kDa respectively. SDS-PAGE of the fractions demonstrated TDP-43 to be present in the first peak corresponding to a higher molecular weight species, but absent in the second smaller peak. Silver stain was used as a more sensitive detection method to confirm the absence of protein in the second peak (Fig 4.7).

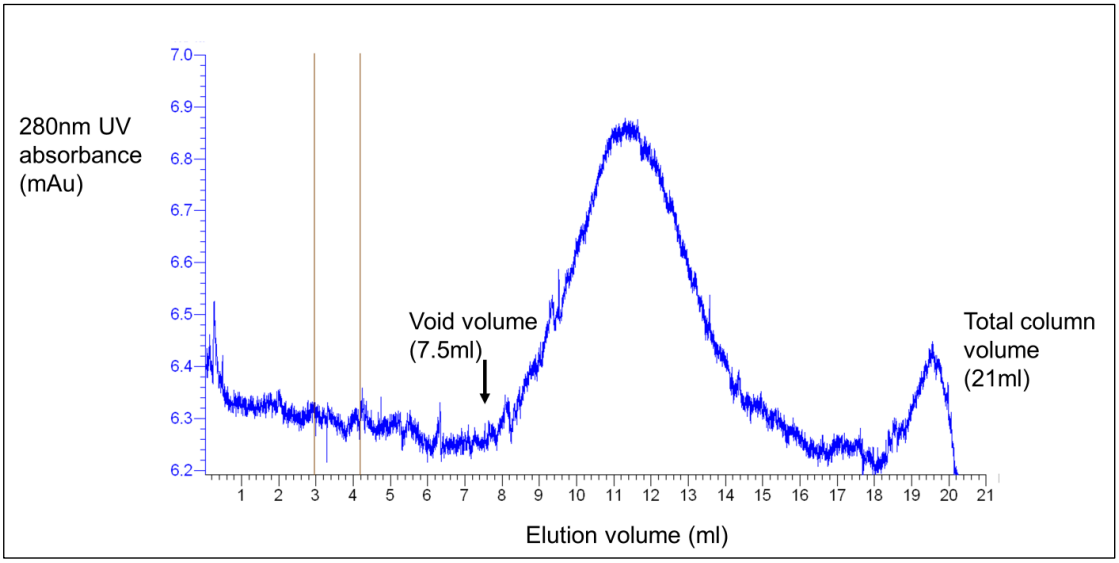


Figure 4.5. Analytical SEC of SDS/MPD refolded TDP-43. 280 nm UV absorbance with elution volume of TDP-43 sample refolded by the SDS/MPD refolding conditions used in Roussel *et al*., (2012), loaded onto a Superdex 200 10/300 gl column.

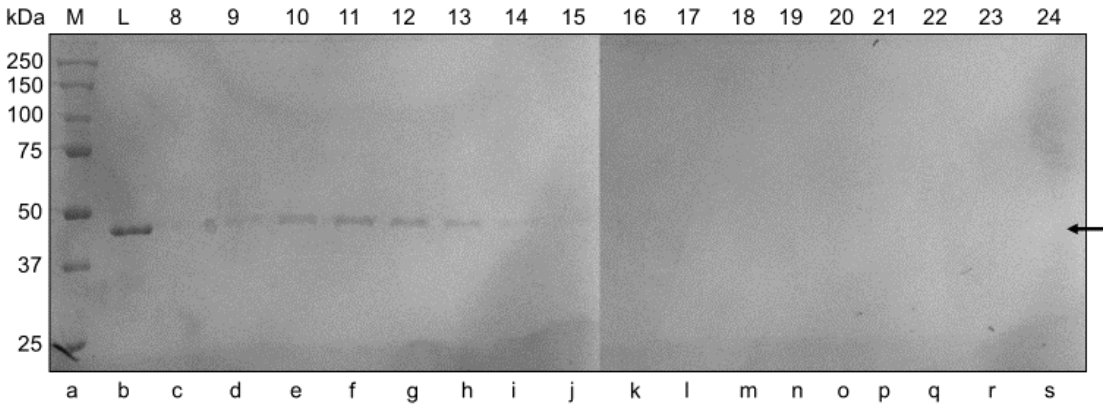


Figure 4.6. Eluted fractions from analytical SEC of SDS/MPD refolded TDP-43. Fractions collected from 8-24 ml of the analytical SEC run shown in Fig. 4.5. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); L (lane b), sample loaded onto the SEC column; 8-24 (lanes c-s), elution fractions 8-24 ml. Arrow indicates TDP-43.

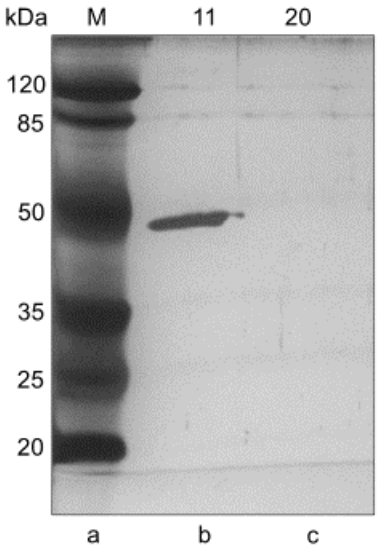


Figure 4.7. Silver stain of peak analytical SEC fractions. Silver stained SDS-PAGE gel of peak centre fractions 11 and 20 from the analytical SEC run shown in Fig. 4.5. M (lane a), Pierce™ Prestained Protein MW Marker (Fisher Scientific); 11 (lane b), elution fraction 11; 20 (lane c), elution fraction 20.

Given the high predicted molecular weight of the species in the first peak, it was hypothesised that the protein had either formed soluble, stable aggregates, or was simply remaining in a denatured state. Size exclusion chromatography separates proteins based on their Stoke's radius, rather than their predicted molecular weight in kDa, and a denatured protein, being in an extended and unfolded conformation, will have a Stoke's radius much larger than in its folded state (Erickson, 2009). It will therefore travel through a size exclusion chromatography column much faster than it would in a globular form, resulting in an earlier peak and a higher molecular weight estimation. To determine if the protein was still denatured by the 60 mM SDS and 800 mM urea final concentration, TDP-43 solubilised in 8 M urea was applied to the column, resulting in the chromatogram shown in Fig. 4.8. As shown, the elution peak once again centred on 11.2 ml, suggesting that the protein may have remained denatured.

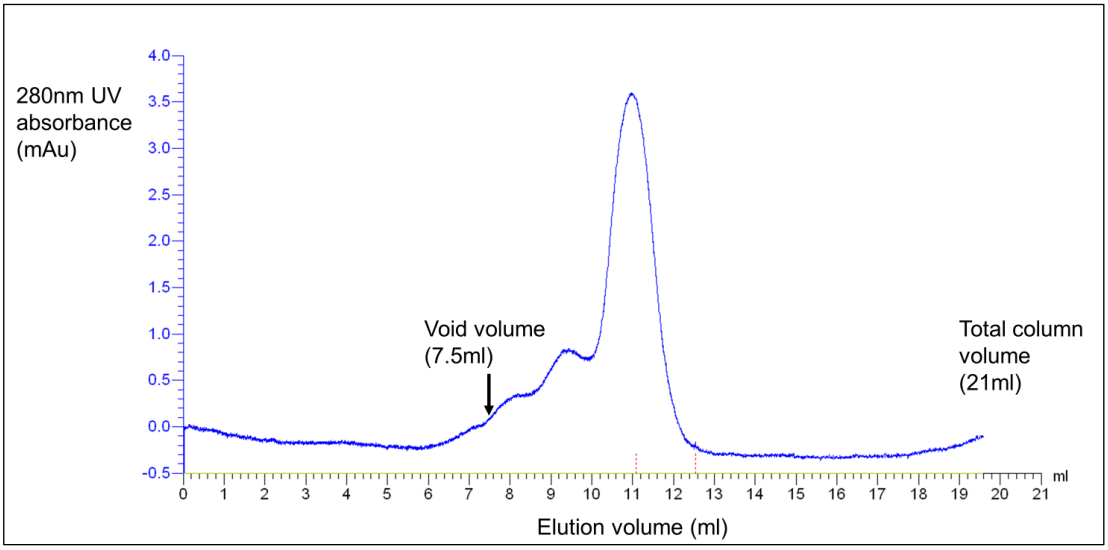


Figure 4.8. Analytical SEC of pure TDP-43 under 8M urea conditions. 280 nm UV absorbance with elution volume of denatured TDP-43 in 8 M urea conditions from the Superdex 200 10/300 gl SEC column.

As demonstrated in Michaux *et al*., (2008) and Roussel *et al*., (2012) different proteins could tolerate different concentrations of SDS before becoming denatured, even in the presence of MPD. Therefore, the final SDS concentration was reduced to 3 mM, and the refolding method attempted again. This however produced a very similar chromatogram, with the TDP-43 elution peak once again centred on 11.2 ml, suggesting that even at this lower SDS concentration, in combination with the 800 mM urea concentration due to the dilution of the protein from an 8 M urea stock, the protein may have remained in a denatured state.

4.5 TDP-43 can be refolding during elution from preparative SDS-PAGE gels

It was then hypothesised that the SDS/MPD refolding method could be applied to the protein as it eluted from the SDS-PAGE gel, where it is in an SDS denatured, monomeric state. This would also remove the acetone precipitation step, during which protein could be lost and which also could result in acetone remaining in the sample if it was not fully removed by evaporation. It would also remove the possibility that the protein remained denatured due to the presence of urea.

To investigate this, inclusion body extract was run on a large scale preparative SDS-PAGE gel, the gel negatively stained and the TDP-43 band excised from the gel. The gel slice was crushed directly into a 5 mM SDS/1.5 M MPD containing buffer and incubated overnight on a rotamixer. The eluted protein remained soluble and the sample was retrieved from the crushed gel mixture and applied to the Superdex 200 10/300 gl for analytical SEC. The initial elution/refolding buffer was designed containing components of the refolding and SEC buffers described in Vega *et al*. (2019), and consisted of 50 mM Tris-HCl, pH 8, 300 mM NaCl, 10% sucrose, 100 mM imidazole, 0.2% PEG 3350, 0.25% Triton X-100, 5 mM SDS, 1.5 M MPD, 5 mM DTT and 100 nM EDTA. DTT was added to prevent the formation of disulphide bonds, and EDTA to act as a chelating agent for any zinc ions that may be present in the gel slice after protein band visualisation using the zinc-based negative stain kit.

Analytical SEC of this sample carried out using the Superdex 200 10/300 gl column, using the BioRad NGC system, demonstrated two peaks centred at approximately 12.6-12.7 ml elution volume and 16.2 ml, as shown in Fig. 4.9. SDS-PAGE gels of the collected fractions show TDP-43 to be present in the early peak centred on 12.6-12.7 ml (Fig. 4.10 (a)). Immunoblot confirmed this species to be TDP-43 (Fig. 4.10 (b)). By column calibration, this peak could be estimated to be between 200 and 215 kDa in

size. The second peak, estimated at 48 kDa, did not appear to contain any protein, and furthermore did not result in a peak using the 215 nm wavelength as would be expected with a protein (215 nm absorbance is used as a detection of peptide bonds). This peak was therefore hypothesised to consist of micellar Triton X-100.

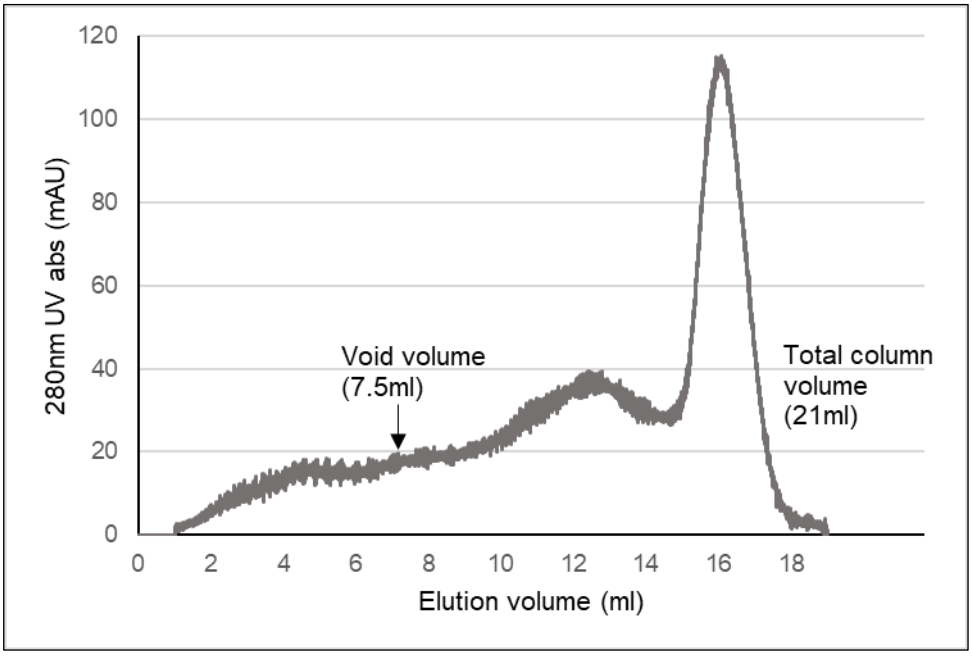


Figure 4.9. Analytical SEC of TDP-43 refolded during preparative SDS-PAGE elution. 280 nm UV absorbance with elution volume of TDP-43 sample directly refolded during elution from a preparative SDS-PAGE gel, run on the Superdex 200 10/300 gl SEC column.

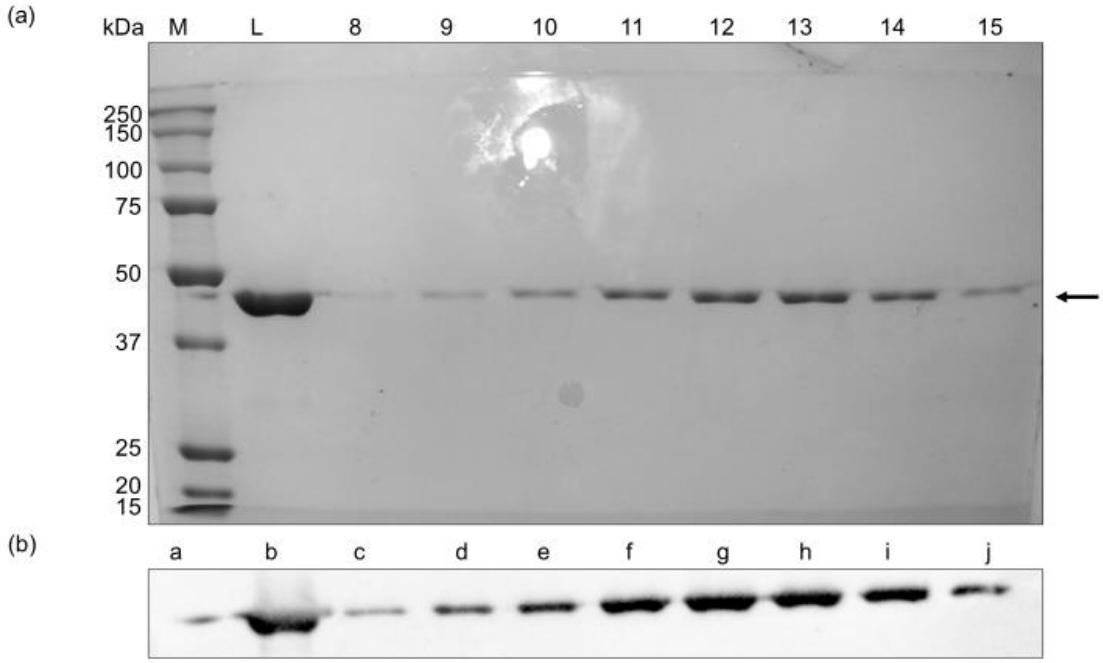


Figure 4.10. Elution fractions obtained from analytical SEC of TDP-43 refolded during preparative SDS-PAGE elution. **(a)** SDS-PAGE gel of fractions 8-15 of the analytical SEC run shown in Fig. 4.9. M (lane a)- Precision Plus Protein™ All Blue protein standards (Bio-Rad); L (lane b), sample loaded onto the SEC column; 8-15 (lanes c-j), elution fractions 8-15 ml. Arrow indicates TDP-43. **(b)** Corresponding immunoblot of sample load (lane b) and fractions 8-15 (lanes c-j) from the analytical SEC run shown in Fig. 4.9. Probed with rabbit polyclonal anti-TDP-43 antibody (Invitrogen).

To confirm this, and to also test what effect the Triton X-100 has on the refolding of TDP-43, sample was refolded directly from the preparative SDS-PAGE gel slice into the same buffer but without Triton X-100. Analytical SEC produced the chromatogram and fractions shown in Fig 4.11 and Fig 4.12. As shown, in the absence of Triton X-100, TDP-43 eluted earlier, with a peak at approximately 10.2 ml, corresponding to a much larger molecular weight species. Furthermore, the second, sharp, 16.2 ml elution peak was absent, suggesting that it was indeed Triton X-100 micelles. From this, it would appear that the absence of Triton X-100 leads to the formation of a larger

molecular weight species of TDP-43, without the observed 200-215 kDa molecular weight species that forms in its presence.

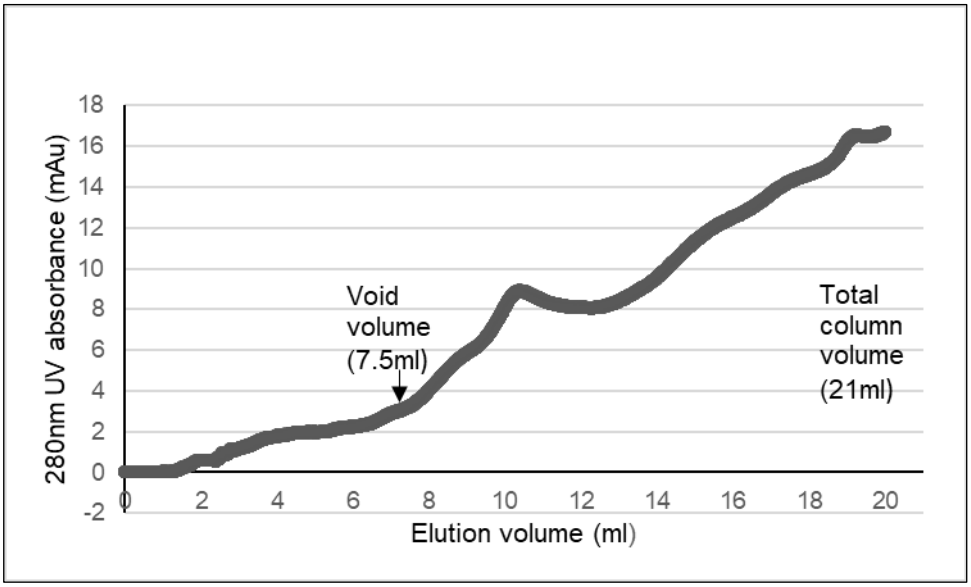


Figure 4.11. Analytical SEC of TDP-43 refolded during preparative SDS-PAGE elution in the absence of Triton X-100. 280 nm UV absorbance with elution volume of TDP-43 sample directly refolded during elution from a preparative SDS-PAGE gel in the absence of Triton X-100, run on the Superdex 200 10/300 gl SEC column.

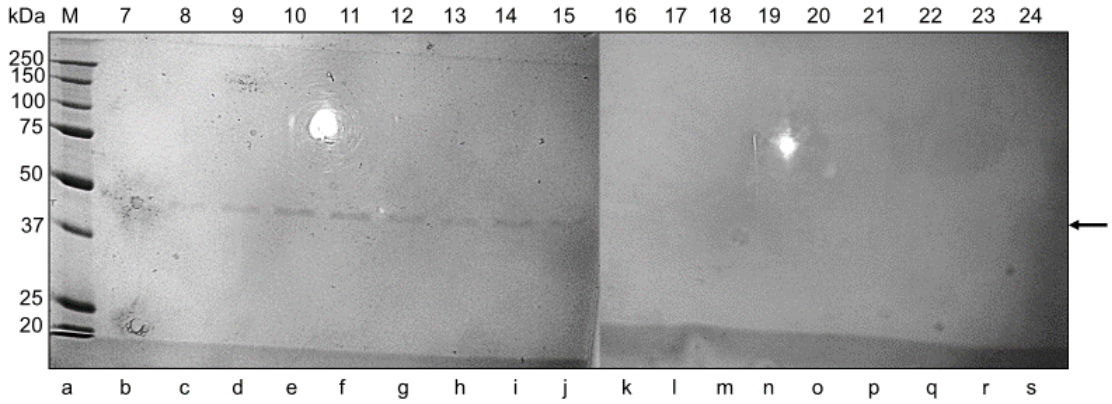


Figure 4.12. Elution fractions obtained from analytical SEC of TDP-43 refolded during preparative SDS-PAGE elution in the absence of Triton X-100. Fractions 7-24 of the analytical SEC run shown in Fig. 4.11. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); 7-24 (lanes b-s), elution fractions 7-24 ml. Arrow indicates TDP-43.

4.6 Conclusion

This chapter presents an investigation into methods to refold full-length, untagged TDP-43 purified under denaturing conditions by preparative SDS-PAGE. Attempts to refold TDP-43 by shock dilution to produce the dimer and monomer states as found by Vega et al., (2019) and Doke and Jha, (2022) were not successful, producing large oligomers and precipitates generated by these two methods respectively.

Alternatives to these methods were tested based on refolding from a detergent denatured state, as in preparative SDS-PAGE, the protein becomes denatured by SDS. Removal and “stripping” of detergent by cyclodextrin artificial chaperones resulted in rapid precipitation of TDP-43, and so methods to modulate the denaturing effect of SDS by the co-solvent MPD were investigated. It was found that by eluting the protein directly into an SDS/MPD containing refolding buffer from the preparative SDS-PAGE gel, refolding to a soluble and potentially tetrameric state could be achieved, but only in the presence of a second, non-ionic detergent, Triton X-100. The findings of this chapter therefore combine preparative SDS-PAGE purification and refolding into a single step, producing a potentially tetrameric species that can be investigated further.

5. Characterisation of refolded TDP-43

5.1 Introduction

To know whether a protein has been successfully refolded, it can be characterised by methods that investigate its oligomeric state, secondary structure and biological activity. For the TDP-43 purified and refolded in this project, this is clearly important- it has already been shown to refold to different soluble and oligomeric species depending on the buffer conditions and refolding method used. Investigation was therefore needed into the species produced by the refolding of TDP-43 during elution from preparative SDS-PAGE gels.

Published literature demonstrates evidence for the existence of TDP-43 in multiple oligomeric states. In terms of recombinant, native, purified full length TDP-43, the protein has been identified to form a stable dimer (Vega *et al*, 2019), a monomer (Doke and Jha (2022), or higher order oligomers (greater than 440 kDa) (Fang *et al*., 2014). The N-terminal domain, often considered to be the region responsible for oligomerisation, has itself been identified in multiple oligomeric states, including monomers, dimers and tetramers (Mompeán *et al*., 2016; Herrera *et al*., 2023), which can be affected by the recombinant protein length, sample concentration and the presence of reducing agents. By chemical crosslinking of proteins in cells and human tissues, a spectrum of oligomeric states, including monomers, dimers, trimers, tetramers and higher order states have been observed (Afroz *et al*., 2017). Crosslinking of the purified and refolded protein produced as described in chapter 4 can be used to investigate its oligomeric state.

Circular dichroism (CD) spectroscopy is a technique used to monitor the secondary structures of proteins. Whilst it cannot provide precise details of structure, it can provide details of the fractions of a protein in α-helical, β-sheet and random coil conformations, and so can be used to estimate the difference in structure between denatured and re-

folded protein samples, and allow comparison to models of folded structures (Anand *et al*., 2011). This form of light absorbance spectroscopy measures the difference of absorbance of right- and left-circularly polarized light by the protein. The conformation of the asymmetric peptide bond, affected by the formation of α-helical, β-sheet and random coil conformations, causes differences in the absorption of the right- and left-handed circularly polarised light, which can be measured (Anand *et al*., 2011). A CD spectrum of the refolded TDP-43 produced as described in the previous chapter can be produced and compared with those found in the literature (Vega *et al*., 2019; Doke and Jha, 2022; Kumar *et al*., 2023), to gain a consensus on the folded content of the protein.

As discussed in the literature review, the main activities characterised for TDP-43 are nucleic acid binding. DNA binding assays can therefore be used to identify biological activity of purified recombinant TDP-43. Binding to DNA probes has been demonstrated in various assays (Fang *et al*., 2014; Vega *et al*., 2019; Doke and Jha, 2022).

This chapter will present the early stages of characterisation of the refolded TDP-43 produced by the method of preparative SDS-PAGE purification and subsequent refolding during gel elution.

5.2 Purified TDP-43 may exist in multiple oligomeric states

The 200-215 kDa refolded TDP-43 species could be produced repeatedly in the presence of Triton X-100, and so attention was turned to determination of its oligomeric state. It was first hypothesised that this peak could correspond to a tetramer of TDP-43.

Fadouloglou *et al*. (2008) demonstrate a method by which chemical crosslinking with glutaraldehyde can be used in conjunction with SDS-PAGE to determine the oligomeric state of proteins. To do this, TDP-43 was purified using preparative SDS-PAGE and eluted and refolded from the gel slice into a Triton X-100 buffer containing HEPES (20mM) as the buffering agent rather than Tris-HCl, due to the incompatibility of the primary amine containing Tris with glutaraldehyde crosslinking. Analytical SEC and SDS-PAGE of the collected fractions confirmed that the same species was produced by refolding in the HEPES based buffer as in Tris (Figs. 5.1 and 5.2). The presence of the Triton X-100 micelle peak at ~16 ml elution volume was also noted.

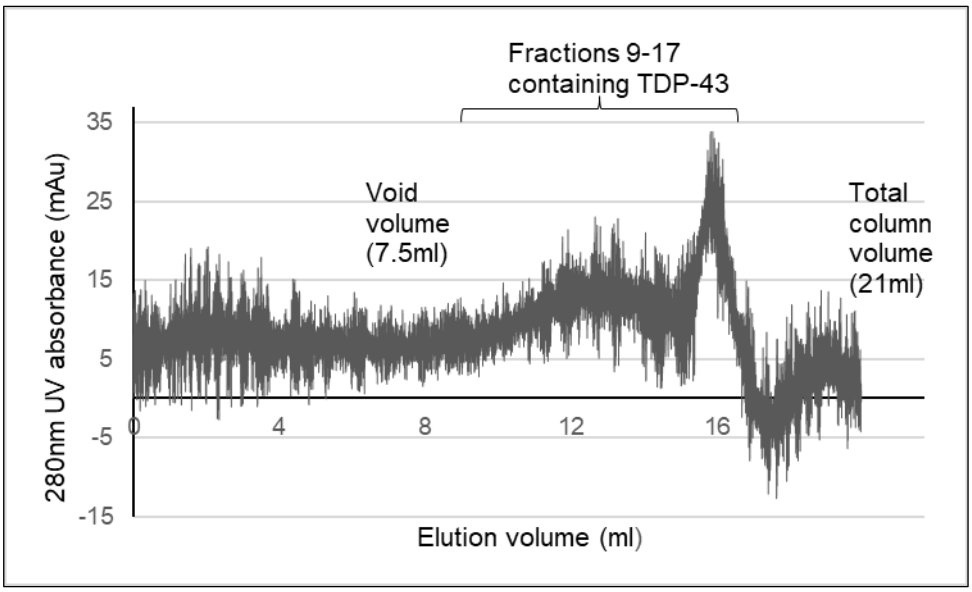


Figure 5.1. Analytical SEC of TDP-43 refolded during preparative SDS-PAGE elution into a HEPES based buffer. 280 nm UV absorbance with elution volume of TDP-43 sample directly refolded during elution from a preparative SDS-PAGE gel into a HEPES based refolding buffer containing Triton X-100, run on the Superdex 200 10/300 gl SEC column.

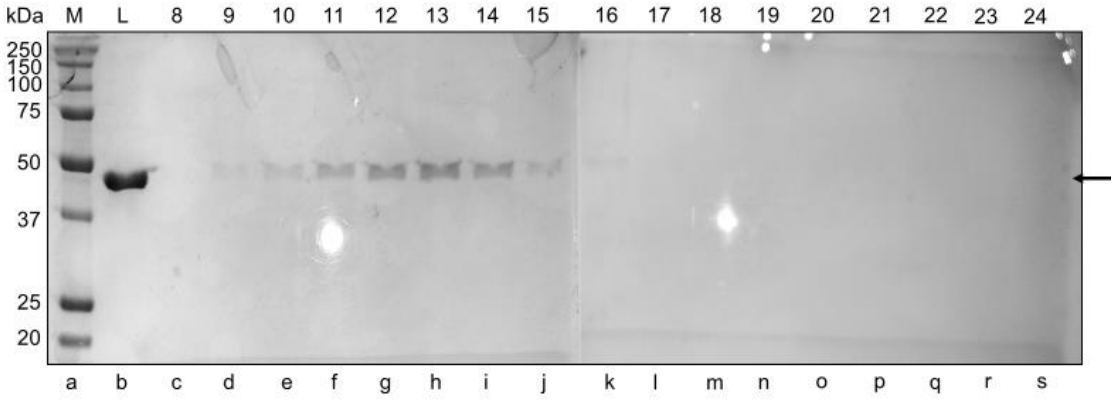


Figure 5.2. Elution fractions obtained from analytical SEC of TDP-43 refolded during preparative SDS-PAGE elution into a HEPES based buffer. Fractions 8-24 of the analytical SEC run shown in Fig. 5.1. M, Precision Plus Protein™ All Blue protein standards (Bio-Rad); L, sample loaded; 8-24, elution fractions 8-24 ml. Arrow indicates TDP-43.

The sample was therefore subjected to a crosslinking time course using the “hanging drop” method from Fadouloglou *et al.*, (2008) and the resulting samples were run on SDS-PAGE. As shown in Fig. 5.3, under these conditions, TDP-43 crosslinks to a species approximately 158 kDa, as determined by calibration of the gel from a standard curve based on the migration of the protein standards in lane M. A larger species also formed, potentially as a consequence of crosslinking itself, as a peak corresponding to a species of this size is not observed in analytical SEC. Glutaraldehyde crosslinking has been found to blur bands and cause a slight increase in the migration distance during SDS-PAGE (Griffith, 1972), so this band could correspond to a tetrameric species of TDP-43 that would be predicted at 178.8 kDa.

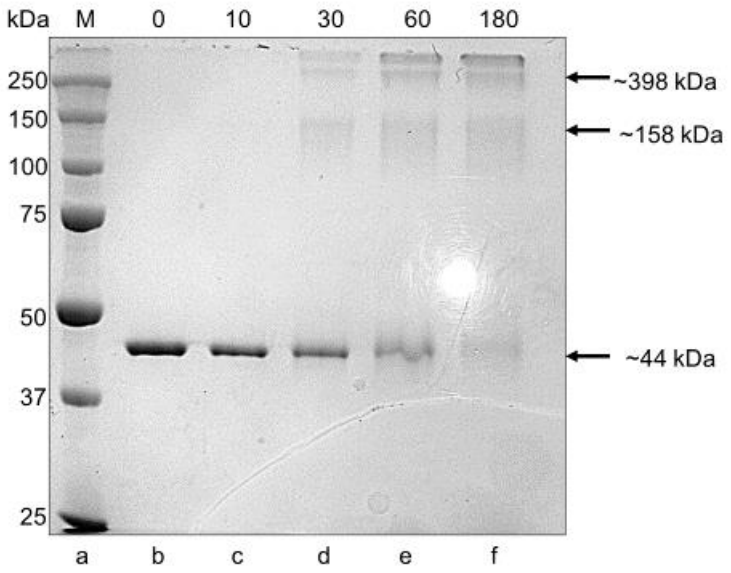


Figure 5.3. Crosslinking of refolded TDP-43. Time-course of crosslinking of TDP-43 refolded from a preparative SDS-PAGE gel into a HEPES- based refolding buffer. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); 0 (lane b), 0 minutes of crosslinking; 10 lane c), 10 minutes of crosslinking; 30 (lane d), 30 minutes of crosslinking; 60 (lane e), 60 minutes of crosslinking; 180 (lane f), 180 minutes of crosslinking. Arrows indicate species with molecular weights determined by gel calibration using the migration of the protein standards in lane a (M).

5.3 Circular dichroism spectroscopy of TDP-43

Circular dichroism (CD) spectroscopy can be used to provide evidence of the secondary structural components of a protein in solution. To carry out CD spectroscopy, the protein had to be exchanged into a suitable buffer. A buffer was designed as follows: 20 mM $NaPO_4$, pH 8, 150 mM NaF, 0.2% PEG 3350, 10% sucrose, 0.25% Triton X-100, 5 mM SDS, 1.5 M MPD. Protein refolded in the HEPES based buffer was exchanged into this CD buffer via HiTrap desalting column. Given the alteration in buffer conditions, crosslinking and SEC was used to confirm that TDP-43 forms similar species in this CD buffer. A time course of glutaraldehyde crosslinking was then carried out to identify if the protein had remained in the tetrameric state. From this point, due

to the distribution of TDP-43 into multiple bands on SDS-PAGE gels, silver stain was used as a more sensitive staining method for crosslinking experiments. As shown in Fig. 5.4, there is an observable shift in the protein band corresponding to a tetramer compared to Fig. 5.3. Furthermore, in Fig 5.4, there is an appearance of a band between the 75 and 100 kDa standards, which could correspond to the formation of a dimer. Analytical SEC produced elution of the protein in the same fractions as the HEPES buffer (Fig. 5.5).

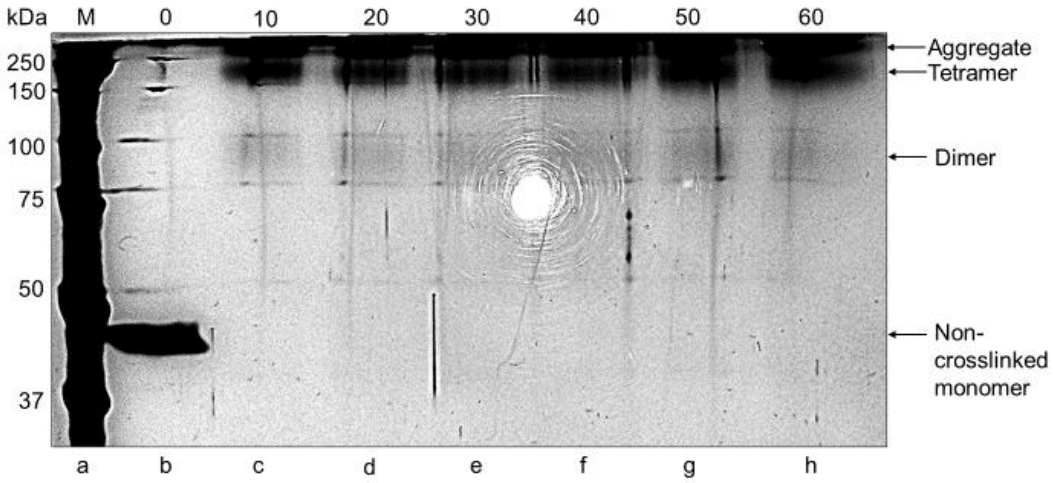


Figure 5.4. Crosslinking of TDP-43 exchanged into CD buffer. A time-course of crosslinking of TDP-43 refolded from a preparative SDS-PAGE gel into a HEPES- based refolding buffer followed by buffer exchange into a CD spectroscopy buffer. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); 0 (lane b), 0 minutes of crosslinking; 10 (lane c), 10 minutes of crosslinking; 20 (lane d), 20 minutes of crosslinking; 30 (lane e), 30 minutes of crosslinking; 40 (lane f), 40 minutes of crosslinking; 50 (lane g), 50 minutes of crosslinking; 60 (lane h), 60 minutes of crosslinking.

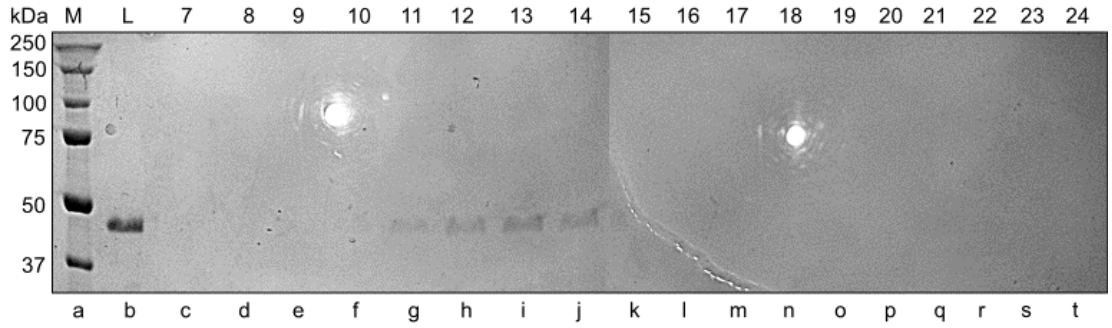


Figure 5.5. Elution fractions obtained from analytical SEC of refolded TDP-43 buffer exchanged into CD buffer. Fractions collected of elution volumes 7-24 ml of analytical SEC of protein refolded in HEPES buffer and buffer exchanged into CD spectroscopy buffer and applied to the Superdex 200 10/300 gl SEC column. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); L (lane b), sample loaded; 7-24 (lanes c-t), elution fractions 7-24 ml.

To determine whether the shift and potential dimer formation was due to the slight dilution factor produced by the HiTrap desalting column buffer exchange, protein was refolded directly into the CD buffer. Chemical crosslinking of this sample appeared to produce the same size band as that produced by refolding into HEPES as shown in Fig. 5.6. It was then considered that during elution from the gel slice, Tris and glycine may elute from the gel slice with the protein, which could have an effect on the refolding or the crosslinking. To remove any solutes that may co elute with the protein, sample refolded directly into the CD buffer was passed through a PD-10 desalting column pre-equilibrated with the CD buffer, using the PD-10 column centrifugation protocol to avoid any dilution. In this way the protein would exist in the CD buffer as designed, without any co-eluted Tris or glycine. Following this, the sample was diluted (1 in 2), incubated at room temperature overnight, and then all dilutions were chemically crosslinked and the results were run on an SDS-PAGE gel, shown in Fig. 5.7. As shown, the undiluted sample produced by buffer exchange on the PD-10 column showed the same tetrameric shift as the previously HiTrap desalting exchanged sample from the HEPES

buffer, with an observation of a potential dimer along with some monomeric TDP-43 that could either be a true monomer, or simply be a product of incomplete crosslinking. The serially diluted samples showed this same crosslinking pattern, with no apparent shift in the proportions of TDP-43 in each oligomeric species.

Also shown in Fig. 5.7 is the crosslinking of a sample incubated with 10 mM DTT (Lanes g and h), to test whether as demonstrated with the NTD by Herrera *et al*., (2023), a reducing environment could shift any tetramer-dimer equilibrium. The results shown do not appear to indicate that this encouraged dimer formation, with the tetramer band still remaining more prominent than that of the dimer.

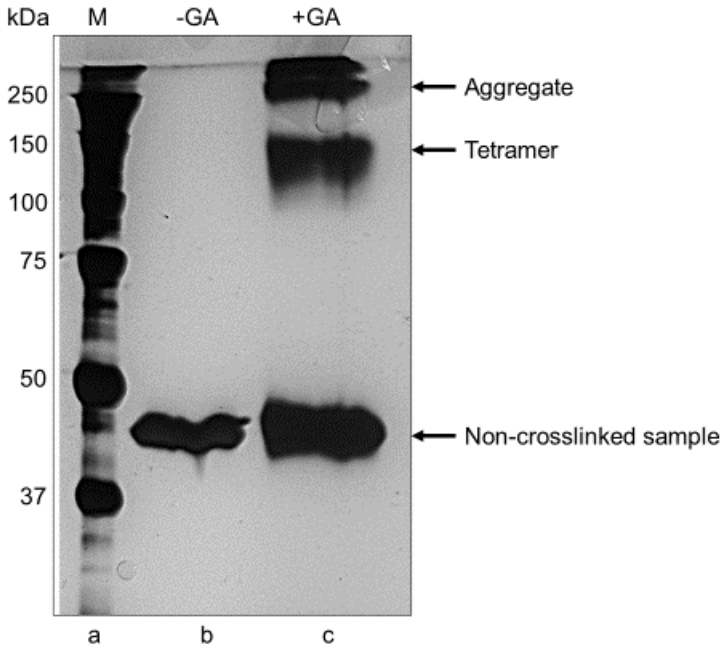


Figure 5.6. Crosslinking of TDP-43 refolded directly into CD spectroscopy buffer. M (lane b), Precision Plus Protein™ All Blue protein standards (Bio-Rad); -GA (lane b), sample before crosslinking; +GA (lane c), sample following 1 hour of crosslinking.

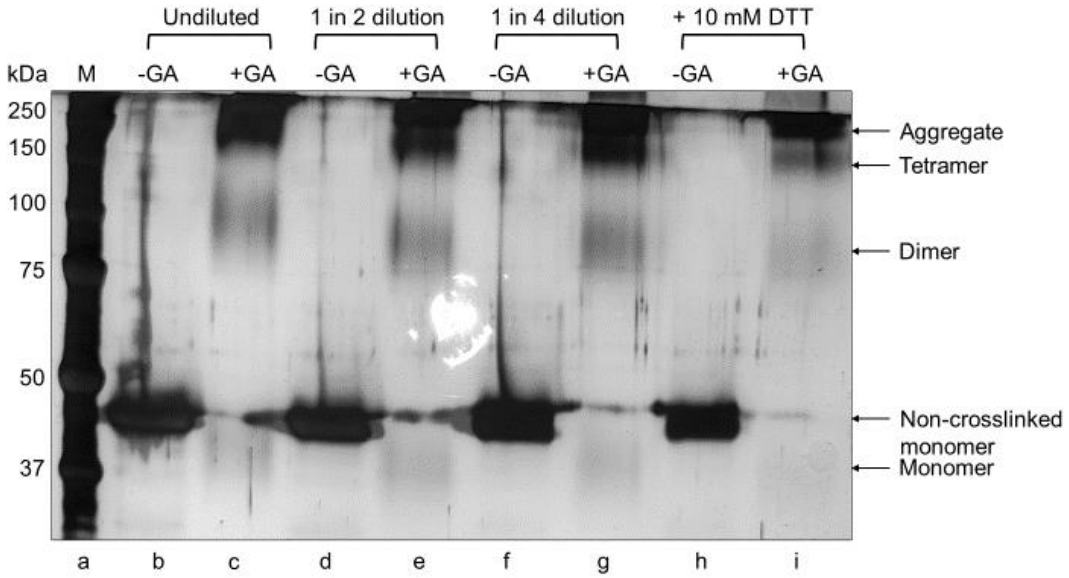


Figure 5.7. **Crosslinking of TDP-43 refolded directly into CD spectroscopy buffer followed by PD-10 column buffer exchange, with varying dilution and reducing environment conditions.** Glutaraldehyde crosslinking of sample desalted by PD-10 column into CD spectroscopy buffer (after elution/refolding into the same buffer) and serially diluted (1 in 2). M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); lane b and c, undiluted buffer exchanged sample before (-GA) and after (+GA) crosslinking respectively; lanes d and e, 1 in 2 diluted sample before (-GA) and after (+GA) crosslinking respectively; lanes f and g, 1 in 4 diluted sample before (-GA) and after (+GA) crosslinking respectively; lanes h and i, Sample incubated overnight with 10 mM DTT before (-GA) and after (+GA) crosslinking respectively.

To determine whether the previously observed tetramer/dimer/monomer bands were distinct species or simply products of incomplete crosslinking, crosslinking followed by SEC could be used. The rationale behind this was as follows: if the protein sample is incompletely chemically crosslinked before applying to an SEC column, but retained in native conditions, then a single tetrameric species will elute with all components of the incomplete crosslinking in one single peak. If the different bands observed do correspond to lower oligomerisation states distinct from the tetramer, then they will elute at different (later) points from the column than the tetramer (dimer then monomer

respectively). An SDS-PAGE gel of elution fractions 9-16 of an analytical SEC run of crosslinked sample is shown in Fig. 5.8. As shown, the aggregate, tetramer, dimer and monomer bands elute in clear, slightly overlapping stages. This suggests that the protein exists in multiple oligomerisation states once completely buffer exchanged into the CD buffer.

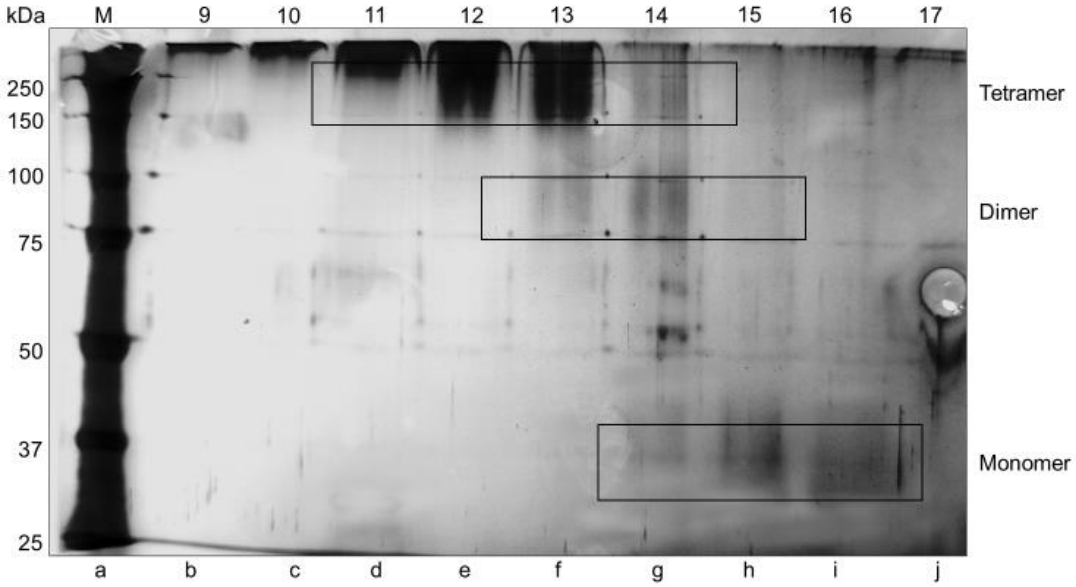


Figure 5.8. Elution fractions obtained from analytical SEC of TDP-43 in CD buffer following glutaraldehyde crosslinking. Fractions collected of elution volumes 9-17 ml of analytical SEC of protein refolded and desalted in CD spectroscopy buffer, which was then crosslinked prior to application to the Superdex 200 10/300 gl column. M (lane a), Precision Plus Protein™ All Blue protein standards; 9-17 (lanes b-j), elution fractions 9-17 ml.

Far-UV CD spectroscopy of the sample was then carried out using the Applied Photophysics Chirascan™. Initially the sample was used at 0.066 mg/ml in CD buffer. This however produced a scan with a large amount of “noise”, due to the buffer components, likely mainly the Triton X-100 which as identified in previous SEC data has a large UV absorbance ability. To reduce noise and produce a useable CD spectrum of the protein sample, the sample was diluted using MilliQ water, with buffer blanks diluted equally until the noise was reduced and a clear spectrum could be produced. As shown in Figs. 5.9 and 5.10, once the protein was diluted down to 0.01 mg/ml (6.6x dilution), the noise was reduced enough to produce usable scans down to approximately 200-205 nm.

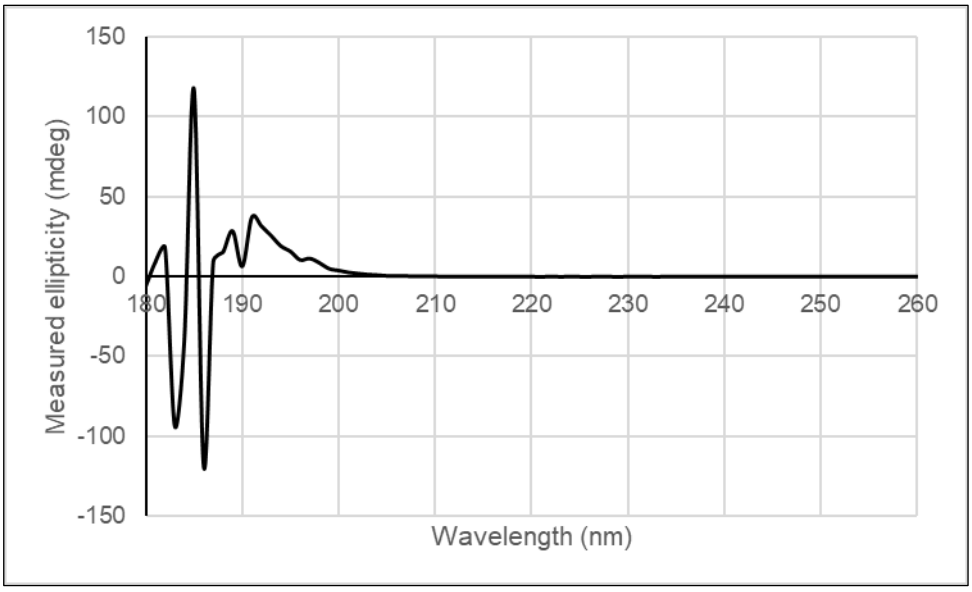


Figure 5.9. Far-UV CD spectrum of diluted CD buffer blank. (6.6 x dilution of CD buffer with MilliQ water), wavelengths 180-260 nm, measured in machine units of mdeg.

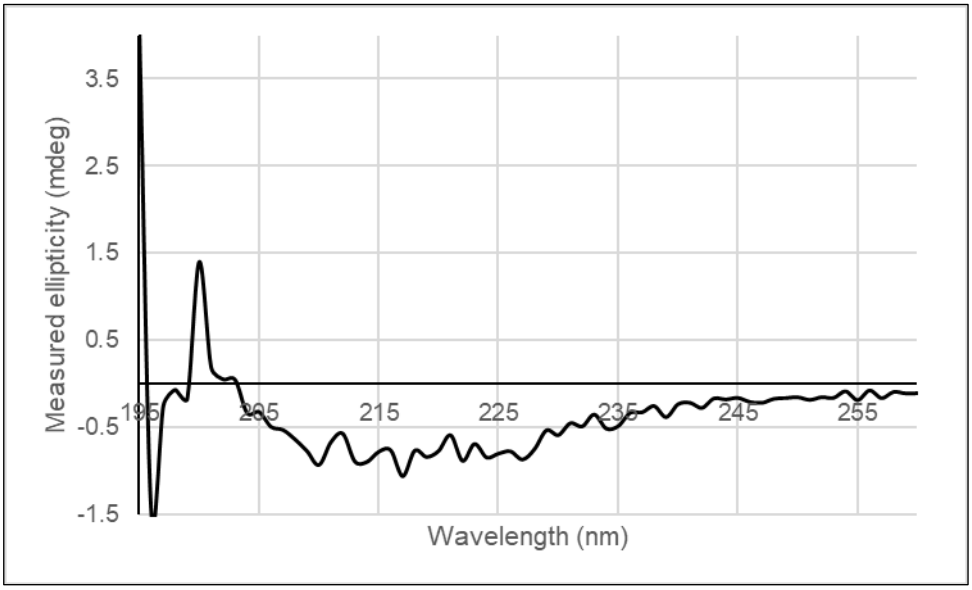


Figure 5.10. Far-UV CD spectrum of 0.01 mg/ml TDP-43. Protein in CD buffer 6.6x diluted with MilliQ water, wavelengths 180-260 nm, measured in machine units of mdeg.

The CD data produced was fitted and smoothed using the BeStSel server, converted to per residue molar absorption units of circular dichroism (Δε) ($M^{-1}cm^{-1}$) and a spectrum was produced shown in Fig. 5.11. Comparison was made with CD spectra of representative confirmations of alpha helical, beta sheet and random coil structures (Fig. 5.12). The spectrum shows a minimum at 218 nm, indicative of β sheet content. Indeed, the spectrum appears representative of a predominantly β sheet structure, which may be expected from the folded regions of TDP-43. The spectrum however also shows signs of helical content. Reading the curve from right left (250-200 nm) the curve decreases from 250 nm rather than rising as with a beta sheet spectrum. There are two sight negative shoulders at approximately 208 nm and 222 nm, and the trace rises more sharply from the 208 nm shoulder crossing the 0 Δε axis below 205 nm, rather than closer to 210 nm like a β sheet trace.

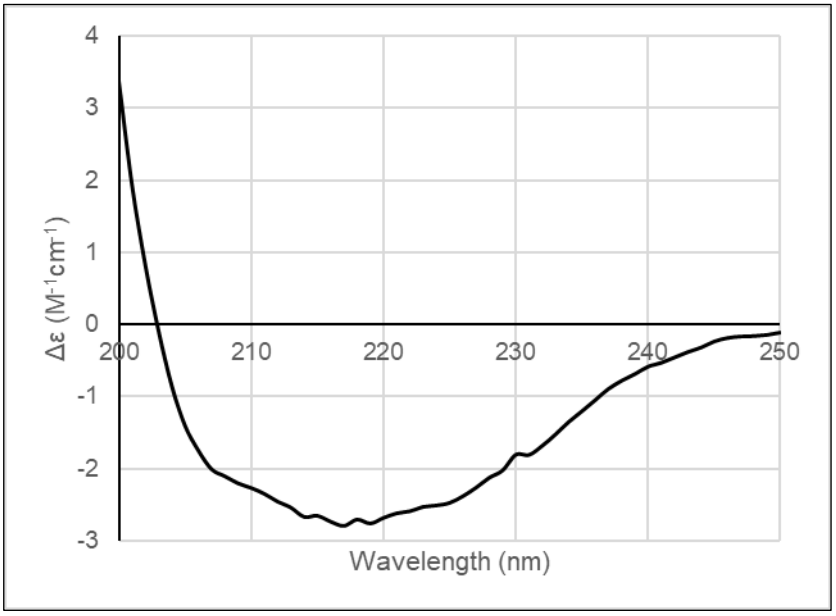


Figure 5.11. BeStSel server fitted and smoothed CD spectrum of TDP-43. 0.01 mg/ml TDP-43 in 6.6x diluted CD buffer, wavelengths 200-250 nm, measured in per residue molar absorption units of circular dichroism (Δε) ($M^{-1}cm^{-1}$).

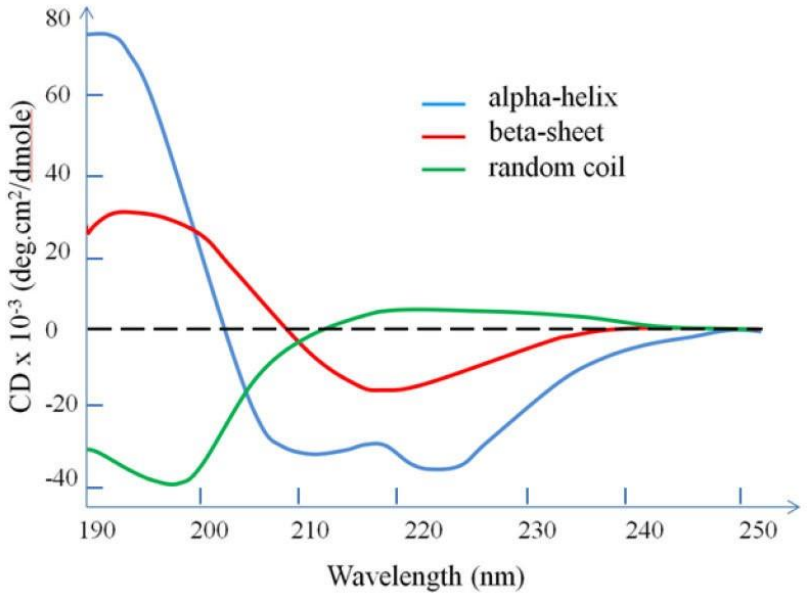


Figure 5.12. Example Far-UV CD spectra. Example Far-UV CD spectra for proteins with a predominantly alpha helical (blue), beta-sheet (red) and random coil (green) structure. From (Wei *et al*., 2014).

5.4 TDP-43 DNA binding activity

To determine whether the refolded protein could have biologically relevant activity, a DNA binding assay was carried out, using Cy3 labelled TG repeat and AC repeat ssDNA probes as binding targets. Co-immunoprecipitation of TDP-43 DNA complexes was tested in CD buffer. Supernatant and pellet fractions of the immunoprecipitation were separated and spotted onto nylon membrane, which was then imaged for Cy3 fluorescence (Fig. 5.13). As demonstrated, in the pellet fractions, Cy3 fluorescence was identified to be strongest in the reaction of TDP-43 incubated with a TG repeat probe. The TDP-43 AC probe reaction also showed a slight increase in the fluorescence compared to the samples which did not contain protein. The supernatant fractions demonstrated that the majority of the DNA probe remained in the supernatant for all reactions, as this membrane required a shorter exposure time to generate the image of fluorescence shown. SDS-PAGE of the Co-IP supernatant and pellet fractions show that the “pulldown” of TDP-43 appears to have been inefficient, with the majority of TDP-43 being present in the supernatant in all reactions where TDP-43 was included in the IP. This can be seen by the prominence of the TDP-43 band in Fig. 5.14 lane b compared to lane c, lane h compared to lane i and lane j compared to lane k. This may have been due to the presence of the detergents in the CD buffer. Attempts to buffer exchange the protein into buffers potentially more suitable for DNA binding resulted in either loss of protein from the sample by precipitation or loss of TDP-43 pulldown by the immunoprecipitation method.

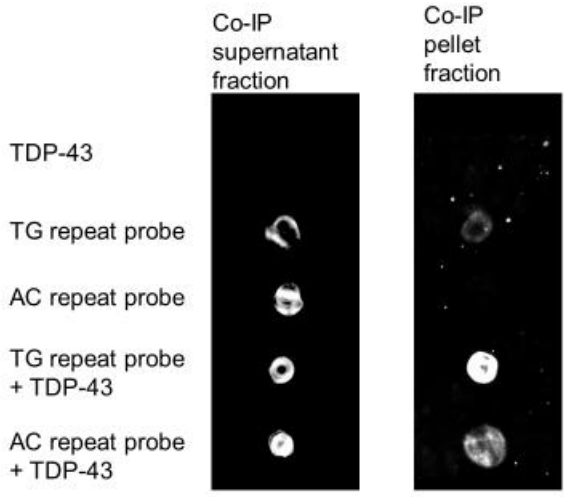


Figure 5.13. Results of Co-IP of TDP-43 and Cy3 labelled TG and AC repeat probes. Images show Cy3 fluorescence of supernatant or pellet fractions spotted onto nylon membrane. TDP-43 sample, TDP-43 protein incubated in the absence of any DNA probe before IP; TG and AC, probe samples included probes incubated in the absence of TDP-43 before IP; TG and AC repeat probe + TDP-43 sample, DNA probe incubated with TDP-43 before IP.

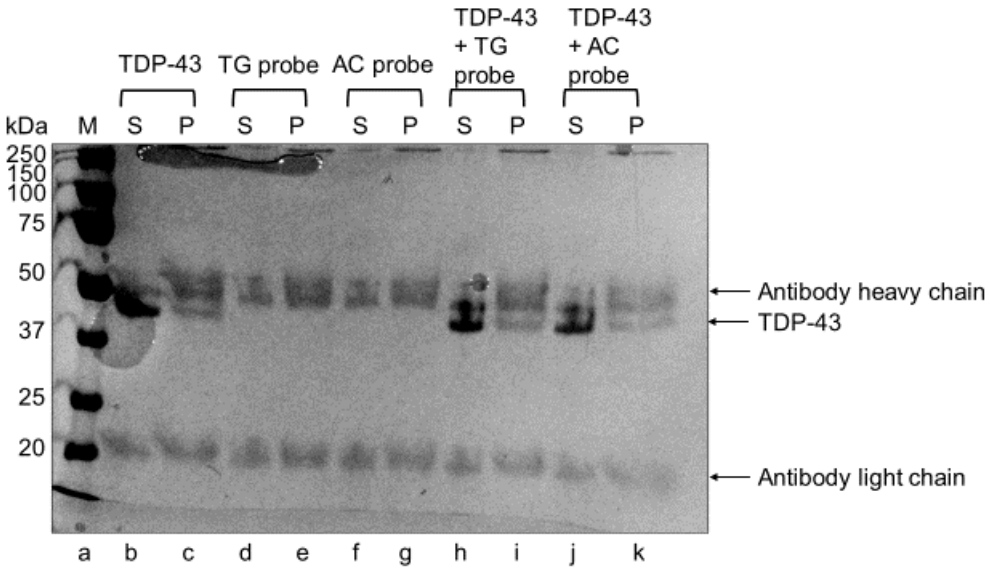


Figure 5.14. Supernatant and pellet fractions of Co-IP reactions. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); lanes b and c , supernatant and pellet respectively of TDP-43 incubated in the absence of DNA probe; lanes d and e , supernatant and pellet respectively of TG probe incubated in the absence of TDP-43; lanes f and g, supernatant and pellet respectively of AC probe incubated in the absence of TDP-43; lanes h and i, supernatant and pellet respectively of TDP-43 incubated with TG repeat probe; lanes j and k, supernatant and pellet respectively of TDP-43 incubated with AC repeat probe.

5.5 Conclusion

This chapter presents characterisation of the TDP-43 produced by purification using preparative SDS-PAGE and refolding during elution from the preparative gel slice. Investigations into the oligomeric state of the species produced by this method indicates a sample containing predominantly tetrameric TDP-43, however exchange into alternative buffers can produce a sample containing a mixture of monomer, dimer, tetramer and potentially higher order aggregates (these however may be artifacts produced by the crosslinking method of oligomeric state determination, given the apparent absence of a peak of corresponding size during analytical SEC).

Circular dichroism spectroscopy indicates that the protein sample contains secondary structural elements corresponding to both alpha helical and beta sheet content, as may be expected for TDP-43. Comparison can be made with spectra produced in the literature, which will be expanded upon in the discussion.

In initial tests of a DNA binding assay, the protein appears to have some biological activity, with an apparent higher affinity for a TG repeat ssDNA molecule than an AC repeat probe. However further optimisation is required to draw a more solid conclusion from these assays, with further assays needed to confirm the results.

6. Purification of His-tagged TDP-43 C-terminal fragment and aggregation assay development

6.1 Introduction

Given the difficulties in purifying a large quantity of untagged full length TDP-43, an alternative TDP-43 recombinant protein was required for developing an aggregation assay and testing of TDP-43 aggregation inhibitors. As discussed in the literature review, one species often identified in aggregates from ALS and FTLD patients is TDP-43 C-terminal fragments, expression of which in cell models recapitulates aggregation and toxicity (Zhang *et al*., 2009). Utilisation of a TDP-43 CTF construct with a C-terminal histidine tag should allow for purification of a large quantity of protein for use in aggregation assays. Prasad *et al*. (2016) utilised a C-terminal fragment for the testing of AIM4, and so as in their work, a purification and aggregation method based on purification under denaturing conditions and aggregation by dilution of the denaturant could be developed.

This chapter will present the development of a routine method to purify a TDP-43 C-terminal fragment, and use of this protein to develop a reliable fluorescence-based aggregation assay.

6.2 Expression and solubility of TDP-43 CTF

The pKMH9 plasmid, encoding a ~21 kDa C-terminal TDP-43 fragment (CTF 208-414) with a C-terminal 6x His tag was purified from stocks and sequencing confirmed a wild type construct. SHuffle T7 Express *E. coli* (New England BioLabs) were transformed with the pKMH9 plasmid and 3 hour small scale expression was tested at 37ºC, with small scale solubility analysis carried out on the resulting 3 hour culture. Fig. 6.1 shows the successful expression of the TDP-43 CTF over 3 hours of IPTG induction at 37ºC, shown by the increasing prominence of the ~21 kDa band with time (lanes b-e). The solubility analysis demonstrated that at this temperature and following 3 hours of induction, the protein fragment appeared predominantly in the insoluble fraction.

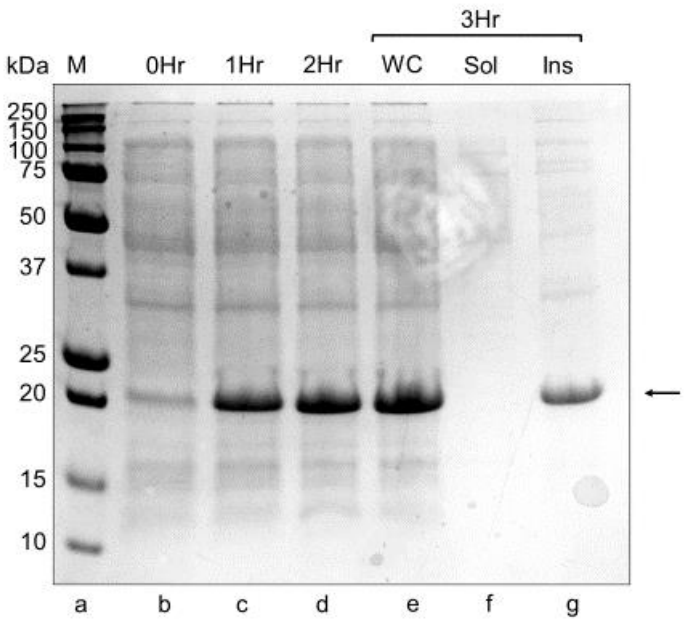


Figure 6.1. Expression and solubility of TDP-43 CTF. Timepoints of 3 hour His-tagged TDP-43 CTF expression from the pKMH9 plasmid. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); 0Hr (lane b), whole cell extract after 0 hours of expression; 1Hr (lane c), whole cell extract after 1 hour of expression; 2hr (lane d), whole cell extract after 2 hours of expression. WC (lane e), whole cell extract; Sol (lane f), Soluble cell fraction; Ins (lane g), Insoluble cell fraction, all after 3 hours of expression. Arrow indicates His-tagged TDP-43 CTF.

6.3 Purification of TDP-43 CTF

Given that the expressed TDP-43 CTF was C-terminally 6x His tagged, the protein could be purified under denaturing conditions using immobilized metal affinity chromatography (IMAC). TDP-43 CTF was expressed for 3 hours at 37ºC in a 1 litre culture, then the cells were harvested and resuspended in denaturing IMAC buffer. Cells were lysed by freeze thaw and sonication, left overnight in the denaturing buffer for insoluble proteins to solubilize, then any remaining insoluble material was removed by centrifugation. The resulting denatured cell extract could then be filtered and applied to an IMAC column.

Figure 6.2 shows the peak fractions obtained from gradient elution IMAC using a 1 ml HisTrap FF column (Cytiva), to which the denatured cell extract from a 1 litre culture was applied. As shown, the TDP-43 CTF eluted across a total of 8 fractions, with low levels of remaining contaminants. To reduce the number of fractions across which the protein elutes and produce a more concentrated elution product, a step elution was tested, the results of which are shown in Fig. 6.3. As shown, the protein did indeed elute in fewer fractions, with comparable levels of remaining contaminants as the gradient elution. The fractions collected from the gradient elution could also be pooled and successfully concentrated using a Vivaspin™ centrifugal concentrator with a 5 kDa molecular weight cut off, however given the potential for sample loss during concentration, and the effectiveness of the step elution IMAC method, step elution was selected for routine use.

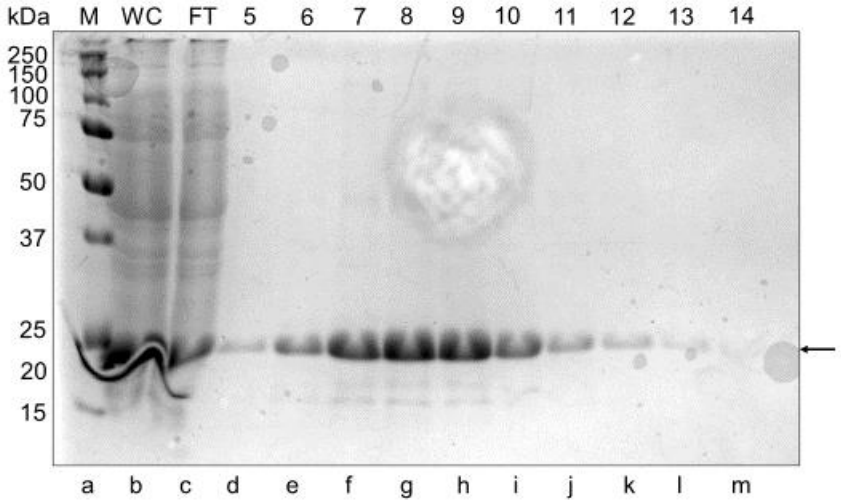


Figure 6.2. Gradient IMAC of denatured whole cell extract following TDP-43 CTF expression. Peak IMAC fractions using gradient elution. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); WC (lane b), denatured whole cell extract; FT (lane c), column flow through; 5-14 (lanes d-m), peak fractions collected during elution. Arrow indicates His-tagged TDP-43 CTF.

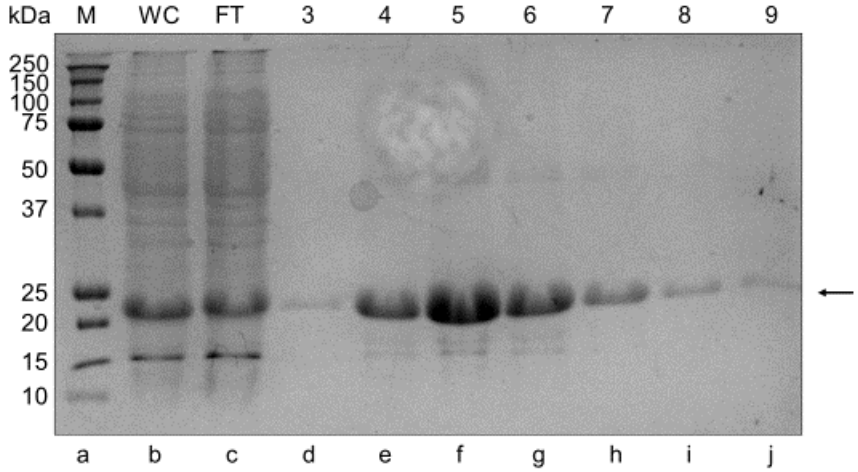


Figure 6.3. Step elution IMAC of denatured whole cell extract following TDP-43 CTF expression. Peak IMAC fractions using step elution. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); WC (lane b), denatured whole cell extract; FT (lane c), column flow through; 3-9 (lanes d-j), peak fractions 3-9 ml collected during elution. Arrow indicates His-tagged TDP-43 CTF.

To further purify the TDP-43 CTF, size exclusion chromatography was carried out under denaturing conditions using a Sephacryl S-300-HR 26/60 column (Cytiva). The TDP-43 CTF eluted across an ~23 ml peak (Fig. 6.4 lanes c-y), the fractions of which could be successfully pooled and concentrated. The steps of the purification procedure of the TDP-43 CTF are demonstrated in Fig. 6.5, with the final sample showing low levels of two contaminants, at approximately 17 and 14 kDa, which could be degradation products. This level of contamination was deemed acceptable for the requirements of the TDP-43 CTF aggregation assays intended.

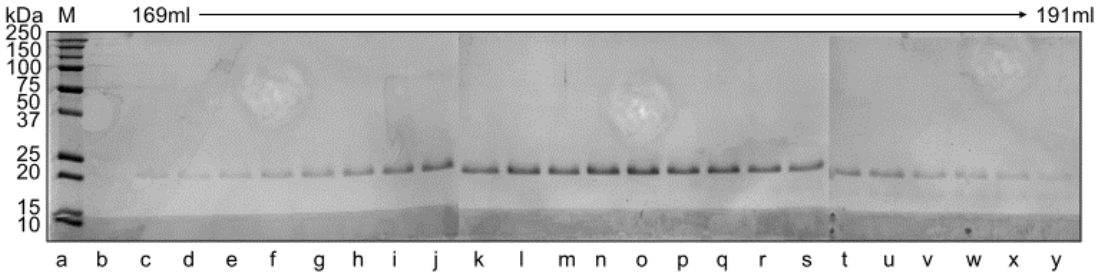


Figure 6.4. Preparative SEC of TDP-43 CTF following IMAC. Fractions of the SEC peak corresponding to TDP-43 CTF. M (lane a)- Precision Plus Protein™ All Blue protein standards (Bio-Rad). Fractions of elution volume 169 ml-191 ml (lanes c-y) contained TDP-43 CTF.

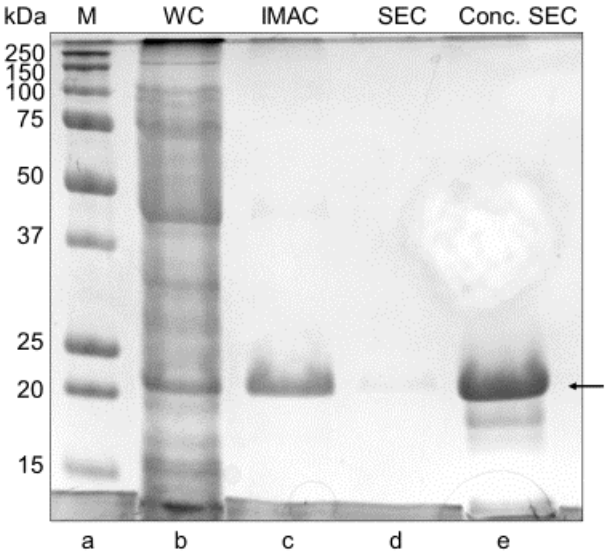


Figure 6.5. Purification procedure for His-tagged TDP-43 CTF. M (lane a), Precision Plus Protein™ All Blue protein standards (Bio-Rad); WC (lane b), denatured whole cell extract following 3 hr expression at 37ºC; IMAC (lane c), pooled IMAC fractions; SEC (lane d), pooled SEC fractions; Conc, SEC (lane e), concentrated pooled SEC fractions (~5 µg loaded). Arrow indicates TDP-43 CTF.

For aggregation assays using the TDP-43 CTF, SEC peaks collected were maintained in TDP-43 CTF preparative SEC buffer and frozen at -20ºC until use. The peak fractions were concentrated when required using Vivaspin™ centrifugal concentrators with a molecular weight cut off of 5 kDa. For aggregation reactions, the stock was adjusted to a concentration that would allow dilution to suitable concentrations of protein and urea to allow aggregation to occur without instant precipitation of the protein.

6.4 Design of a TDP-43 CTF aggregation assay

Given the difficulties in purifying and working with TDP-43 and its fragments, it is perhaps not surprising that there are few published reports of investigations of purified recombinant TDP-43 aggregation *in vitro.* For amyloid proteins, fluorescent amyloid aggregate binding dyes such as Thioflavin S (ThS), Thioflavin T (ThT) and Congo red are often used for aggregation assays where the aggregation kinetics of a protein sample can be monitored with time. There is contrasting evidence in the literature on the usefulness of these dyes for monitoring the aggregation of TDP-43 constructs. Johnson *et al*., (2009), demonstrated aggregated TDP-43 to be minimally fluorescent when incubated with both ThT and Congo red, suggesting these dyes are not suitable for use in an aggregation assay. Similarly, Saini and Chauhan, (2011), also showed a TDP-43 CTF (220-414) to show minimal ThT fluorescence, but did demonstrate significant ThT fluorescence of aggregates produced from short peptides consisting solely of aggregation sequences. Contrastingly, Prasad *et al*., (2016) successfully used the ThT assay to monitor TDP-43 CTF aggregation with time when testing AIM4 and other acridine derivatives as aggregation inhibitors.

The potential usefulness of ThT to monitor TDP-43 CTF aggregation was therefore investigated. As in Prasad *et al*. (2016), protein was purified and maintained in a 4 M urea stock, before being concentrated and diluted in a non-urea buffer (50mM HEPES, pH 8) to dilute down the urea so allowing protein aggregation. The protein, initially diluted to 20 μM (1.5 M urea final concentration), was incubated at 25ºC with 20 μM ThT in a Tecan Infinite plate reader, reading fluorescence with 440 nm excitation and 485 nm emission. This concentration was determined using the literature surrounding ThT aggregation assays, such as its use with α-synuclein (Wördehoff and Hoyer, 2018), Aβ40, Aβ42 and prion Ure2 (Xue *et al*., 2017) and its tendency to become self-fluorescent at higher concentration (Xue *et al*., 2017). Repeat wells of the same initial reaction were measured for 4 hours of aggregation initially (Johnson *et al*. (2009)

identified TDP-43 recombinant proteins to aggregate within 2 hours) and the data for this is shown in Fig. 6.6. As shown, the ThT fluorescence between repeat wells varied fairly broadly with a difference of up to 12 000 RFU for repeat wells of identical reactions. After multiple reactions, including longer incubation periods (up to 18 hours), these differences remained and did not appear to show any pattern based on well position that could indicate errors in the plate reader. Furthermore, the plate reader gain settings required to generate this level of fluorescence was very high, at 210 AU (arbitrary units), indicating that the aggregation of the TDP-43 CTF in these conditions is only very weakly ThT fluorescent.

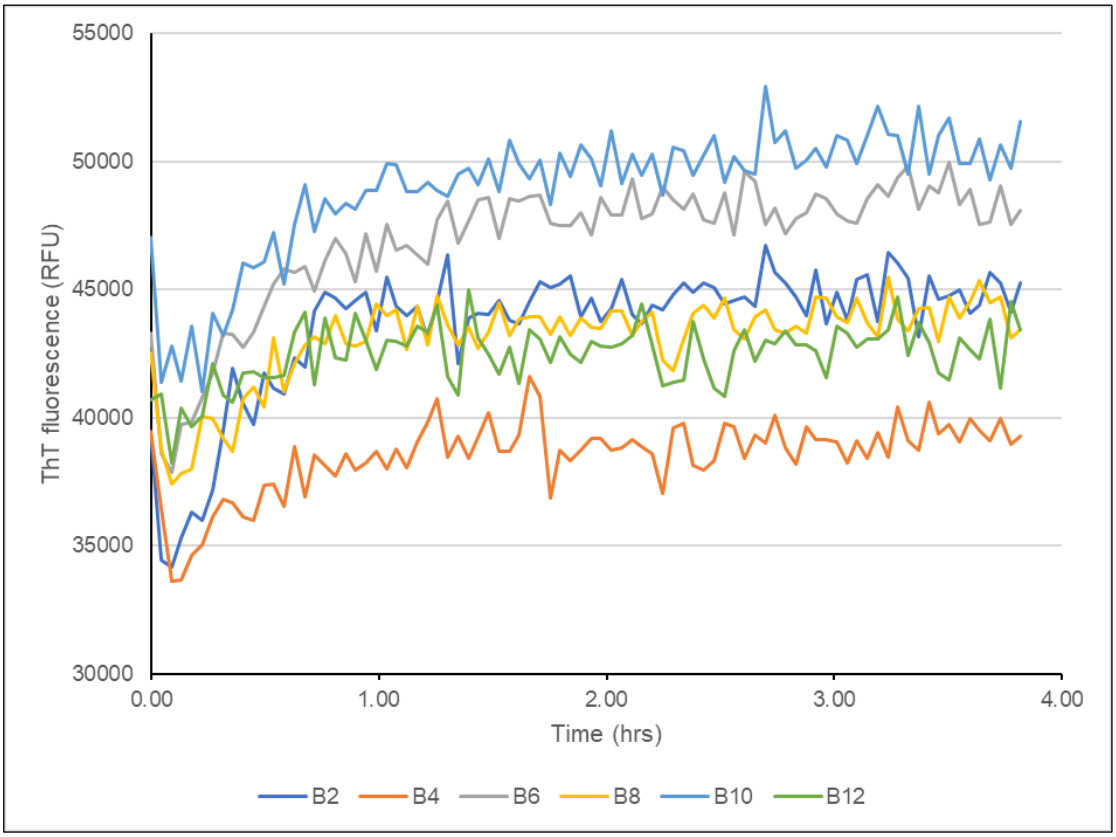


Figure 6.6. ThT TDP-43 CTF aggregation assay of repeat wells. 4 hour ThT aggregation assay of 20 μM TDP-43 CTF in 1.5 M urea, showing repeat wells (B2-B12) of identical starting reactions.

An alternative to this assay was therefore required to more reliably detect and monitor aggregation, without the level of data spread seen with ThT. ProteoStat® (Enzo Life Sciences) is an alternative aggregate detection dye that is reported to have greater sensitivity to levels of protein aggregates and can detect both amyloid and non-amyloid protein aggregates in cell extracts and inside cells (Oshinbolu *et al*., 2018; Navarro and Ventura, 2014).TDP-43 CTF protein stock was therefore diluted to 20 μM (concentration as recommended by the ProteoStat® manual) (1.5 M urea), and aggregated with 20 μM ProteoStat® for initially 7 hours in a Tecan Infinite plate reader - the ThT data suggested the protein continued to aggregate beyond the 4 hours first tested. As shown in Fig. 6.7, the ProteoStat® fluorescence of repeat wells containing aggregating TDP-43 CTF was far less spread than the reactions carried out with ThT, with the maximum difference between wells being approximately 1000 RFU. This was repeated across different reactions, indicating ProteoStat® could be more reliably used to monitor TDP-43 CTF aggregation than ThT.

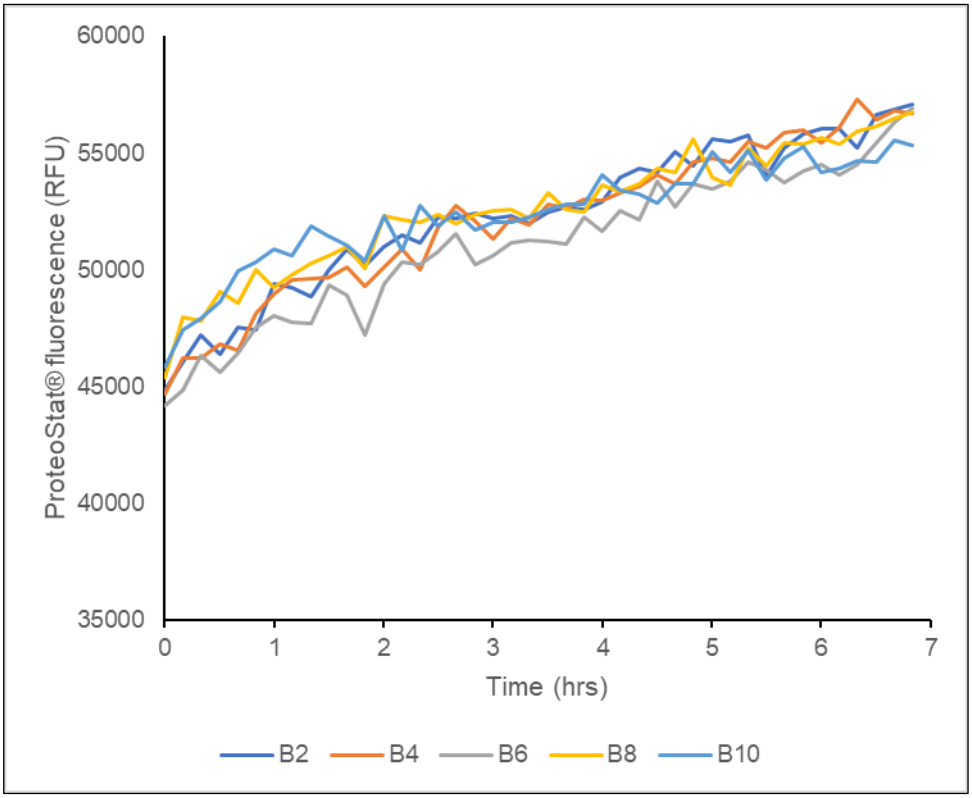


Figure 6.7. ProteoStat® TDP-43 CTF aggregation assay of repeat wells. 7 hour ProteoStat® aggregation assay of 20 µM TDP-43 CTF in 1.5 M urea, showing repeat wells (B2-B10) of identical starting reactions.

6.5 Optimisation of the TDP-43 CTF aggregation assay

Consideration of the reaction conditions for aggregation had to be made. One such consideration was the concentration of TDP-43 CTF to use in the assay. Prasad *et al*., (2016), in their ThT aggregation assays, used a protein concentration of 400 µM- the protein used in this method was however purified by only a single IMAC step, meaning the protein sample was likely more contaminated than that produced by the TDP-43 CTF purification presented in this chapter. A concentration of 400 µM could not be reached by our method. Contrastingly Johnson *et al*., (2009), used a protein concentration of only 3 µM in their aggregation assays, however their assay was based on turbidity and was not a fluorescence-based assay. The concentration suggested for

ProteoStat® assays by EnzoLifeSciences is 20 μM. This is similar to the concentrations of Aβ40, Aβ42 and Ure2 (40 μM, 15 μM and 20 μM respectively) used in ThT assays in Xue *et al*., (2017), 50μM α-synuclein used in Wördehoff and Hoyer, (2018) and 20 μM tau Δ1-250 used previously in our lab (Aggidis, 2019). Another consideration is the final concentration of urea in the aggregation reaction following dilution from the 4 M urea protein stock. Prasad *et al*., (2016), diluted a 4 M urea stock to an end concentration of 2.5 M in the aggregation reaction, with a protein concentration of 400 μM. Due to the lower concentrations of protein to be used in the ProteoStat® assay this urea concentration may be too high. With these considerations in mind, initial aggregation reaction conditions were screened. First, using the EnzoLifeSiences recommended protein concentration of 20 μM, different urea concentrations were tested, including the 2.5 M urea used in Prasad *et al*., (2016). Results of three urea concentrations are shown in Fig. 6.8. At the 2.5 M urea concentration, there was only a low level of ProteoStat® fluorescence, with minimal increase with time, suggesting the urea concentration was inhibiting aggregation. 2 M urea demonstrated higher fluorescence, with a slight increase over time, while 1.5 M final urea concentration exhibited the most fluorescence with a noticeable increase of almost 30 000 RFU over 18 hours. With reduction of the endpoint urea further to 1 M, the protein noticeably rapidly precipitated before it could be pipetted into the microplate for the assay, suggesting this urea concentration was not suitable, and the dilution of the urea destabilised the protein too rapidly. These assays showed that there was a rapid increase in fluorescence during the dead time of the experiment (i.e., during dilution and before the plate could be placed in the reader). This is demonstrated by Fig. 6.9, showing that dilution to 1.5 M urea causes a dramatically increased level of ProteoStat® fluorescence when compared to sample diluted with a 4 M urea or a control sample of 4 M urea buffer alone.

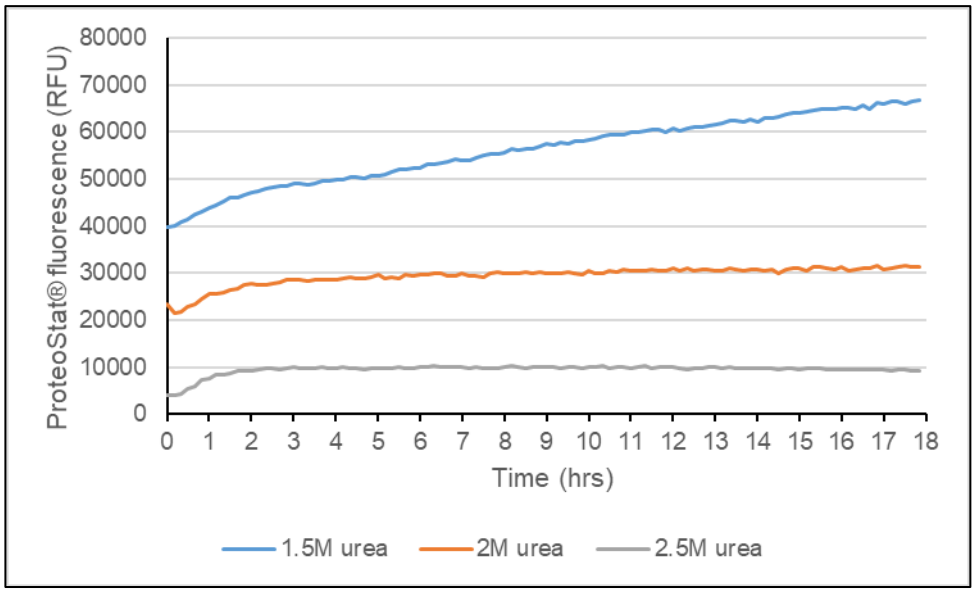


Figure 6.8. Effect of urea concentration on aggregation of a constant TDP-43 CTF concentration. 18 hour ProteoStat® aggregation assay of 20 μM TDP-43 CTF in varying final urea concentrations. Lines show mean averages of triplicate repeats of reactions. Blue- 1.5 M urea, orange- 2 M urea, grey- 2.5 M urea. Lines represent mean averages of triplicate reactions.

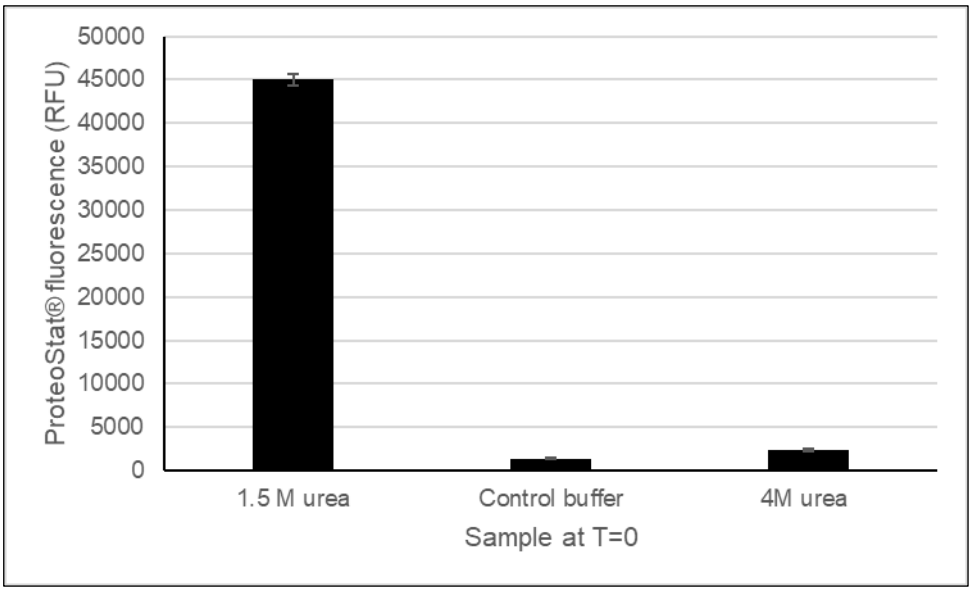


Figure 6.9. Effect of urea dilution on assay start point fluorescence. ProteoStat® fluorescence of samples at T=0 (start point) of reactions diluted to 1.5 M urea, maintained in 4M urea or a sample of 4 M urea buffer with no TDP-43 CTF present. Protein containing samples consisted of 20 μM TDP-43 CTF. Reactions were conducted in triplicate, error bars represent standard deviation of the mean.

To see if this initial “burst” of fluorescence could be avoided, the TDP-43 CTF concentration used in the assay was varied while maintaining 1.5 M urea. The results of this are shown in Fig. 6.10. Reduction in protein concentration did lead to a reduction in the initial fluorescence increase, but as shown, the fluorescence did not appear to be increasing further with time for the 10 µM and 5 µM TDP-43 CTF reactions. This could be due to the urea concentration being too high for these protein concentrations to aggregate, or simply a lack of detection of aggregation due to the reduced concentration below that recommended for ProteoStat® by the manufacturer. From these data, across repeated runs of the assay, the best assay conditions found for monitoring protein aggregation with ProteoStat® were to use 20 µM purified TDP-43 CTF in 1.5 M final urea concentration.

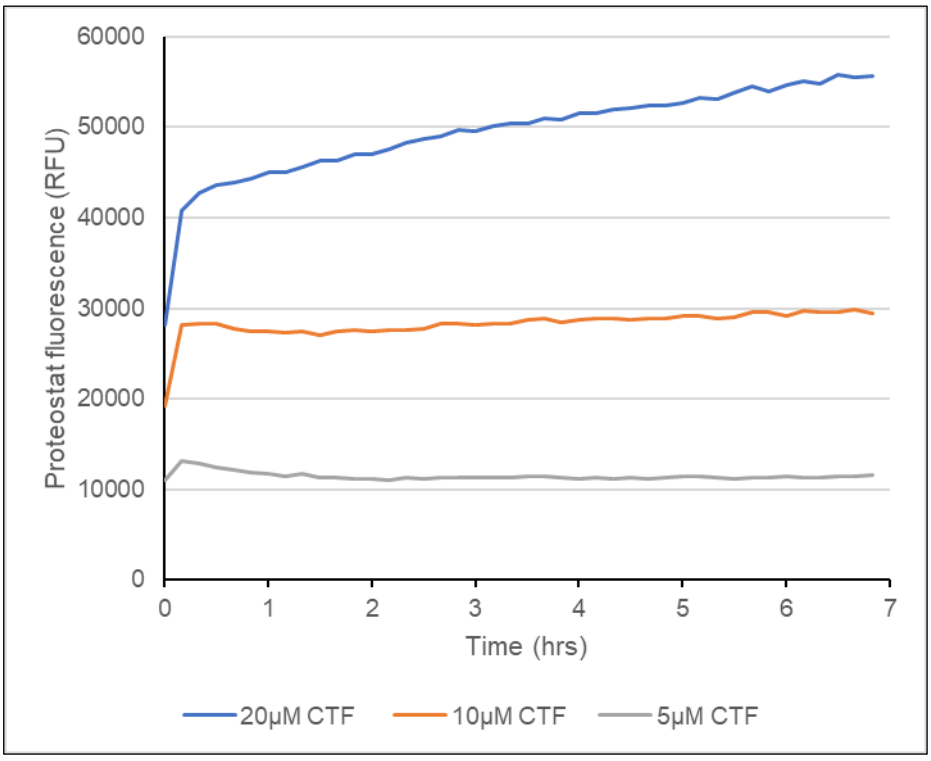


Figure 6.10. Effect of varying TDP-43 CTF concentration on ProteoStat® fluorescence at constant urea concentration. 7 hour ProteoStat® aggregation assay of varying concentrations of TDP-43 CTF diluted to 1.5M urea final. Blue- 20 µM TDP-43 CTF, orange- 10 µM TDP-43 CTF, grey- 5 µM TDP-43 CTF. Lines represent mean averages of triplicate reactions.

Finally, the duration of the experiment was considered. As shown in Figs 6.7, 6.8 and 6.10 so far aggregation time was varied from 4 to 18 hours. Further reactions, as shown in Fig 6.11 demonstrated that aggregation continued even past 48 hours. Going forward, 18 hours was decided upon as the duration of aggregation assays, for practical reasons of plate reader access and usage.

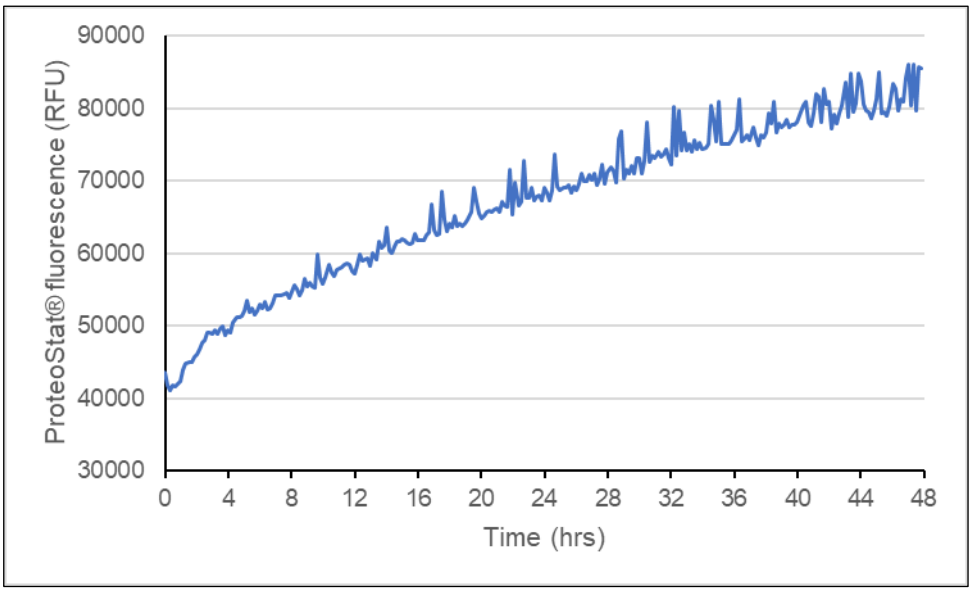


Figure 6.11. 48 hour ProteoStat® aggregation assay. Aggregation of 20 μM TDP-43 CTF in 1.5 M urea final concentration. Line represents the mean average of triplicate reactions.

6.6 TEM of TDP-43 CTF aggregates

Aggregates of TDP-43 formed in cells (Cascella *et al*., 2022), by full length, His-tagged TDP-43 in vitro (Cascella *et al*., 2022) and those identified from ALS and FTLD patients (Arseni *et al*., 2021) do not necessarily demonstrate classical amyloid morphology. Structures of aggregates formed from TDP-43 fragments and peptides differ further; by TEM imaging, Saini and Chauhan (2011) showed “worm-like” fibrils of a TDP-43 CTF construct (220-414) (which were also not ThT fluorescent), but also amyloid-like fibrillar aggregates from peptides consisting of amino acids 246-258 and 311-323. Aggregates of TDP-43 CTFs with deletions of amino acids 311-320, and of both 246-255 and 311-

320 together, appeared instead to be highly amorphous and granular in their appearance (Saini and Chauhan, 2011). Full length and mutant full-length proteins have been shown by TEM imaging to form multiple different structures including granular oligomers, small filamentous structures and large thread-like structures (Johnson *et al*., 2009) and filamentous structures with a granular surface coating (Kumar *et al*., 2023).

The products of the aggregation reactions were therefore imaged using transmission electron microscopy (TEM). Reaction endpoints were removed from the microplate wells and adsorbed onto formvar/carbon coated copper mesh grids (300 mesh), stained with phosphotungstic acid (PTA), and imaged using the JEOL JEM 1010. Images from various timepoints are shown in Figs. 6.12, 6.13 and 6.14. As shown across all the images, a consistent aggregate morphology was observed with the TDP-43 CTF aggregated under the reaction conditions used. The aggregates appear with a fairly straight, rod like morphology, with curved ends, up to approximately 1 μm in length and up to approximately 100 nm in width. The aggregates appear to cluster into larger masses at the later timepoints of 18 and 48 hours. In fact, following 18 hours of aggregation, solid aggregates were visible to the naked eye in the microplate wells.

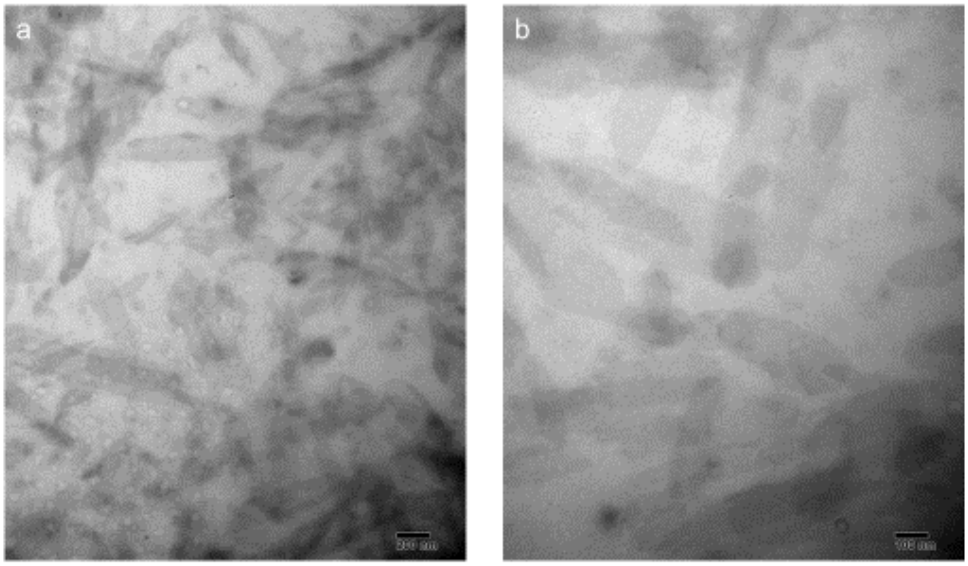


Figure 6.12. TEM images of TDP-43 CTF aggregation reaction samples after 5 hours of aggregation. Imaged at (a) 30 000x (200 nm scale bar) and (b) 50 000x (100 nm scale bar) magnification.

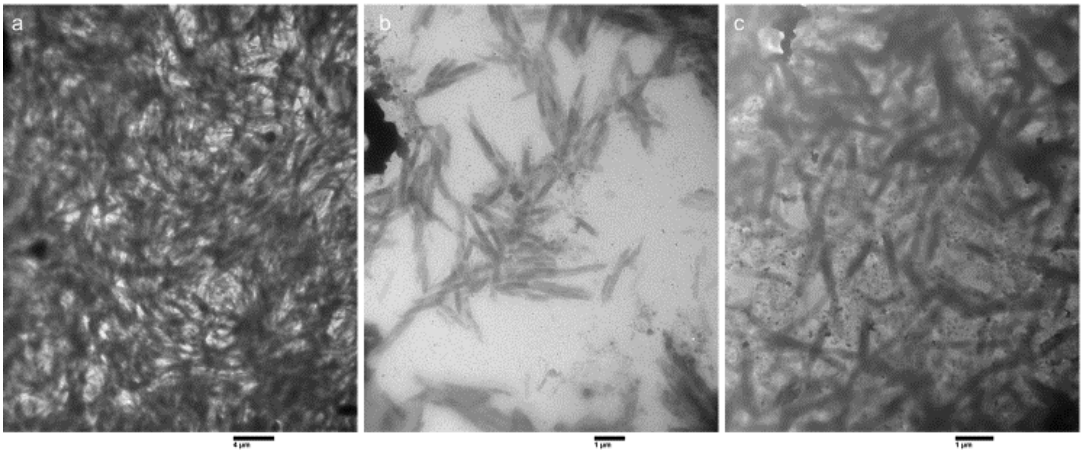


Figure 6.13. TEM images of TDP-43 CTF aggregation reaction samples after 18 hours of aggregation. Imaged at (a) 4000x (4 µm scale bar) (b) 12 000x (1 µm scale bar) and (c) 15 000x (1 µm scale bar) magnification.

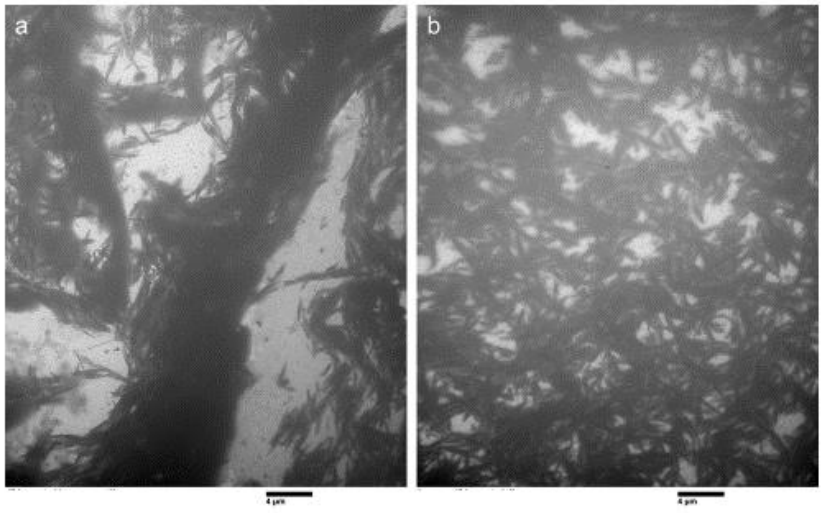


Figure 6.14. TEM images of TDP-43 CTF aggregation reaction samples after 48 hours of aggregation. Both (a) and (b) imaged at 4000x magnification (4 µm scale bar).

6.7 Conclusion

This chapter presents the development of a routine method for purifying TDP-43 CTF (208-414) and monitoring aggregation. The method proceeds by purification under denaturing conditions by IMAC and SEC, followed by concentration to produce a stock of protein in 4 M urea that can be stored at -20ºC before use.

This protein can be aggregated by dilution of the 4 M urea down to 1.5 M followed by incubation with shaking. Concentrations of urea and protein to use in the assay were determined empirically. The aggregation process can be monitored reliably using the fluorescent dye ProteoStat®, but not by using Thioflavin T. Aggregation begins immediately on dilution of the 4 M urea, and continues beyond 48 hours of incubation time, though for routine use of the assay in the laboratory, 18 hours was selected as the final assay length.

Transmission electron microscopy identifies a consistent morphology of aggregates produced by this method, which appear rod-like and group into larger masses at the later timepoints.

7. Design and testing of TDP-43 aggregation inhibitors

7.1 Introduction

As discussed in the literature review, one potential therapy target for ALS and FTLD could be TDP-43 aggregation. Small molecule inhibitors could potentially prevent the toxic TDP-43 aggregation seen in these diseases. Although some molecules in the literature such as AIM4, QBP1, antibodies and short peptides have shown potential as aggregation inhibitors, they are still in the early stages of development (Liu *et al.*, 2013; Prasad *et al*., 2016; Mompeán *et al*., 2019; Esposto & Martic, 2021; Kamagata *et al*., 2023). Testing and development of further molecules is therefore useful to produce more potential aggregation inhibitors for therapeutic use.

This chapter will present two approaches to identifying and developing aggregation inhibitors. Firstly, the repurposing of molecules identified in the literature to have inhibitory ability against the aggregation of other proteins, that have not yet been tested for their potential against TDP-43. A second approach is to design inhibitors specifically targeting regions of TDP-43 known or predicted to have a role in aggregation. Potential inhibitors will be tested in the TDP-43 CTF aggregation assay developed in the previous chapter.

7.2 Repurposing of generic aggregation inhibitors

One approach to inhibiting the aggregation of TDP-43 could be to re purpose molecules that have been identified to have anti-aggregation propensity against other proteins. Indeed, this was the approach for the development of AIM4 (Prasad *et al*., 2016), who repurposed acridine derivatives that had been shown to have aggregation inhibitory potential against several other proteins such as insulin and the prion PrP^{SC}.

7.2.1 Investigation of the inhibition of aggregation of TDP-43 CTF by repeat polypeptides

One such type of molecule are repeat polypeptides. Poly-L-arginine and Poly-L-ornithine are cationic polypeptides that have shown anti-aggregation potential against Tau peptides, Tau mutant protein and mutant p53 aggregation (Nadimidla *et al*., 2017; Chen *et al*., 2017). As a polyanionic alternative, poly-L-glutamic acid was also investigated. The structures of these molecules are shown in Fig. 7.1.

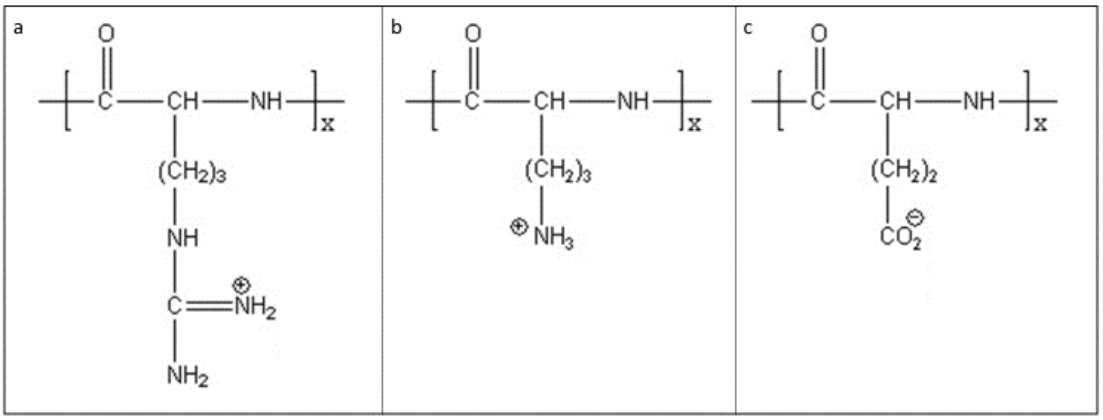


Figure 7.1. Repeat poly-peptide structures. a- Poly-L-arginine, b- poly-L-ornithine and c- poly-L-glutamic acid.

These ionic polypeptides, (5000-15,000 kDa), were tested for TDP-43 CTF aggregation inhibition potential. They were initially tested at equimolar concentrations to TDP-43 CTF in the ProteoStat® assay. Across repeat assays, poly-ornithine was found to accelerate the aggregation of the TDP-43 CTF to a higher fluorescence endpoint than control samples. Polyarginine and polyglutamic acid showed variation in their effect on TDP-43 CTF aggregation- in some cases as in Fig. 7.2 (a), they led to an increase in aggregation as indicated by ProteoStat® fluorescence, but in others as in Fig. 7.2 (b), they had either a minimal effect or led to marginal reduction in fluorescence. However as shown by the error bars, there was a large spread of the data, overlapping with the control reactions, suggesting there was a large variability in the effect of the poly-ions on the TDP-43 CTF aggregation. On multiple occasions, it did appear that the poly-cations poly-L-arginine and poly-L-ornithine caused a reduction in the initial “burst” of fluorescence increase in the dead time of the experiment. However, this may not be due to a stabilising factor of these poly-cations, as it was noticed that on dilution with HEPES buffer to begin the reaction, these reactions would appear to begin to precipitate. This is demonstrated by the start point absorbances at 395 nm shown in Fig. 7.3. Interestingly, the poly-anionic poly-L-glutamic acid did not cause such precipitation and in fact demonstrated a lower initial absorbance than the TDP-43 CTF control. Furthermore, poly-L-arginine, with a greater positive charge than poly-L-ornithine, demonstrated a higher absorbance. Together, these results may suggest a charge related precipitation effect by the poly-cationic peptides.

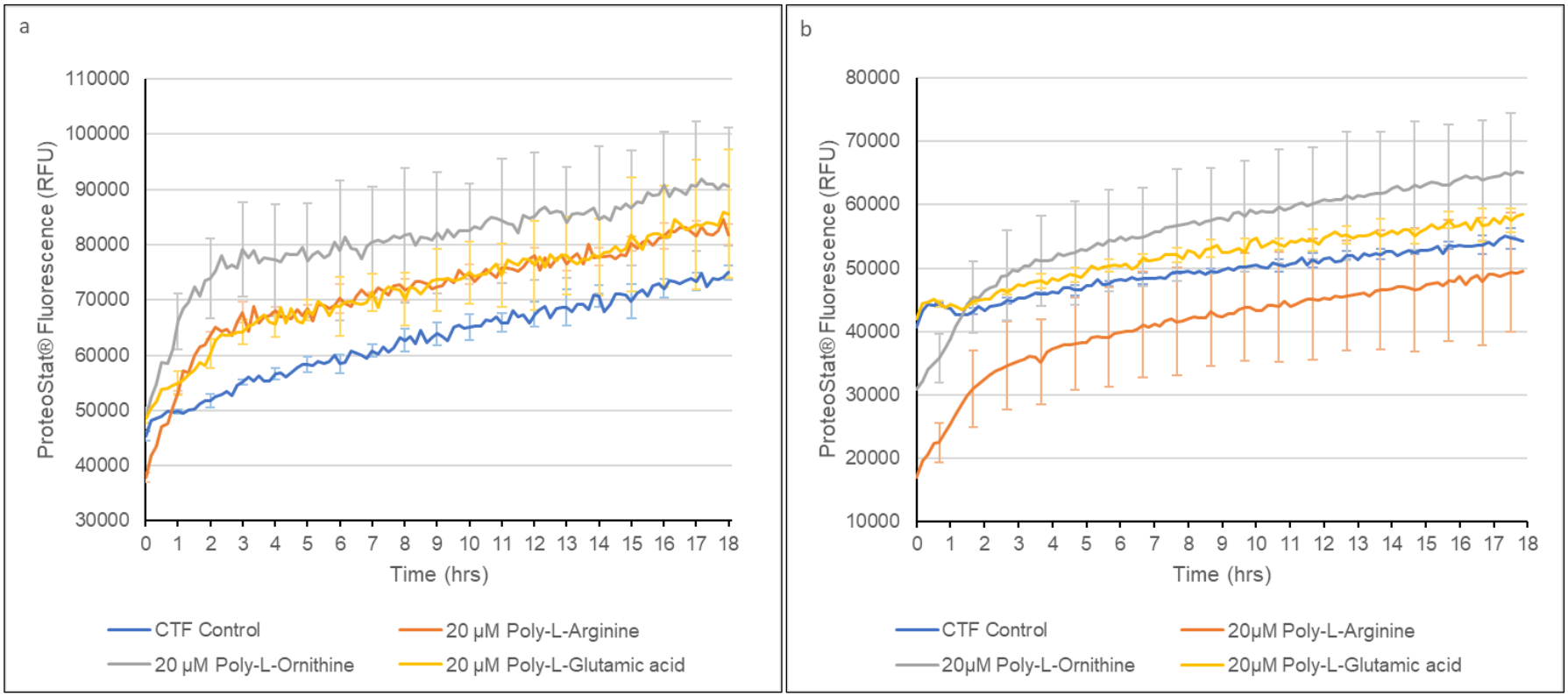


Figure 7.2. The effect of repeat polypeptides on the TDP-43 CTF aggregation assay. ProteoStat® aggregation assay, showing ProteoStat® fluorescence with time of aggregating 20 µM CTF with and without the addition of 20 µM poly-L-arginine, poly-L-ornithine and poly-L-glutamic acid. (**a**) and (**b)** are repeats of the same reaction conditions. Lines represent mean fluorescence of triplicate reactions and error bars represent standard deviation of the means.

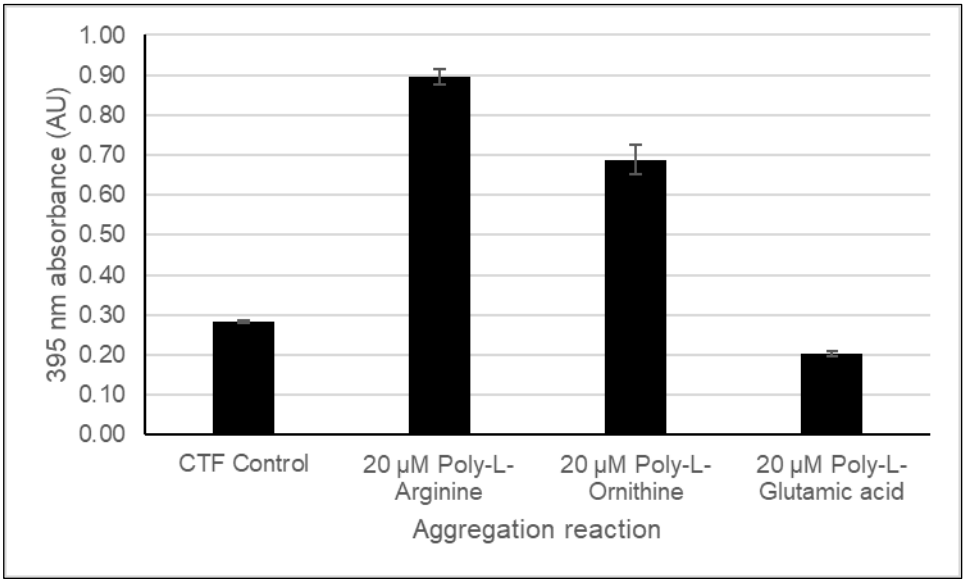


Figure 7.3. The effect of repeat polypeptides on TDP-43 CTF aggregation reaction turbidity. 395 nm absorbance of start point aggregation reactions for untreated CTF control with and without the addition of 20 µM poly-L-arginine, poly-L-ornithine and poly-L-glutamic acid. Bars represent mean absorbance of triplicate reactions with error bars representing standard deviation of the means.

Endpoints of aggregation reactions were imaged by TEM, examples of which are shown in Fig. 7.4. The inclusion of the poly-ionic peptides did not appear to alter the morphology of the aggregates formed.

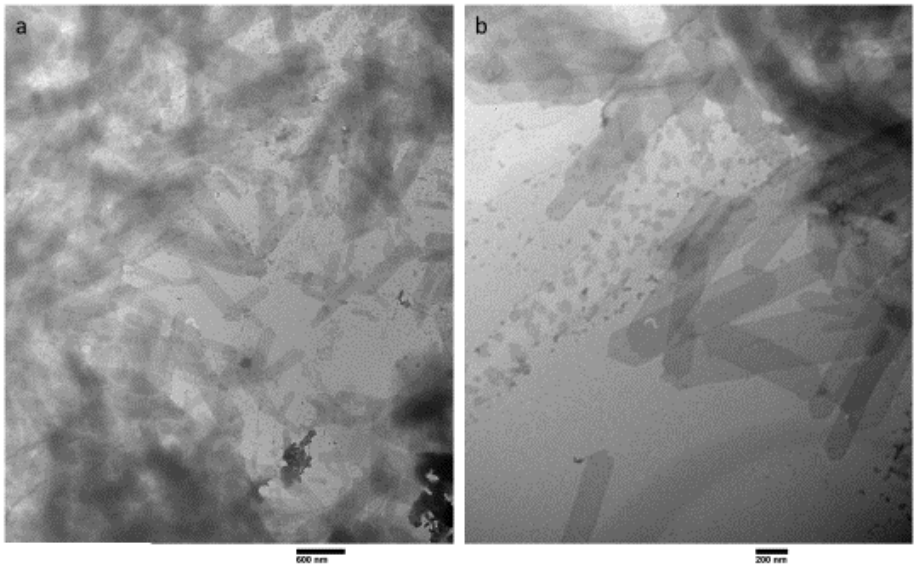


Figure 7.4. Representative TEM images of endpoint aggregation reactions for CTF treated with poly-ionic peptides. a- endpoint sample of a poly-L-arginine reaction, 25 000x magnification, scale bar represents 600 nm. b- endpoint sample of a poly-L-ornithine reaction, 50 000x magnification, scale bar represents 200 nm.

7.2.2 Investigation of the inhibition of aggregation of TDP-43 CTF by polyphenols

A second class of compounds identified to have general aggregation inhibitor potential against multiple proteins is the polyphenols. Two examples of these are curcumin and rosmarinic acid, the structures of which are shown in Fig. 7.5. Curcumin, a turmeric extract, has been shown to have anti-aggregation ability against amyloid-beta, tau and α-synuclein (Yang *et al*., 2005; Pandey *et al*., 2008; Rane *et al*., 2017). Derivatives of curcumin, dimethoxy curcumin and monocarbonyl dimethoxycurcumin C, have been shown to reduce hyperexcitability induced by TDP-43 expression in a motoneuron-like cellular model of ALS and reduce aggregation of a 25 kDa TDP-43 fragment termed TDP-25 in NSC34 cells respectively (Duan *et al*., 2013; Dong *et al*., 2014). Rosmarinic acid is a polyphenol extracted from *Lamiaceae* (sage family) plants or *Isodon*

japonicus. It, and its derivatives, have been shown to have anti aggregation ability against amyloid and hiAPP peptides (Taguchi *et al*., 2017; Sun *et al*., 2019).

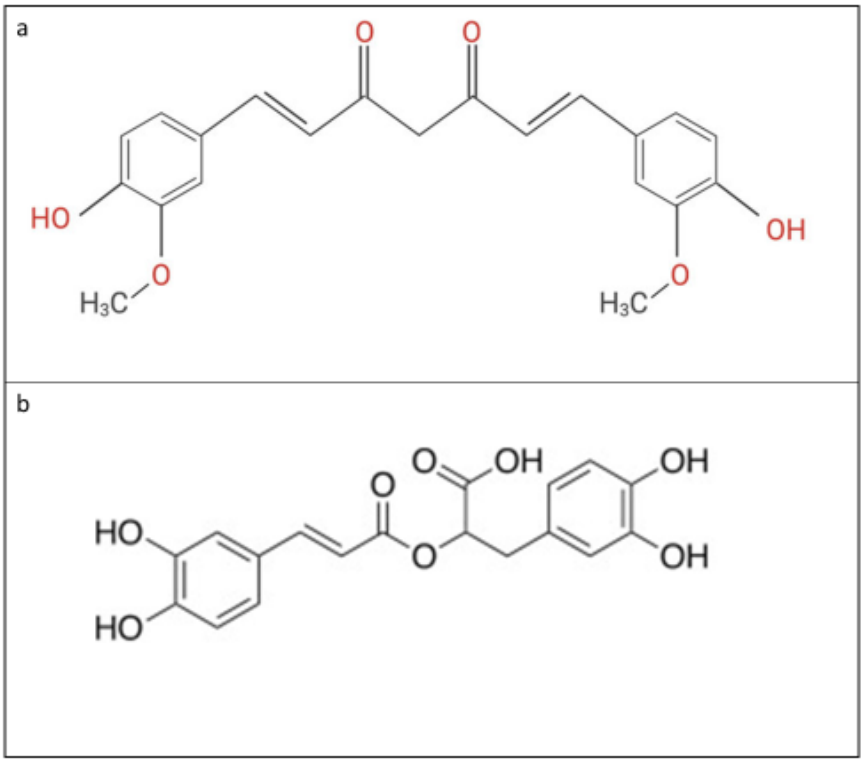


Figure 7.5. Polyphenol structures. The structure of a- curcumin (from Abd. Wahab *et al*., 2020) and b- rosmarinic acid (from Taguchi *et al*., 2017).

These compounds were therefore tested for any potential anti-aggregation ability against TDP-43 CTF at multiple concentrations with a maintained 2% final DMSO concentration in the ProteoStat® assay. The addition of curcumin to the aggregation reaction appeared to dramatically reduce ProteoStat® fluorescence, in a dose dependent manner (Fig. 7.6). However, several observations were made from these reactions; in the 100 μM and 200 μM curcumin reactions, the curcumin could be seen to precipitate on dilution to 2% DMSO in the aggregation buffer. Given the highly pigmented yellow/orange colouring of curcumin, this may have affected the assay. Furthermore, in the reactions containing lower curcumin concentrations (which were

also yellow in colour), protein aggregates were visible to the naked eye as in untreated samples. It was therefore considered that curcumin could interfere with the aggregation reactions in some way, by either affecting the excitation/emission with some form of quenching effect, or by interfering with the binding of ProteoStat® to aggregates. To test this, 20 μM curcumin was added to a sample of 20 μM of pre-aggregated lysozyme pre-incubated with ProteoStat® (aggregated lysozyme is provided as a positive control in the ProteoStat® kit). As shown in Fig. 7.7, curcumin dramatically reduced the fluorescence of the aggregated lysozyme compared to a sample with just 2% DMSO added. It therefore appeared that curcumin interferes with the assay.

Rosmarinic acid also reduced the ProteoStat® fluorescence in the TDP-43 CTF aggregation assay (Fig. 7.8), again with increasing concentrations leading to increased reduction. Furthermore, for concentrations higher than 20 μM, the fluorescence decreased over time. Unlike curcumin, rosmarinic acid is not visibly coloured and did not precipitate with the DMSO concentration reduction occurring on dilution. However, again there were still visible aggregates in the end point reactions, and so rosmarinic acid was also added to pre-aggregated lysozyme sample and the ProteoStat® fluorescence measured. Like curcumin, rosmarinic acid led to a reduction in the fluorescence of the aggregated lysozyme compared to control sample and so also appeared to interfere with the assay. However, this may not explain the decrease in fluorescence over time seen with the rosmarinic acid reactions.

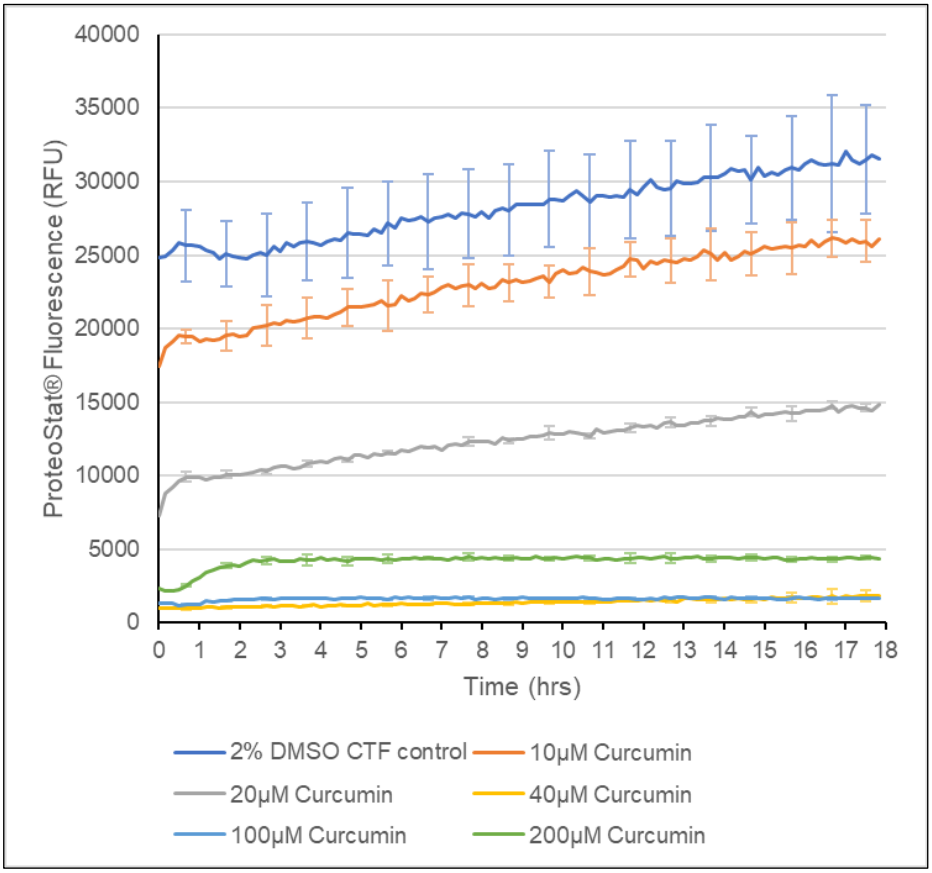


Figure 7.6. The effect of curcumin on the TDP-43 CTF aggregation assay. ProteoStat® aggregation assay, showing ProteoStat® fluorescence with time of aggregating 20 μM TDP-43 CTF with and without the addition of various concentrations of curcumin. Lines represent mean fluorescence of triplicate reactions and error bars represent standard deviation of the means.

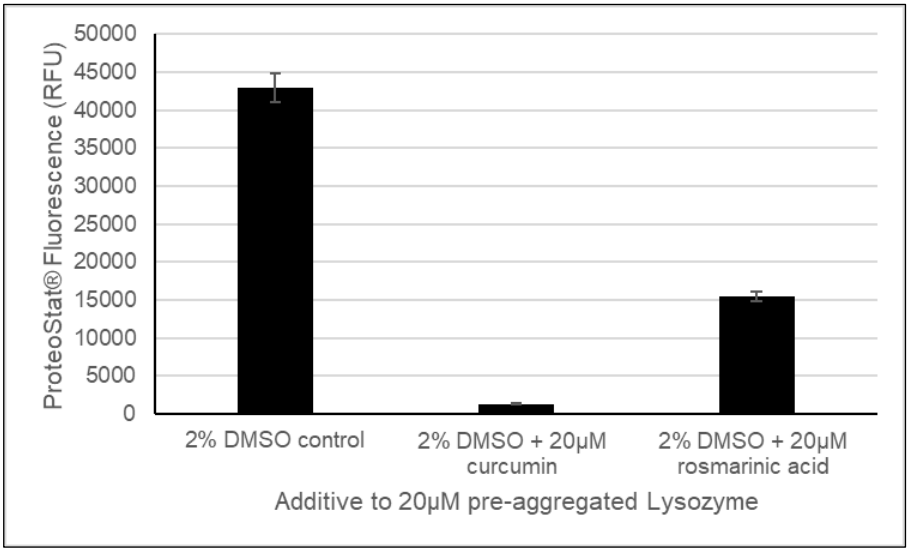


Figure 7.7. Interference of ProteoStat® fluorescence by curcumin and rosmarinic acid. The effect of the addition of 20 μM curcumin and rosmarinic acid on the ProteoStat® fluorescence of 20 μM aggregated lysozyme, with a control sample of 2% DMSO addition. Bars represent mean averages of triplicate reactions and error bars represent standard deviation of the means.

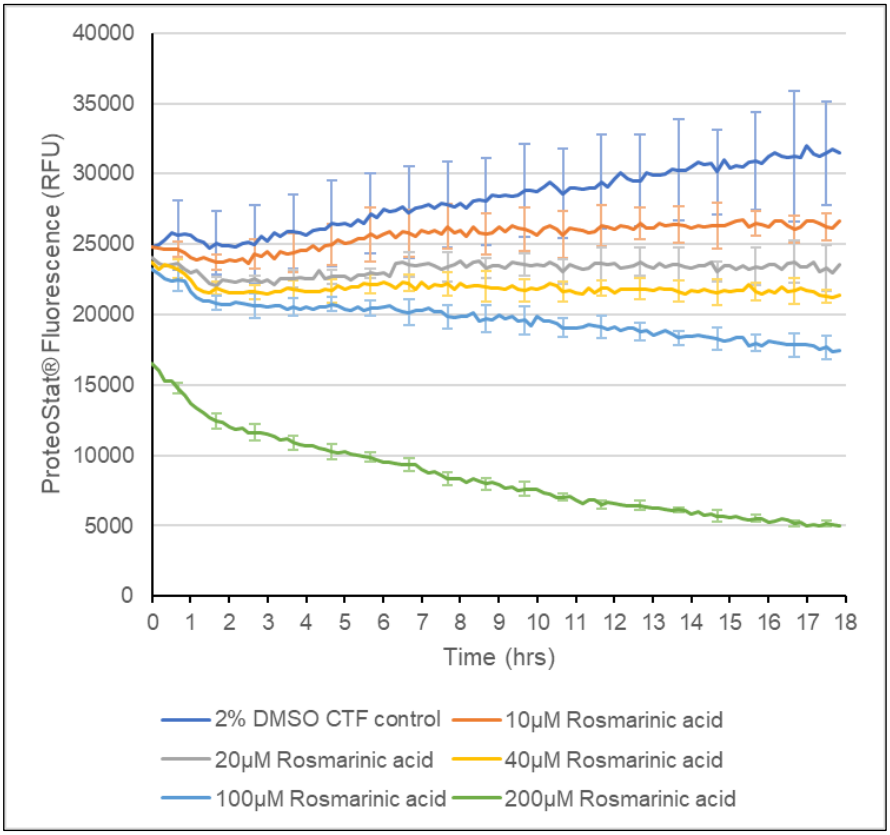


Figure 7.8. The effect of rosmarinic acid on the TDP-43 CTF aggregation assay. ProteoStat® aggregation assay, showing ProteoStat® fluorescence with time of aggregating 20 μM TDP-43 CTF with and without the addition of varying concentrations of rosmarinic acid. Lines represent mean fluorescence of triplicate reactions and error bars represent standard deviation of the means.

To see whether curcumin and rosmarinic acid had any effect on the aggregate morphology, end point samples were imaged by TEM. Representative images are shown in Fig. 7.9. For samples treated with curcumin, no alterations were seen in the morphology of the aggregates formed. For rosmarinic acid samples, the morphology of the aggregates also appeared the same, however there did appear to be many smaller “fragments” of aggregates, indicated with arrows in Fig. 7.9 (c). One suggestion for the reduction in ProteoStat® fluorescence with time could be some kind of fragmentation or break down of the aggregates, however there clearly was still a large amount of aggregate formation occurring.

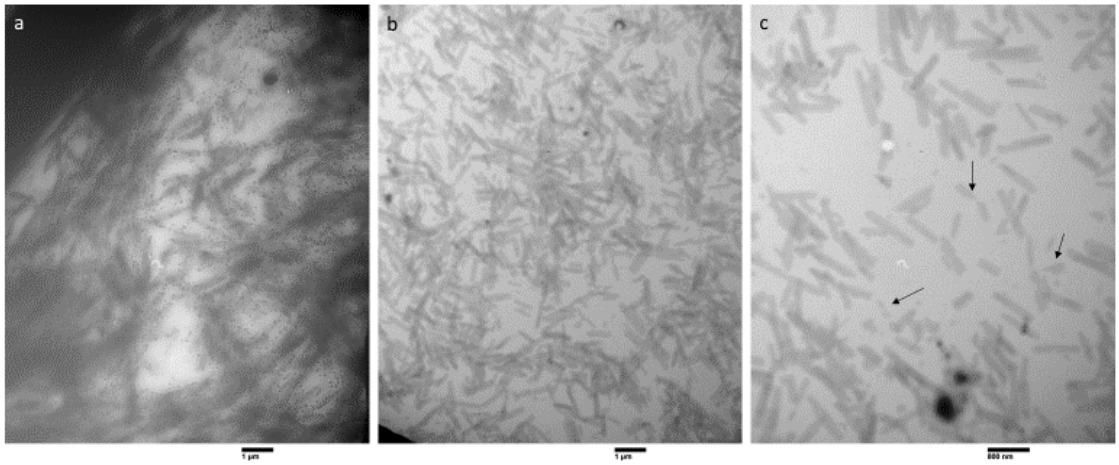


Figure 7.9. Representative TEM images of endpoint aggregation reactions for TDP-43 CTF treated with curcumin and rosmarinic acid. a- endpoint sample of a 20 μM curcumin reaction, 12 000x magnification, scale bar represents 1 μm. b- endpoint sample of a 200 μM rosmarinic acid reaction, 12 000x magnification, scale bar represents 1 μm. c- endpoint sample of a 200 μM rosmarinic acid reaction, 20 000x magnification, scale bar represents 800 nm. Arrows indicate apparent fragments of aggregate.

To identify if this could be a valid observation, TDP-43 CTF was aggregated for a longer time period in the presence or absence of rosmarinic acid, with fluorescence taken at timepoints of 24 hours and 48 hours rather than continual readings (Fig. 7.10). As shown, the fluorescence of the rosmarinic acid treated sample continued to decrease with more time, while the fluorescence of the TDP-43 CTF control continued to increase. TEM images were taken of reaction endpoints (Fig. 7.11). For the TDP-43 CTF control reactions, after 48 hours the aggregates with normal morphology had formed large masses, and while the rosmarinic acid treated sample had many aggregates with normal morphology, there were few areas of these larger masses identified. There were however some areas where the aggregates were predominantly granular in structure. Therefore, over a longer time period rosmarinic acid may have an effect on the aggregation of the TDP-43 CTF, however this is not conclusive- clearly

the normal aggregate morphology still occurs to a large extent, and for any noticeable change in TEM images rosmarinic acid had to be used at a 10 times molar excess to the TDP-43 CTF. From the extent of the reduction of fluorescence, more inhibition of aggregation or dissolution of aggregates would be expected- the reduction seen to below 4000 RFU is a similar level of fluorescence seen as the start point of a TDP-43 CTF sample diluted in 4 M urea buffer.

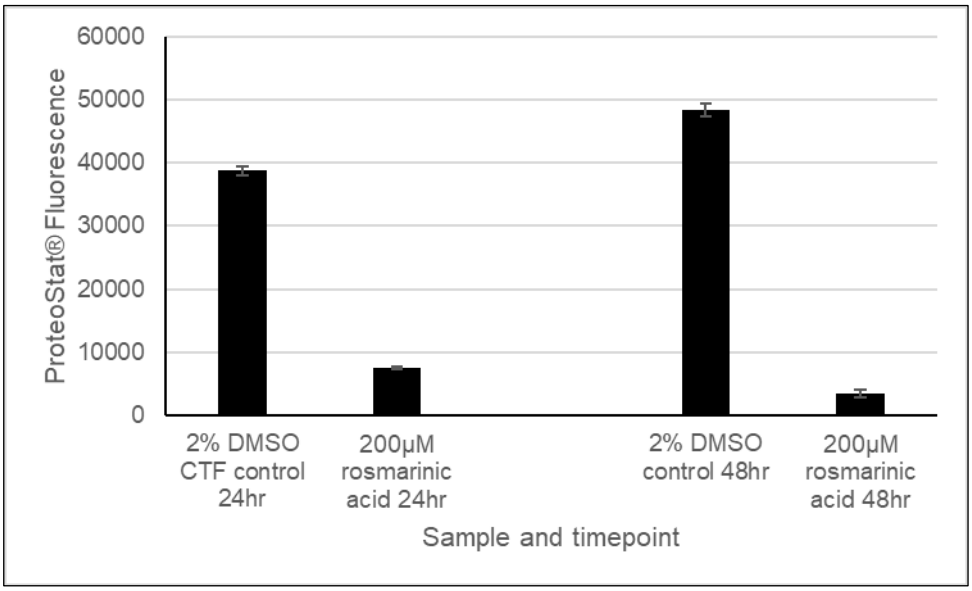


Figure 7.10. ProteoStat® fluorescence of 20 μM TDP-43 CTF with and without treatment with 200 μM rosmarinic acid following 24 and 48 hours of aggregation. Bars represent mean averages of triplicate reactions and error bars represent standard deviation of the means.

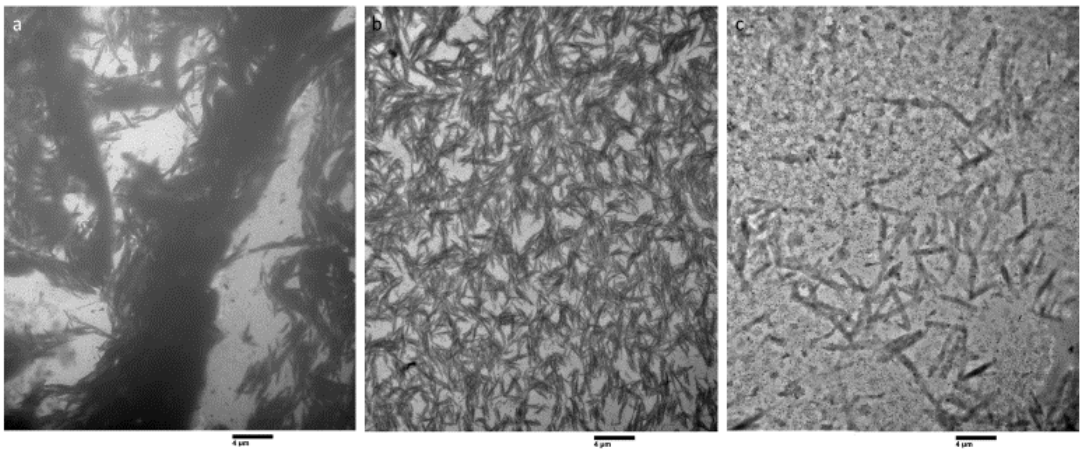


Figure 7.11. Representative TEM images of aggregation reactions for CTF following 48 hours of aggregation. a- endpoint sample of an untreated CTF control reaction b- endpoint sample of a 200 μM rosmarinic acid reaction. c- endpoint sample of a 200 μM rosmarinic acid reaction. All images taken at 4000x magnification and scale bars represent 4 μm.

7.2.3 Investigation of the inhibition of aggregation of TDP-43 CTF by Polyethylene Glycol-phosphatidylethanolamine (PEG-PE)

A final molecule identified in the literature is polyethylene glycol-phosphatidylethanolamine (PEG-PE). This is a hydrophilic polyethylene glycol moiety (PEG) conjugated to a hydrophobic phosphatidylethanolamine (PE), and readily forms micelles in solution. The PEG-PE micelle has been shown to inhibit the aggregation of hiAPP peptides (Fang *et al*., 2018) and also block the aggregation of denatured insulin and encourage its proper folding (Fang *et al*., 2016). The structure of PEG2000PE, as used in Fang *et al*., (2018) is shown in Fig. 7.12.

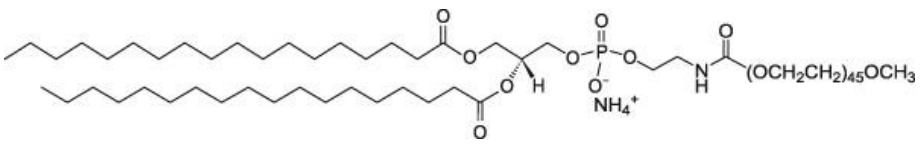


Figure 7.12. The structure of PEG2000PE ammonium salt. From *18:0 peg2000 PE* (2023) *Avanti Polar Lipids.*

PEG2000PE (Avanti Lipids) was resuspended in 50 mM HEPES pH 8 (the same as the dilution buffer for the aggregation reactions) and tested in the TDP-43 CTF ProteoStat® assay (Fig. 7.13). The lower concentrations (10 μM and 20 μM) of PEG2000PE did appear to reduce the levels of fluorescence, with minimal overlap in error bars with the control, as shown. In contrast, increased concentrations (40 μM and 200 μM) appeared to accelerate the aggregation detected in the assay. Despite the increased levels of fluorescence however, after an initial burst of approximately 3-4 hours the fluorescence did not increase any further.

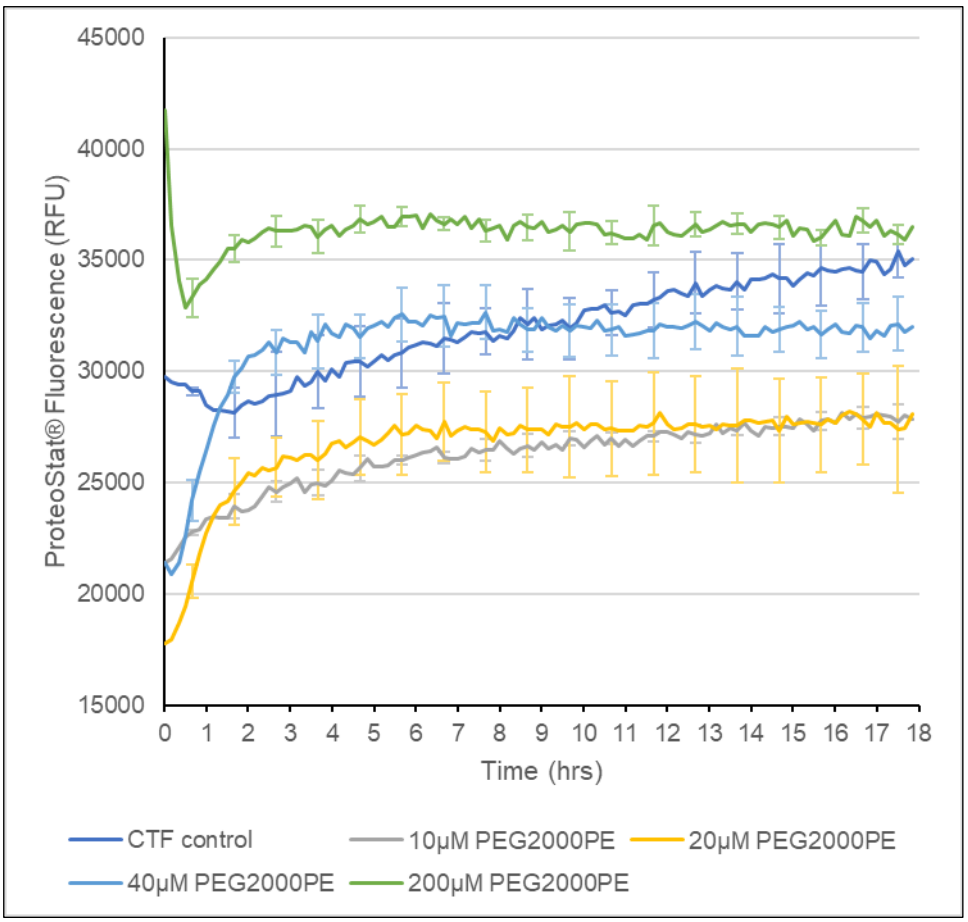


Figure 7.13. The effect of PEG2000PE on the TDP-43 CTF aggregation assay. ProteoStat® aggregation assay, showing ProteoStat® fluorescence with time of aggregating 20 μM TDP-43 CTF with and without the addition of varying concentrations of PEG2000PE. Lines represent mean fluorescence of triplicate reactions and error bars represent standard deviation of the means.

TEM imaging of the reaction endpoints did not however demonstrate a change to the usual morphology of the aggregates, but did show the presence of some unusual cube shape structures (Fig. 7.14)- it is unclear whether these are some alternative aggregates, simply artefacts in these samples or a product of the presence of the micellar PEG2000PE.

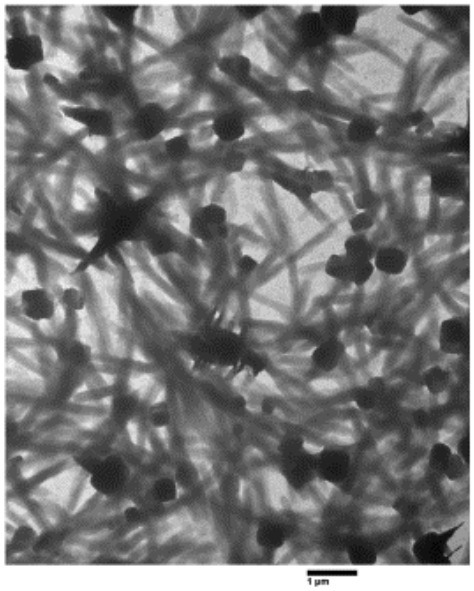


Figure 7.14. Representative TEM image of an endpoint aggregation reaction for TDP-43 CTF treated with 20 µM PEG2000PE. Imaged at 15 000x magnification and scale bar represents 1 µm.

7.3 Targeted peptide aggregation inhibitors

An alternative to the molecules investigated above, inhibitor molecules can be specifically designed. As particular regions of TDP-43 have been identified to be particularly aggregation prone, or are found in the core regions of aggregates, they are likely important regions of the protein to target. Another approach is to predict points of interaction, and target regions that theoretically could nucleate aggregation. Peptide based inhibitors were therefore designed and tested against the TDP-43 CTF.

7.3.1 First generation peptide inhibitor design

Previous work in the Allsop laboratory has demonstrated that short peptide inhibitors can act as aggregation inhibitors for neurodegenerative disease associated proteins. These include peptide and retro-inverso peptides for amyloid-beta and for tau protein, which were designed based on identified sequences of the protein that appear responsible for its self-association (Austen *et al*., 2008; Taylor *et al*., 2010; Aggidis, 2019). The premise is that these peptides block protein self-association by binding to their target sequences and preventing interaction with other protein molecules by way of added amino acids such as arginine that are highly charged and hydrophilic.

It was therefore considered whether further peptides, alternative to those designed in Liu *et al*. (2013), could be produced and optimised for the inhibition of TDP-43 aggregation. Two sequences of TDP-43 were found to be highly aggregation prone by Saini and Chauhan (2011). These sequences, when produced as synthetic peptides, produce ThT fluorescent, fibril like structures when aggregated and therefore may represent part of an amyloid-like core around which TDP-43 may aggregate. The two sequences are shown in Fig. 7.15. Interestingly, the MNFGAFSINPAMM sequence is central in the “double spiral fold” structure of TDP-43 aggregates extracted from ALS with FTLD patients, imaged by cryo-EM (Arseni *et al*., 2021). Therefore, peptide

inhibitors were designed based on these two aggregation regions. These first-generation inhibitors contained a hexapeptide core based on their respective target sequences, with Arg-Gly/Gly-Arg sequences at their termini. The inhibitor peptides, with their denominations, sequence and target sequence are shown in Table 7.1.

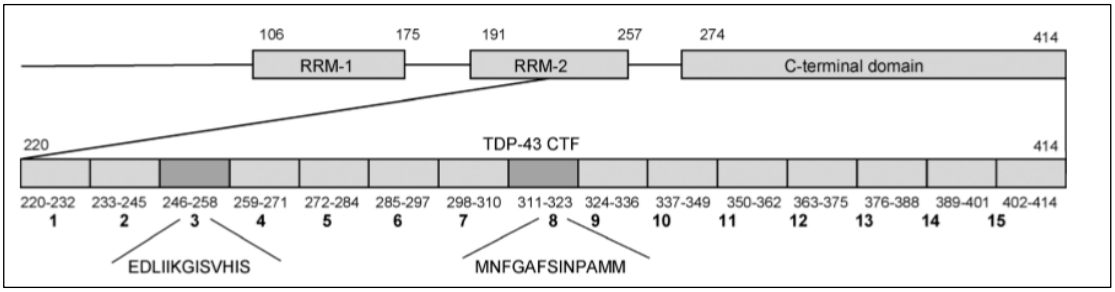


Figure 7.15. Two highly aggregation prone regions of TDP-43, residues 246-258 and 311-323. From Saini and Chauhan, (2011), from which the initial peptide aggregation inhibitors were designed.

Table 7.1. The first-generation peptide aggregation inhibitors and their TDP-43 aggregation regions targeted.

Peptide Inhibitor	Peptide Inhibitor sequence	TDP-43 aggregation sequence targeted	TDP-43 residue number targeted
MT01	H-Arg-Gly-Leu-Ile-Ile-Lys-Gly-Ile-Gly-Arg-OH	Leu-Ile-Ile-Lys-Gly-Ile	248-253
MT02	H-Arg-Gly-Ile-Ile-Lys-Gly-Ile-Ser-Gly-Arg-OH	Ile-Ile-Lys-Gly-Ile-Ser	249-254
MT03	H-Arg-Gly-Ile-Lys-Gly-Ile-Ser-Val-Gly-Arg-OH	Ile-Lys-Gly-Ile-Ser-Val	250-255

MT04	H-Arg-Gly-Phe-Gly-Ala-Gly-Ser-Ile-Gly-Arg-OH	Phe-Gly-Ala-Phe-Ser-Ile	313-318
MT05	H-Arg-Gly-Gly-Ala-Phe-Ser-Ile-Asn-Gly-Arg-OH	Gly-Ala-Phe-Ser-Ile-Asn	314-318
MT06	H-Arg-Gly-Ala-Phe-Ser-Ile-Asn-Pro-Gly-Arg-OH	Ala-Phe-Ser-Ile-Asn-Pro	315-320

7.3.2 Self-aggregation of first-generation peptide inhibitors

Given that the peptide inhibitors are designed based on TDP-43 aggregation sequences, they may therefore have some propensity to self-aggregate. If self-aggregation occurred, they would be unable to carry out their role as aggregation inhibitors by binding to the TDP-43 aggregation regions they target. In an attempt to prevent self-aggregation, the peptide inhibitors are designed with Arg-Gly/Gly-Arg sequences at their termini. This acts to prevent self-aggregation by repulsion between the charged arginine residues on each peptide inhibitor.

To confirm that the inhibitors do not self-aggregate, aggregation assays were carried out. The resuspended peptide inhibitors were subjected to Thioflavin T (ThT) assays, to identify any amyloid self-aggregation. As shown in Fig. 7.16 (a), there was no apparent increase in ThT fluorescence for any of the six inhibitors, suggesting that no spontaneous amyloid fibril formation occurred. The fluorescence was far lower than the positive control, a peptide of amino acid sequence VQIVYK (an aggregation sequence of Tau protein, known to aggregate in the presence of heparin), and similar to a negative control containing no peptide (HEPES+ThT). Similarly, as shown by Fig. 7.16 (b), when heparin (a polyanion used as an aggregation inducer in Tau studies) was added to the reaction, there was again no increase in ThT fluorescence of any of the peptide inhibitors, again suggesting no amyloid fibril formation.

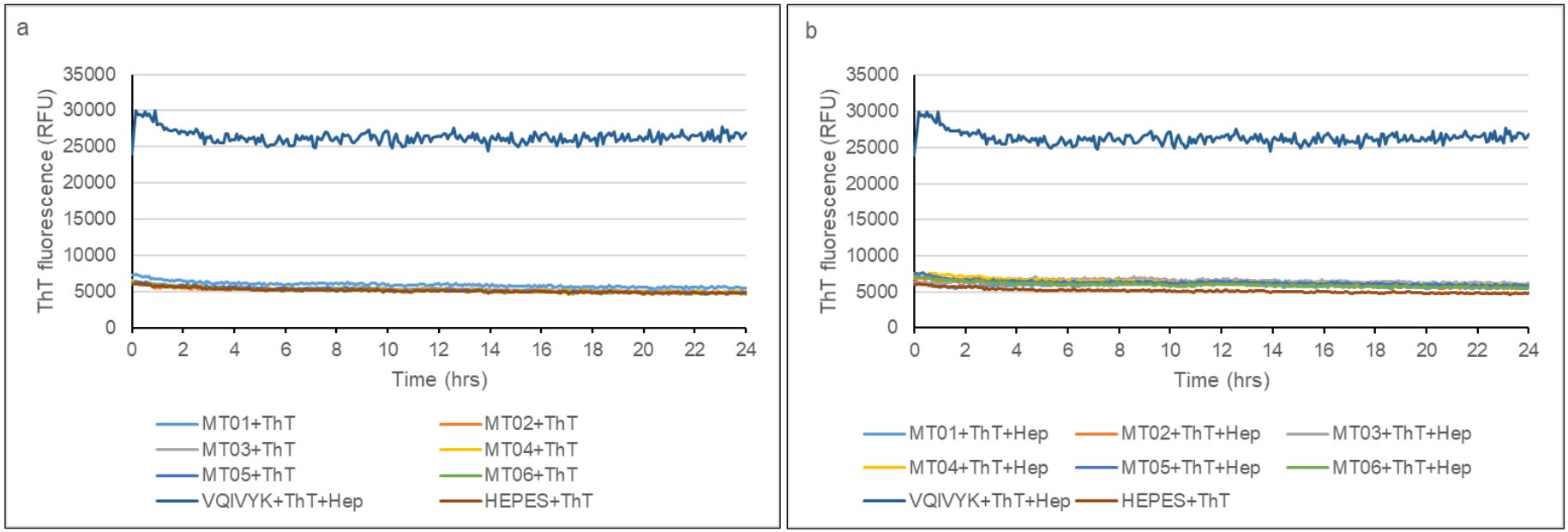


Figure 7.16. ThT assay for testing self-aggregation of first-generation peptide inhibitors. Relative fluorescence against time for each peptide inhibitor incubated with ThT (Thioflavin T) in the absence (**a**) or presence (**b**) of heparin. N=3 for all reactions.

In order to confirm the results of the ThT assays, 395 nm light absorbance was used as a second aggregation assay. Incubation with mixing was used to encourage any aggregation, and any increased absorbance at 395 nm was used as an indicator of aggregate formation. Fig. 7.17 shows the results of this absorbance assay. Once again, the inhibitors (with and without heparin) did not appear to show self-aggregation within 24 hours when compared to the positive control of the VQIVYK peptide aggregated with heparin. MT01 with heparin present showed some increase in absorbance (average 0.054 AU, n=3) compared to the other inhibitors and negative controls. Therefore, this inhibitor may show some self-aggregation when in the presence of the aggregation inducer heparin, however at a much lower level than the positive control.

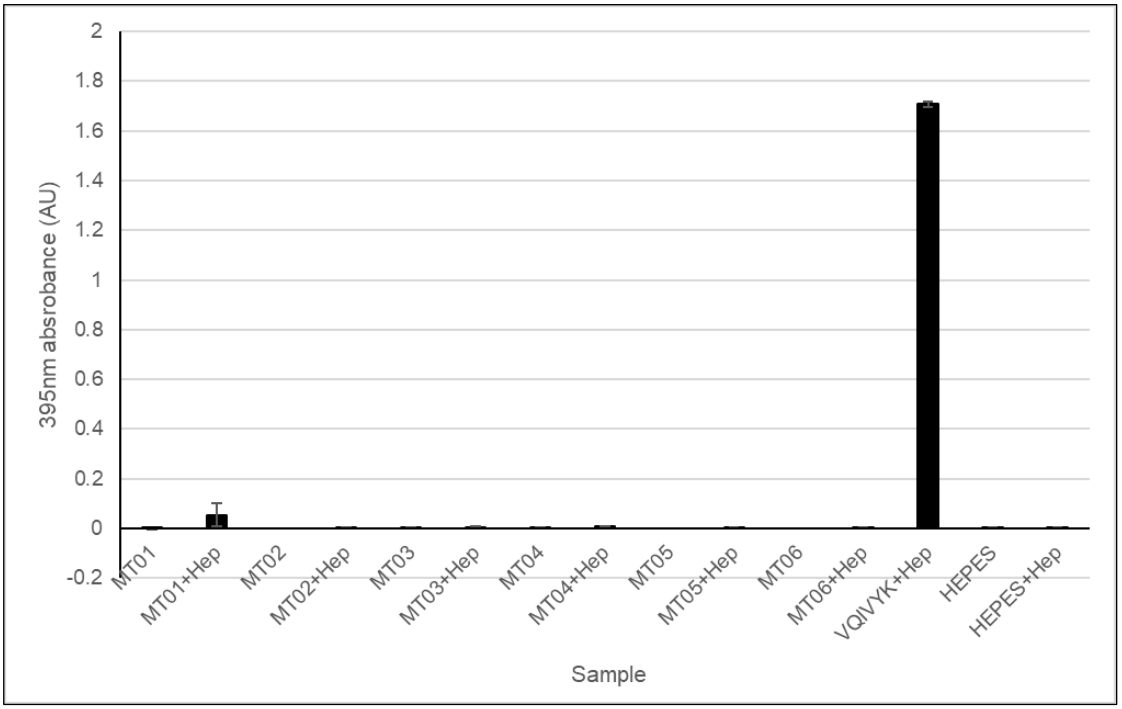


Figure 7.17. 395 nm absorbance for peptide inhibitor samples incubated with and without heparin (Hep). Error bars show standard deviation of the mean. N=3 for all samples.

7.3.3 Investigation of the inhibition of aggregation of TDP-43 CTF by first-generation peptide inhibitors

Given that the inhibitors did not appear to self-aggregate, they could be tested against aggregating TDP-43 protein. The ProteoStat® TDP-43 CTF aggregation assay developed in the previous chapter was used to assess any effect the inhibitors had on the detectable levels of aggregation, and TEM used to observe any potential alterations to the aggregate morphology. Peptides were initially used at equimolar concentration to TDP-43 CTF (20 μM). Across multiple tests, no consensus on the effectiveness of the inhibitors at this concentration was reached. As shown in Fig. 7.18 there was noticeable variation in the effect of the inhibitors. In some cases (Fig. 7.18 (a)) the MNFGAFSINPAMM targeting inhibitors MT04, MT05 and MT06 appeared to show an ability to reduce the extent of ProteoStat® detected aggregation compared to an untreated control, by up to approximately 10 000 RFU. However, in other assays, the inhibitors either had no effect or even accelerated aggregation compared to untreated TDP-43 CTF control (Fig. 7.18 (b)). Furthermore, there was no consistent pattern shown across the inhibitors- Fig. 7.18 (a) shows MT06, MT04 and MT05 to cause the most fluorescence reduction, but Fig. 7.18 (b) shows MT06 and MT05 to accelerate the aggregation the most, but reaching a similar endpoint as untreated TDP-43 CTF.

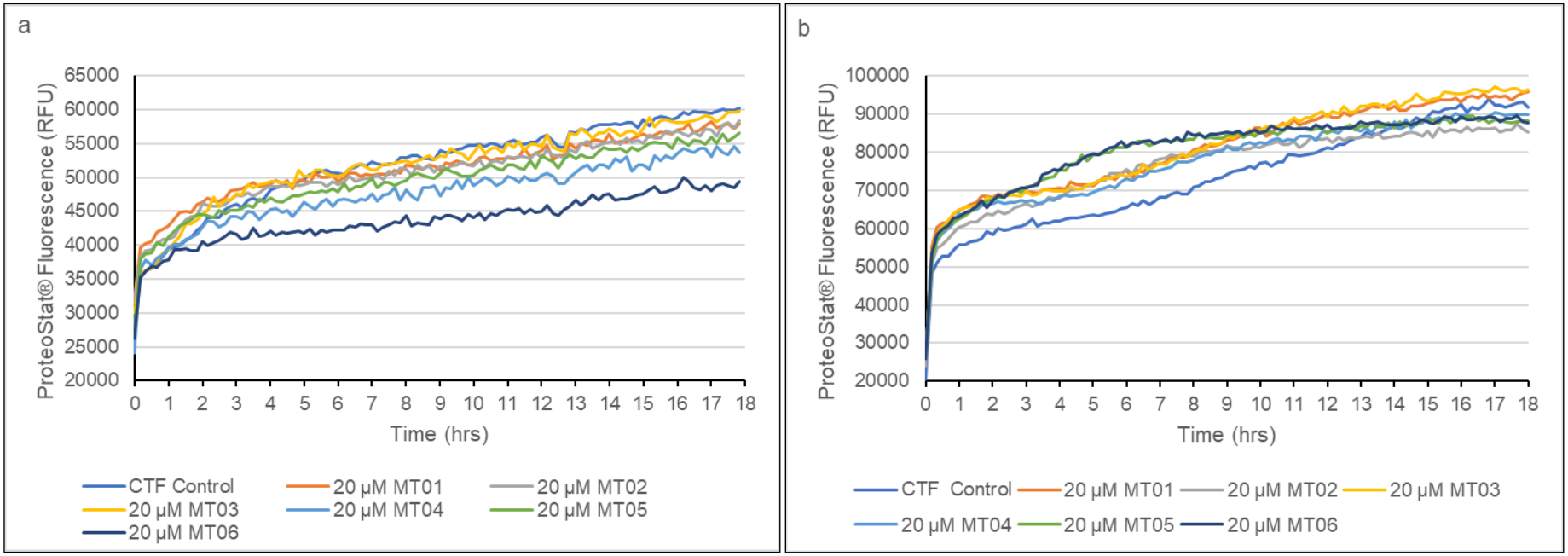


Figure 7.18. The effect of first-generation peptide inhibitors on the TDP-43 CTF aggregation assay. ProteoStat® aggregation assay, showing ProteoStat® fluorescence with time of aggregating 20 μM TDP-43 CTF with and without the addition of 20 μM peptide inhibitors. (**a**) and (**b**) are repeats of the same reaction conditions Lines represent mean fluorescence of triplicate reactions.

To potentially amplify any effect of the inhibitors on TDP-43 CTF aggregation, higher concentrations of inhibitor were tested. Often in aggregation assays, inhibitors are more efficient at higher concentrations than their target proteins, as demonstrated by AIM4 against TDP-43 CTF (Prasad *et al*., 2016) and RI-AG03 against TauΔ1-250 (Aggidis, 2019), though this may limit their therapeutic use. Therefore, inhibitors were tested at 200 μM, ten times the concentration of the target TDP-43 CTF. Equimolar reactions were run in parallel for comparison. As shown in Fig. 7.19 (b), at 200 μM, all inhibitors accelerated aggregation to a higher endpoint than untreated TDP-43 CTF control reactions, and higher than the inhibitors at 20 μM (Fig. 7.19 (a)). Fig. 7.19 (a) also demonstrates further variability in the results of inhibitor testing at equimolar concentration, with MT03 this time appearing the most effective, but with only a modest reduction of fluorescence of 4000 RFU.

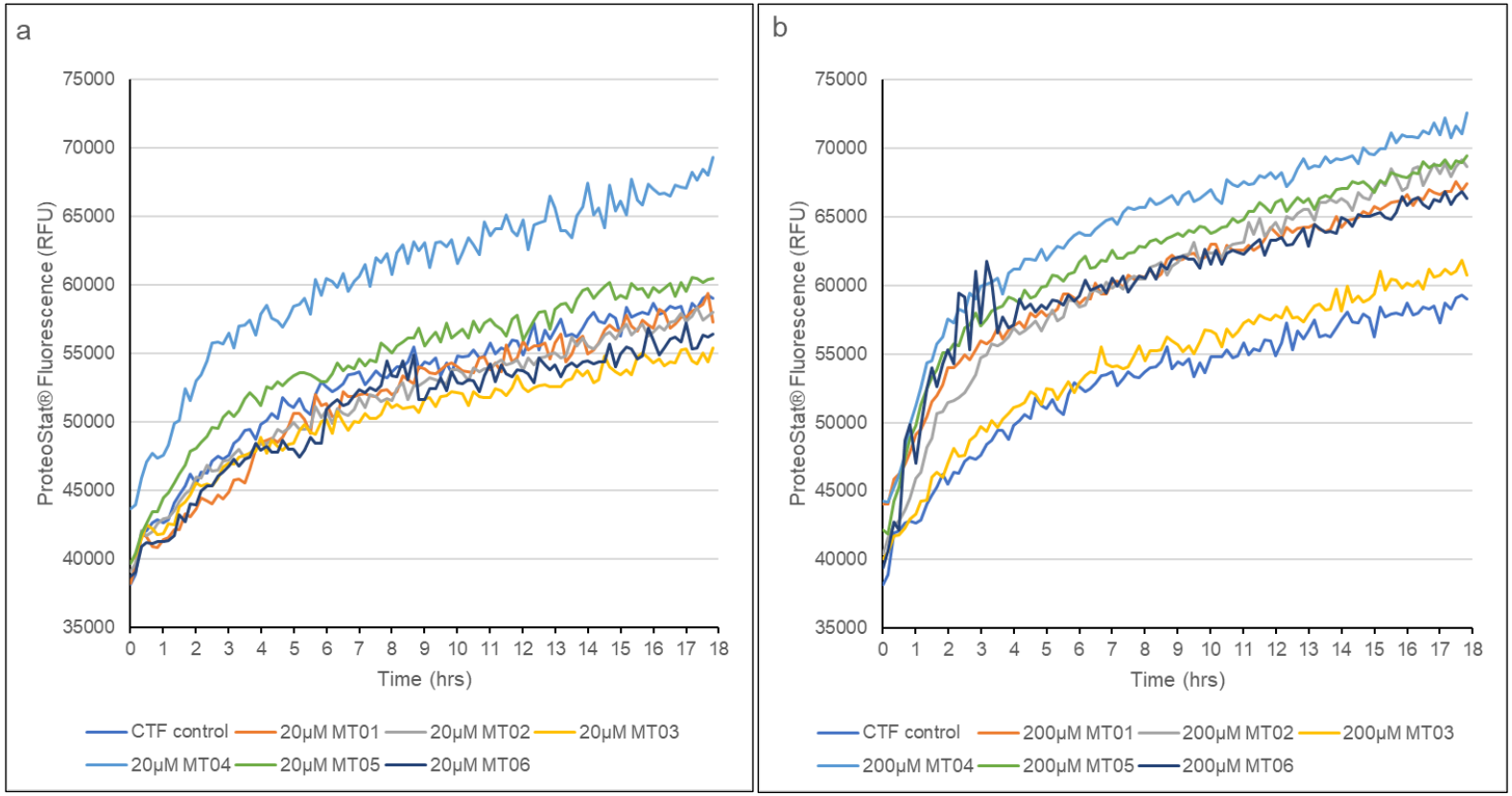


Figure 7.19. The effect of increased concentrations of first-generation peptide inhibitors on the TDP-43 CTF aggregation assay. ProteoStat® aggregation assay, showing ProteoStat® fluorescence with time of aggregating 20 μM TDP-43 CTF with and without the addition of (**a**) 20 μM or (**b**) 200 μM peptide inhibitors. Lines represent mean fluorescence of triplicate reactions.

End points of reactions were taken of the MT06 reaction from Fig. 7.18 (a) and the MT03 reaction from Fig. 7.19 (a) and prepared for TEM imaging. Representative images from each reaction are shown in Fig. 7.20. The previously observed morphology of the aggregated TDP-43 CTFs remained present with equimolar treatment by the first-generation inhibitors, indicating that these inhibitors have minimal effect on TDP-43 CTF aggregation.

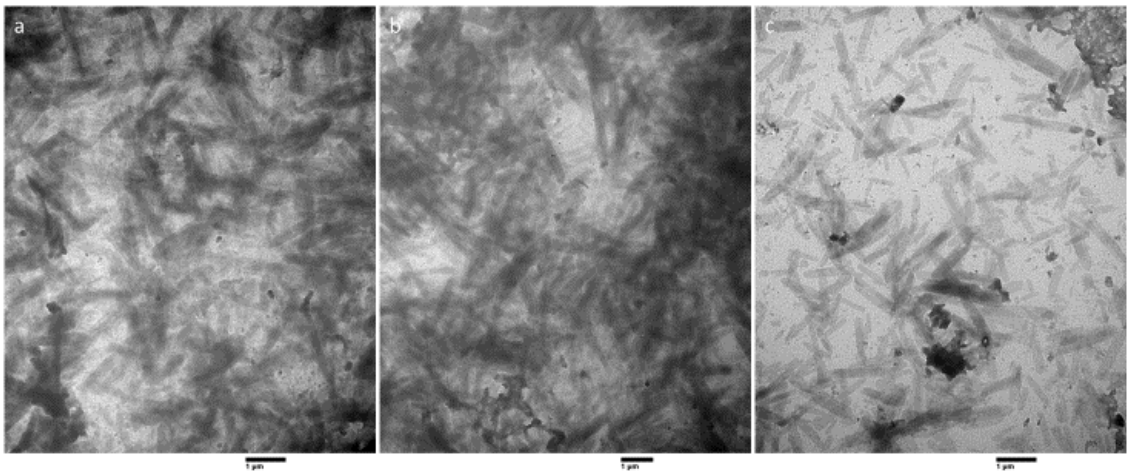


Figure 7.20. Representative TEM images of endpoint aggregation reactions for TDP-43 CTF treated with first-generation inhibitors. a- endpoint sample of an MT06 reaction from Fig. 7.18 (a), 12 000x magnification, scale bar represents 1 µm. b- endpoint sample of an MT03 reaction from Fig. 7.19 (a), 12 000x magnification, scale bar represents 1 µm. c- endpoint sample of an MT06 reaction from Fig. 7.18 (a), 15 000x magnification, scale bar represents 1 µm.

7.3.4 Design of poly-arginine tagged peptide inhibitors

In an attempt to improve the first-generation of peptide inhibitors, they were redesigned with the addition of a poly-arginine stretch onto the C-terminus. This decision was based on several factors. Firstly, in previous peptide inhibitor development in the Allsop group, particularly against Tau protein, the addition of a charged stretch of arginine (4-8 residues) to early generations of peptide inhibitors led to an increase in their effectiveness, likely due to the introduction of stretch of charge and hydrophilicity into an aggregation prone area (Aggidis, 2019). Secondly, in the case of TDP-43, for the aggregation regions targeted, the core of aggregation may stretch further towards the C-terminus than just the hexapeptide sequences, clearly demonstrated by the “double spiral fold” structure demonstrated by Arseni *et al*., (2019), where there is a continued stretch of hydrophobicity in the fold following the GAFSINPAMM sequence. Therefore, extending a stretch of hydrophilic, charged arginine residues into this area could prove beneficial. Finally, an octa-arginine sequence can act as a cell penetrating peptide (Schmidt *et al*., 2009; Aggidis, 2019), meaning that the peptide could potentially cross the cell membrane to act intracellularly where it would be needed *in vivo*. The second-generation inhibitors are outlined in Table 7.2.

Table 7.2. The second-generation poly-arginine peptide aggregation inhibitors and their TDP-43 aggregation regions targeted.

Peptide Inhibitor	Peptide Inhibitor sequence	TDP-43 aggregation sequence targeted	TDP-43 residue number targeted
MT01R	H-Arg-Gly-Leu-Ile-Ile-Lys-Gly-Ile-Gly-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-OH	Leu-Ile-Ile-Lys-Gly-Ile	248-253
MT02R	H-Arg-Gly-Ile-Ile-Lys-Gly-Ile-Ser-Gly-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-OH	Ile-Ile-Lys-Gly-Ile-Ser	249-254
MT03R	H-Arg-Gly-Ile-Lys-Gly-Ile-Ser-Val-Gly-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-OH	Ile-Lys-Gly-Ile-Ser-Val	250-255
MT04R	H-Arg-Gly-Phe-Gly-Ala-Gly-Ser-Ile-Gly-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-OH	Phe-Gly-Ala-Phe-Ser-Ile	313-318
MT05R	H-Arg-Gly-Gly-Ala-Phe-Ser-Ile-Asn-Gly-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-OH	Gly-Ala-Phe-Ser-Ile-Asn	314-318
MT06R	H-Arg-Gly-Ala-Phe-Ser-Ile-Asn-Pro-Gly-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-OH	Ala-Phe-Ser-Ile-Asn-Pro	315-320

7.3.5 Investigation of the inhibition of aggregation of TDP-43 CTF by second-generation poly-arginine tagged inhibitors

The poly-arginine tagged peptides were then tested in the ProteoStat® TDP-43 CTF aggregation assay, again initially at equimolar concentrations to the CTF at 20 μM. The results are shown in Fig. 7.21. Surprisingly, the addition of equimolar poly-arginine inhibitors rapidly accelerated the aggregation of the TDP-43 CTF to a higher fluorescence endpoint. This occurred repeatedly across multiple reactions of triplicate repeats, with some variation as to which peptide accelerated aggregation the most.

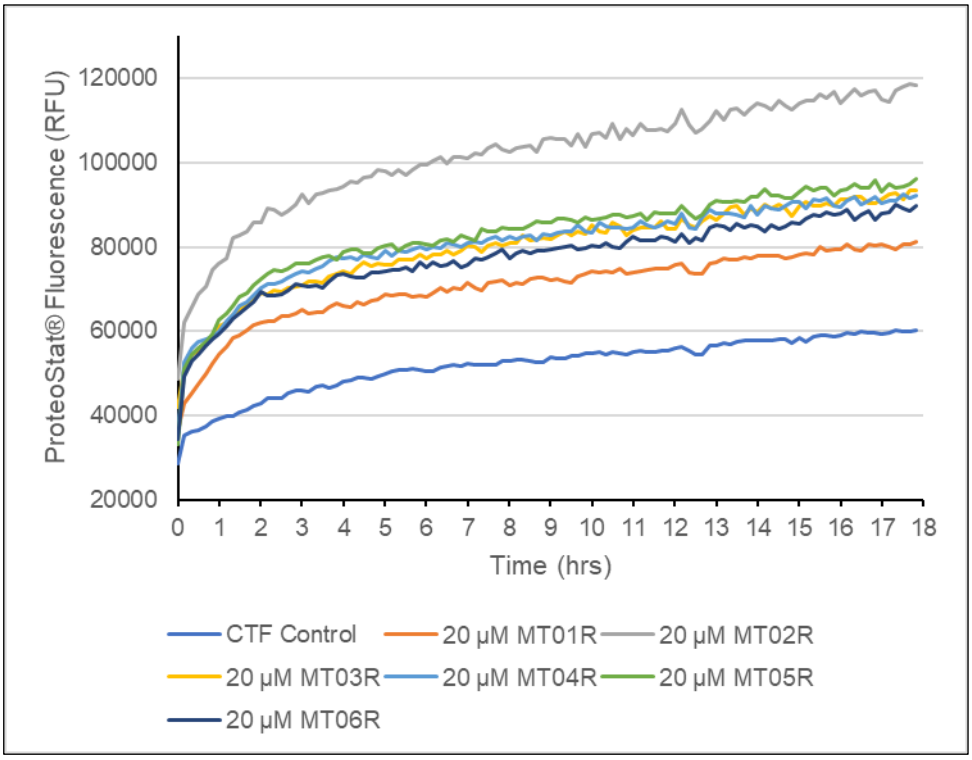


Figure 7.21. The effect of second-generation peptide inhibitors on the TDP-43 CTF aggregation assay. ProteoStat® aggregation assay, showing ProteoStat® fluorescence with time of aggregating 20 μM TDP-43 CTF with and without the addition of 20 μM polyarginine-tagged peptide inhibitors. Lines represent mean fluorescence of triplicate reactions.

TEM imaging of end point reactions revealed a change in morphology of the aggregates produced by aggregation with the poly-arginine tagged inhibitor peptides. Multiple aggregate morphologies were identified in these reactions, including irregular, granular morphologies producing large masses (Fig. 7.22) but also some species appearing filamentous and fibrillar (Fig. 7.23). These poly-arginine tagged inhibitors therefore appear to alter the aggregation process of TDP-43 CTF, triggering a more rapidly aggregating species with an altered morphology when compared to TDP-43 CTF aggregated in the absence of any inhibitor.

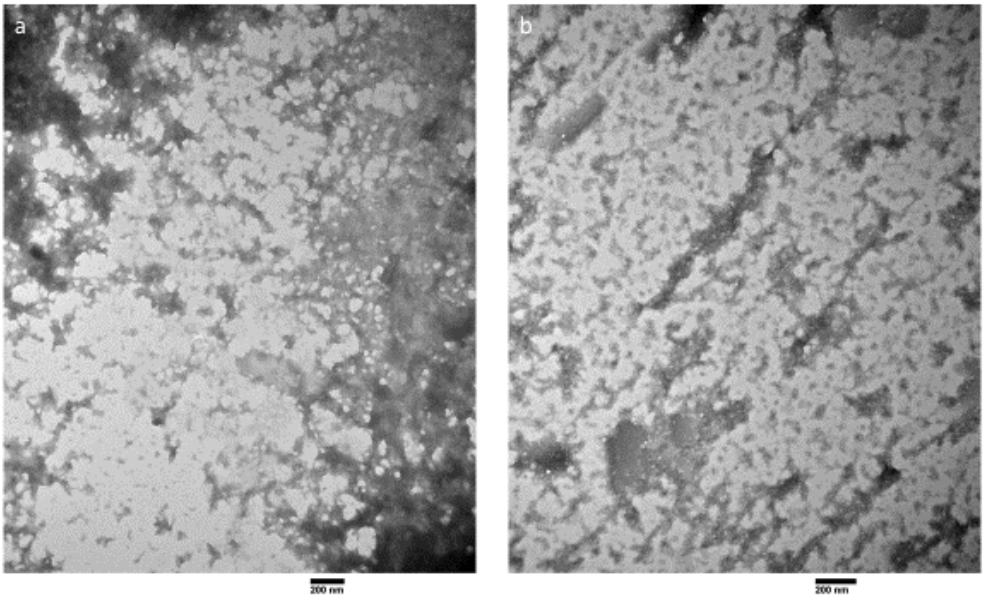


Figure 7.22. Representative TEM images of endpoint aggregation reactions for TDP-43 CTF treated with second-generation inhibitors. a- endpoint sample of an MT06R reaction, 50 000x magnification, scale bar represents 200 nm. b- endpoint sample of an MT02R reaction, 60 000x magnification, scale bar represents 200 nm.

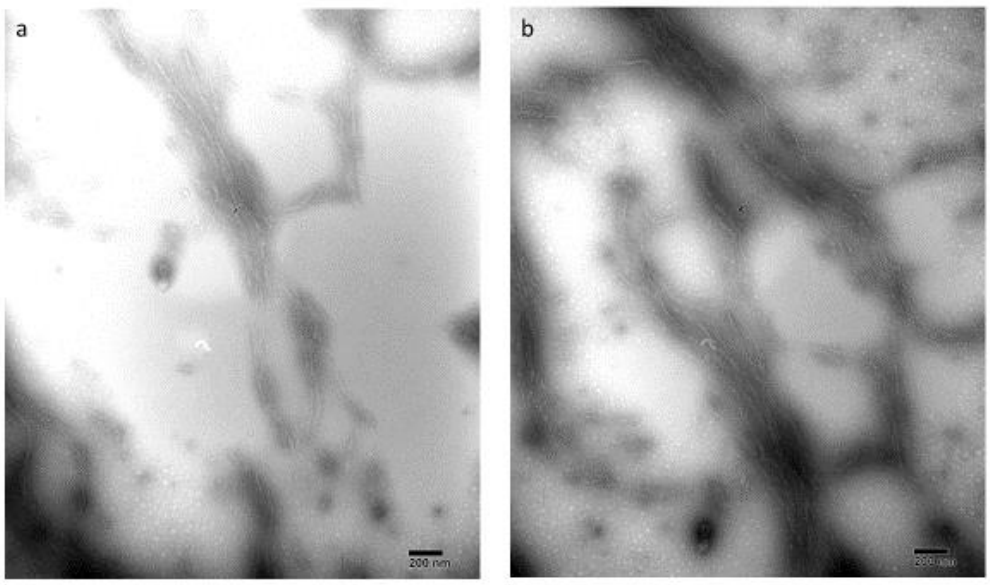


Figure 7.23. Representative TEM images of endpoint aggregation reactions for TDP-43 CTF treated with second-generation inhibitors. a- endpoint sample of an MT06R reaction, 50 000x magnification, scale bar represents 200 nm. b- endpoint sample of an MT02R reaction, 50 000x magnification, scale bar represents 200 nm.

7.3.6 AlphaFold based design of third-generation peptides

As previously discussed, the first two generations of peptide inhibitors were designed based on aggregation prone sequences identified in the literature. However, these sequences were determined mainly by aggregation of short peptide sequences, which it could be argued may not represent how a larger protein fragment may aggregate due to their lack of structure. Furthermore, other aggregation regions have been identified in a similar manner (Wang *et al*., 2013; Shimonaka *et al*., 2016; Kumar *et al*., 2019), five regions of which do not correspond with the sequence involved in the "double spiral fold" identified from aggregates in ALS with FTLD patients (Arseni *et al*., 2021). Furthermore, the TDP-43 CTF being aggregated in this project is of a different length to fragments aggregated previously (Saini and Chauhan, 2011; Prasad *et al*., 2016) and certainly of a larger size than the short peptides aggregated to identify aggregation prone sequence.

Consideration was therefore made of how the TDP-43 CTF molecules being aggregated may actually interact to aggregate. Given the lack of full structures of TDP-43 and the TDP-43 CTF region, the AI programme AlphaFold v2 was used to predict a structure of the TDP-43 CTF used in this project from its amino acid sequence, shown in Fig. 7.24. By comparison, this structure is similar to the corresponding region in the full structure of TDP-43_{WtoA} (Wright *et al*., 2020), that is, containing the characteristic β-sheet and helix fold of RRM2, with a highly unstructured CTD containing a single α-helix.

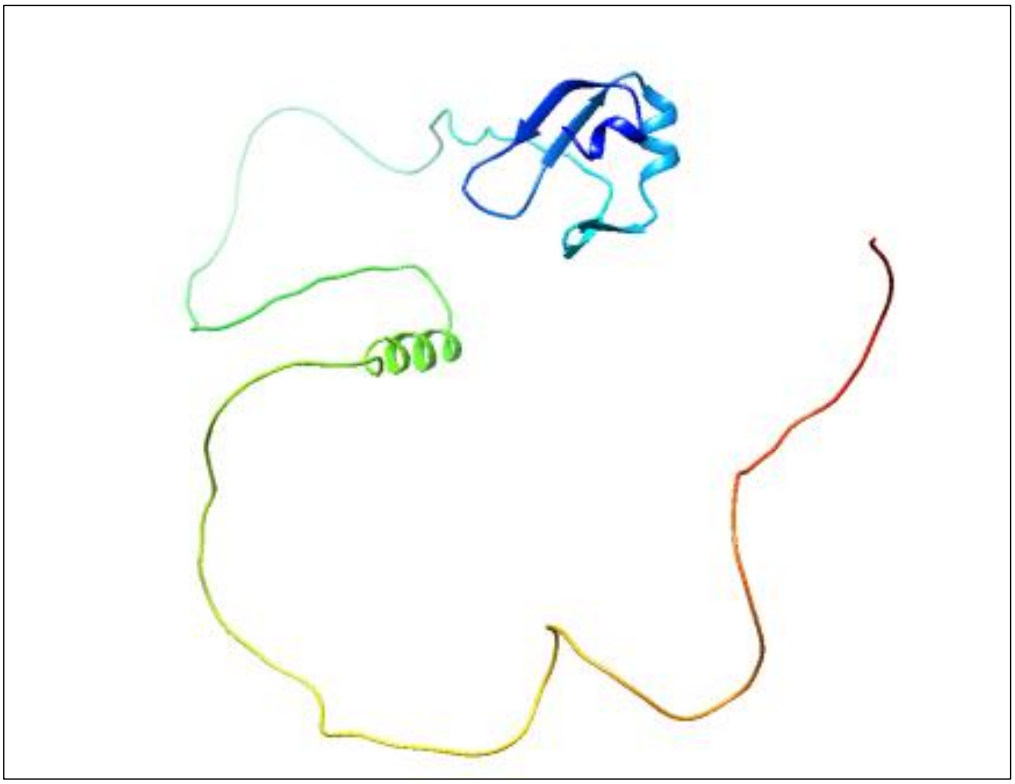


Figure 7.24. AlphaFold v2 generated structure for TDP-43 CTF (208-414). Viewed in UCSF Chimera. Blue to red indicates N-terminus to C-terminus.

The point of interaction between two TDP-43 CTF molecules was then predicted. As shown in Fig. 7.25, AlphaFold predicts a point of interaction in the RRM2 section of the fragment, where β strands in different molecules can coordinate into a potential sheet like structure. Interestingly, the predicted β-strand mainly involved corresponds to a sequence GISVHIS, part of the 246-258 aggregation sequence identified by Saini and Chauhan (2011), and GISV was targeted by first generation inhibitors. This 246-258 aggregation sequence, EDLIIKGISVHIS, itself appears as a two-strand antiparallel sheet, turning at the central KG amino acids. By viewing the TDP-43 CTF by hydrophobicity surface in UCSF Chimera, these sheets appear as exposed hydrophobic (red) surfaces (Fig. 7.26).

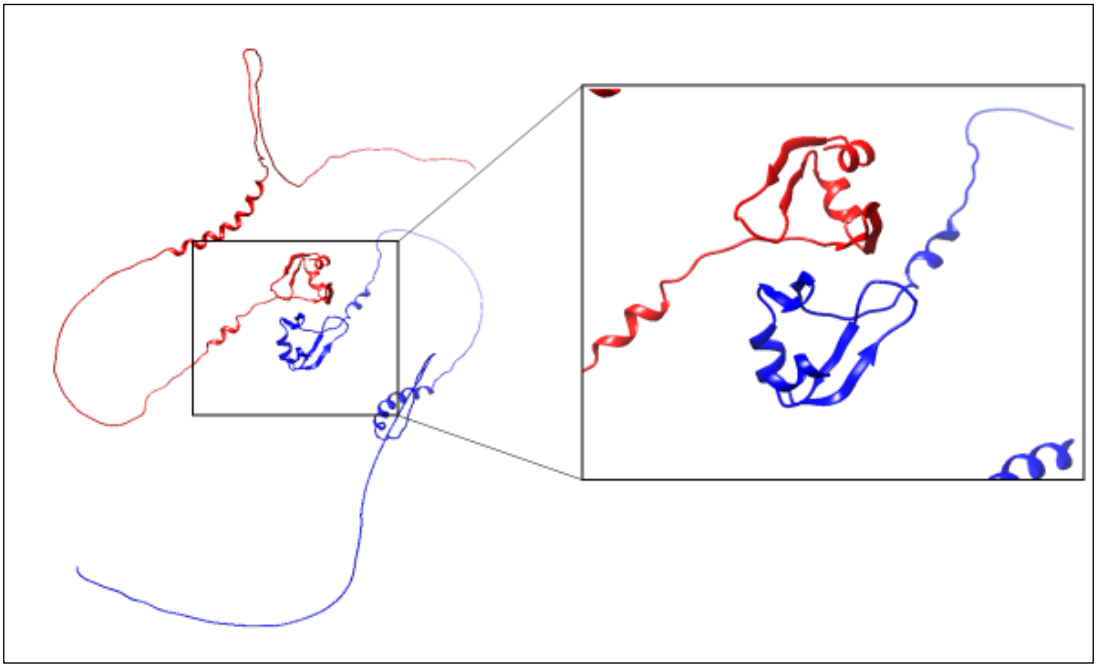


Figure 7.25. AlphaFold v2 generated structure for the interaction of two TDP-43 CTF (208-414) molecules. Viewed in UCSF Chimera. Blue and red ribbons indicate individual molecules.

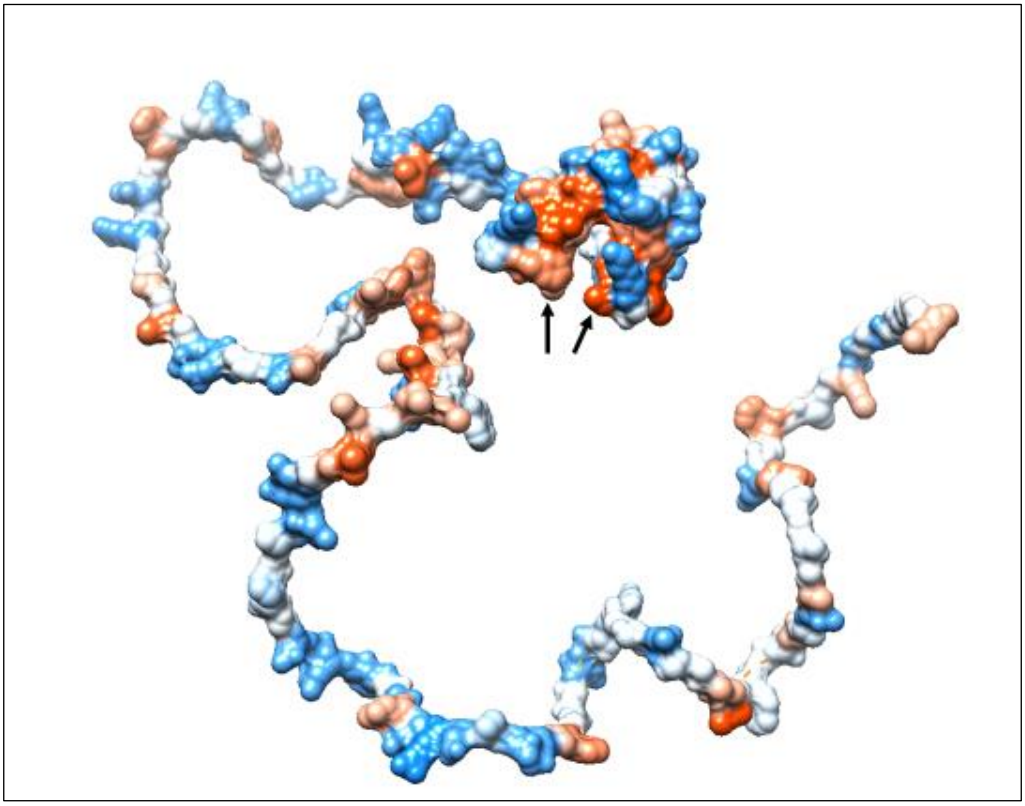


Figure 7.26. AlphaFold v2 generated structure for TDP-43 CTF (208-414), viewed in UCSF Chimera displaying hydrophobicity surface. Arrows indicate the exposed hydrophobic (red) residues involved at the predicted point of interaction between two TDP-43 CTF molecules.

When first- and second-generation inhibitor peptide sequences were added to the AlphaFold multimer prediction, a consistent point of interaction was identified. For both generations of inhibitors, the peptide was inserted into the space between the two stranded β-sheets (Fig. 7.27), with the space acting like a “pocket”. This was predicted for both the peptides targeting amino acids 246-258 (MT01,02 and 03) and for those targeting 311-323 (MT04,05 and 06). The point of interaction remained unaltered for first generation inhibitors, with the peptide sitting in the sheet structure (Fig. 7.27 (a) and (b)). For the poly-arginine tagged inhibitors, the peptide targeting sequence again sits in the structure, with the octa-arginine tag extending out away from the protein molecules, and either causing no change to the predicted interaction (Fig. 7.27 (c) and (d)), or causing a slightly altered, but potentially closer interaction, (Fig. 7.27 (e)).

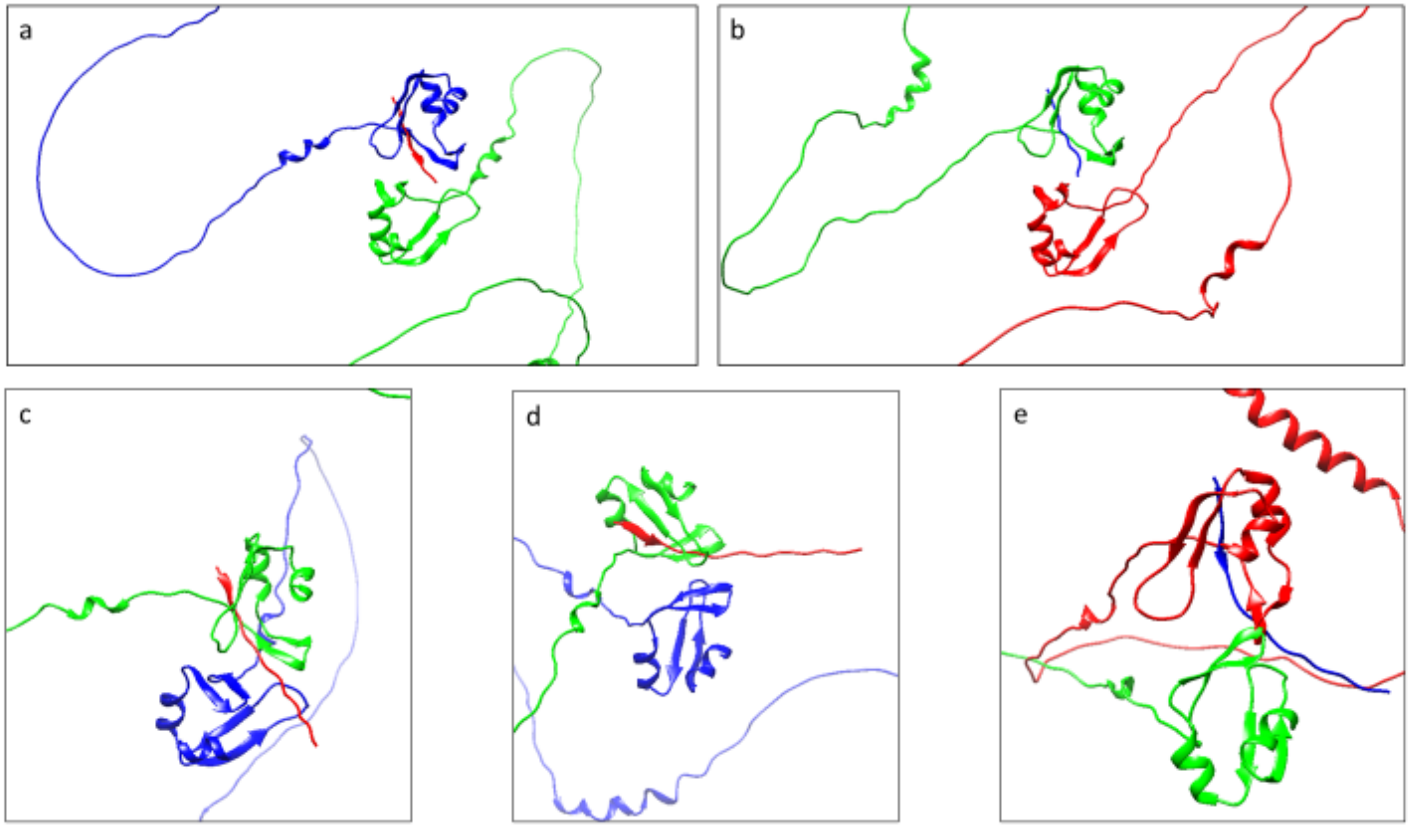


Figure 7.27. AlphaFold v2 generated structures for interactions between TDP-43 CTF (208-414) molecules when peptide inhibitors are included. Viewed in UCSF Chimera. a- 2 TDP-43 CTF molecules (blue and green) with MT01 (red). b- 2 TDP-43 CTF molecules (red and green) with MT06 (blue). c and d – 2 TDP-43 CTF molecules (blue and green) with MT01R (red) (the same prediction at two different viewing angles). e- 2 TDP-43 CTF molecules (red and green) with MT06R (blue).

Following these observations from AlphaFold, it was hypothesised that a peptide could be designed to target the “pocket” space between β strands, that would bend back on itself to fill the space between the two TDP-43 CTF molecules with a hydrophilic, charged strand. In this way, rather than the hydrophobic targeting core of the peptide being included into a sheet structure between the molecules, the hydrophilic stretch would interrupt the sheet and “mask” the exposed area of hydrophobicity.

First, a binding sequence was designed. As noted, the apparent point of interaction was the GISVHIS β strand and so a binding sequence was designed based on this. Due to the orientation of the strand, and how the peptide would need to bend back into the space between molecules, the sequence was reversed to IHVSI- it was identified by AlphaFold predictions that this length of binding sequence aligns well into the space between the β strands (Fig. 7.28 (d)). Consideration was also made as to the strength of binding of the peptide vs another TDP-43 CTF molecule. Conversion of the serine in the binding sequence to the more hydrophobic alanine, producing a sequence IHVAI, better disrupted the interaction between two GISVHIS peptides than the IHVSI sequence (Fig. 7.28 (a) (b) and (c)). Next, consideration of how a “harpin” structure could be produced was made. After the binding sequence, a glycine-proline-glycine turn could be included, followed by a chain of hydrophilic residues. At the TDP-43 CTF aggregation reaction conditions of pH8, the target sequence, and the TDP-43 CTF as a whole will exhibit a net negative charge (-0.5 for GISVHIS, -0.6 for the TDP-43 CTF). A hydrophilic stretch consisting of a polycation at this pH such as polyarginine, would be electrostatically attractive to the aggregation sequence of another TDP-43 CTF. Therefore, a poly-glutamic acid stretch was added to the sequence, to provide hydrophilicity and a net negative charge.

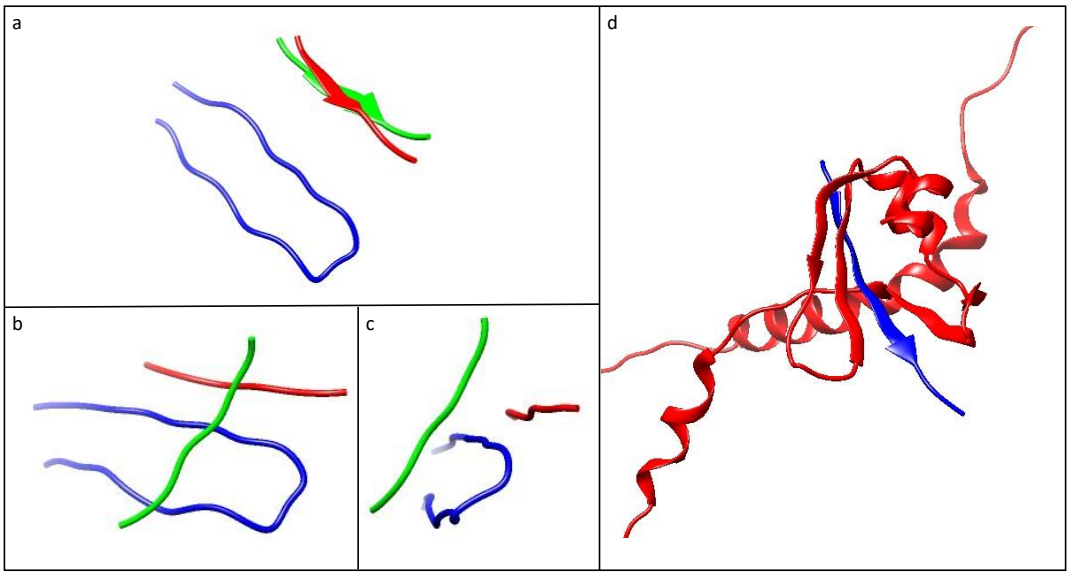


Figure 7.28. AlphaFold v2 generated structures for interactions of early designs of third generation peptides with target sequences, viewed in UCSF Chimera. a- prediction of the interaction of inhibitor peptide (DCIHVSIGPGEEEEECE) (blue) containing IHVSI binding sequence with 2 GISVHIS peptides (red and green)- note the strong interaction of the GISVHIS sequences with exclusion of the inhibitor. b and c- prediction of the interaction of inhibitor peptide (blue) containing IHVAI binding sequence (DCIHVAIGPGEEEEECE) with 2 GISVHIS peptides (red and green)- note the preferential alignment of the red GISVHIS strand with the blue IHVAI sequence. d- Insertion of a DCIHVSIGD peptide (blue) into the TDP-43 CTF molecule “pocket” (red).

Cysteine residues were added as the second and penultimate amino acids in the sequence, allowing the peptide to be linked into a “hairpin” shape bending at the central Gly-Pro-Gly turn. It is possible to “staple” cysteine residues together covalently into a stable thioether bond using spacer molecules such as di-bromo-xylenes, with the orientation of the xylene providing different linker lengths (Fig. 7.29 (b), (c) and (d)). A post linkage purification step removes any intermolecularly bonded peptides that could form in this reaction.

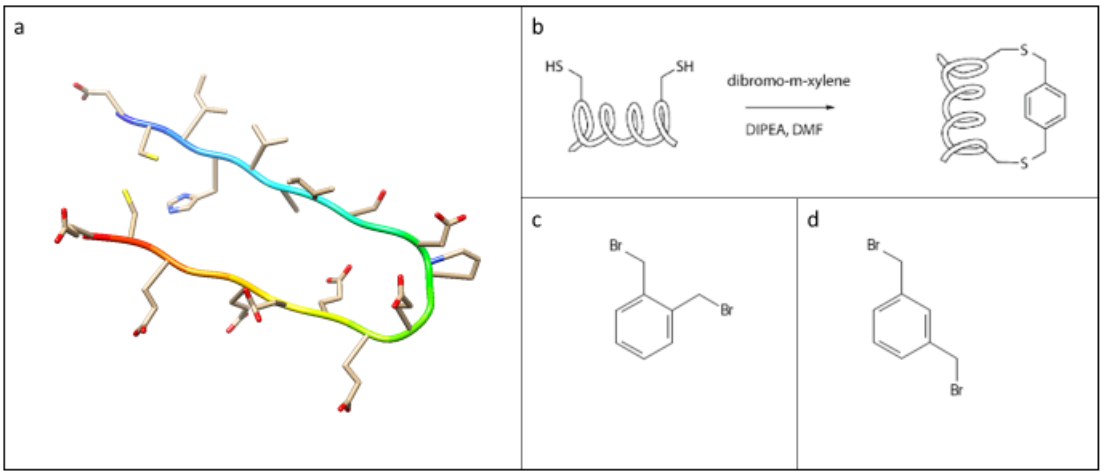


Figure 7.29. Third generation peptide design and linker chemistry. a- final design of the third-generation inhibitor peptide sequence DCIHVAISGDPGEEEEEGECE, generated in AlphaFold v2 and imaged in UCSF Chimera, showing side chain atoms. b- Diagrammatic representation of a cysteine stapling reaction using di-bromo-xylene, from (Moiola *et al.*, 2019). c- structure of di-bromo-o-xylene, generated in PubChem Sketcher V2.4. d- structure of di-bromo-p-xylene generated in PubChem Sketcher V2.4.

Together, these considerations produced a peptide with sequence DCIHVAISGDPGEEEEEGECE, shown in Fig. 7.29 (a). When this peptide was included into an AlphaFold multimer prediction, the peptide inserted into its binding pocket, with the poly glutamic acid stretch bending down and back towards the N-terminus, filling the space under the sheet with a hydrophilic stretch (Fig. 7.30 (a)). The predicted interaction between the two TDP-43 CTF molecules was altered, with a new interaction being apparently mediated around a single exposed isoleucine side chain on each TDP-43 CTF molecule (Fig. 7.30 (b)).

Final considerations of the design were self-aggregation and spacer length. AGGRESCAN was used to predict whether the binding sequence was an aggregation hotspot. The AGGRESCAN result for DCIHVAISGDPGEEEEEGECE is shown in Fig.

7.31 (a). As shown, the target sequence is identified as an aggregation hotspot (5 residues above the “hotspot threshold”, and so this peptide may have a propensity to self-aggregate. Therefore, the sequence was altered to DC**A**HVAISGDPGEEEEEGECE, which reduced the predicted aggregation propensity to no longer be a predicted “hotspot” (4 residues above the threshold (Fig. 7.31 (b)). In case the original sequence did in fact not self-aggregate, or if the altered sequence was less effective due to the change, both sequences were ordered for synthesis.

To test spacer arm size between the cysteines, peptides linked with two different di-bromo xylene orientations were synthesised; a smaller di-bromo-o-xylene (Fig. 7.29 (c)) and larger di-bromo-p-xylene (Fig. 7.29 (d)). The third-generation peptide sequences, their denominations and linkages are shown in Table 7.3.

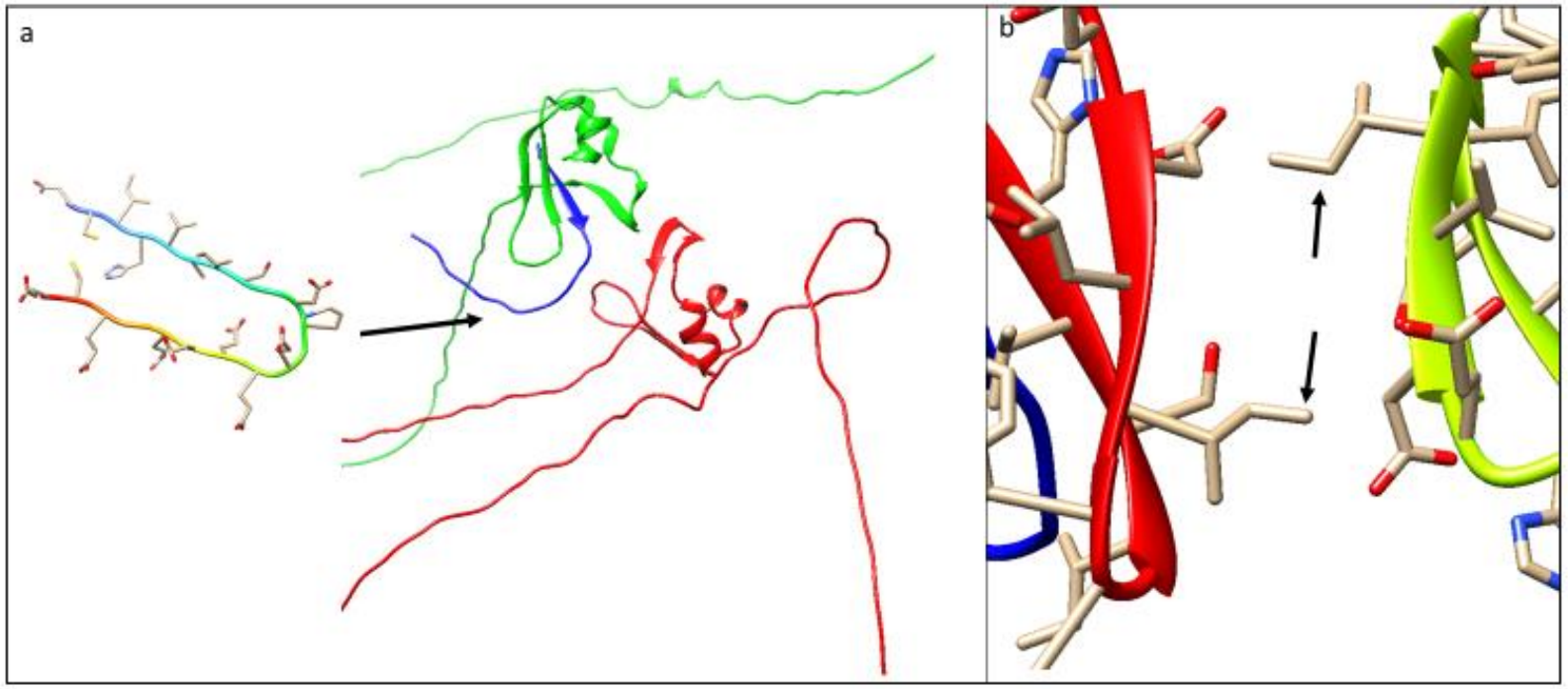


Figure 7.30. Predicted interaction of a third-generation peptide inhibitor with TDP-43 CTF molecules. a- Diagrammatic representation showing the predicted insertion of a third-generation peptide sequence DCIHVAISGDPGEEEEEGECE (blue ribbon) into its binding site, with an alteration of the interaction between two TDP-43 CTF molecules (red and green). Structures generated using AlphaFold v2 and viewed in UCSF Chimera. b- New predicted point of interaction mediated by isoleucine side chains (indicated with arrows) following inclusion of third generation peptide sequence in AlphaFold multimer with 2 TDP-43 CTF molecules (red and green) Structures generated using AlphaFold v2 and viewed in UCSF Chimera.

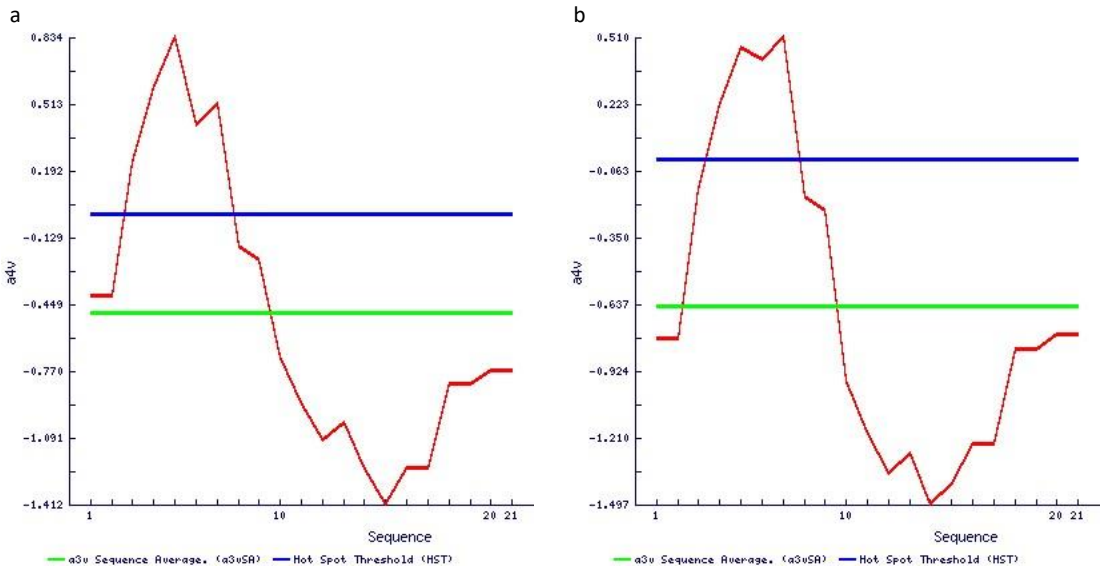


Figure 7.31. Peptide inhibitor AGGRESCAN results. a- DC**I**HVAISGDPGEEEEEGECE and b- DC**A**HVAISGDPGEEEEEGECE.

Table 7.3. The third-generation inhibitor sequences, their denominations and linkers.

Peptide Inhibitor	Peptide Inhibitor sequence	Linker for cysteine stapling
TMH1o	DCIHVAISGDPGEEEEEGECE	Di-bromo-o-xylene
TMH1p	DCIHVAISGDPGEEEEEGECE	Di-bromo-p-xylene
TMH2o	DCAHVAISGDPGEEEEEGECE	Di-bromo-o-xylene
TMH2p	DCAHVAISGDPGEEEEEGECE	Di-bromo-p-xylene

7.3.7 Investigation of the inhibition of aggregation of TDP-43 CTF by third-generation peptide inhibitors

The synthesised third-generation peptides were resuspended in DMSO as recommended by the manufacturer (Peptide Synthetics, Peptide Protein Research Ltd). First, due to the concerns over self-aggregation arising from AGGRESCAN data, the peptides were tested for self-aggregation using the ProteoStat® assay, the results of which are shown in Fig. 7.32, with aggregating TDP-43 CTF used as a control. As shown, the inhibitors do not appear to show self-aggregation and produce minimal ProteoStat® fluorescence, not increased above the fluorescence of a buffer control containing no inhibitor or protein.

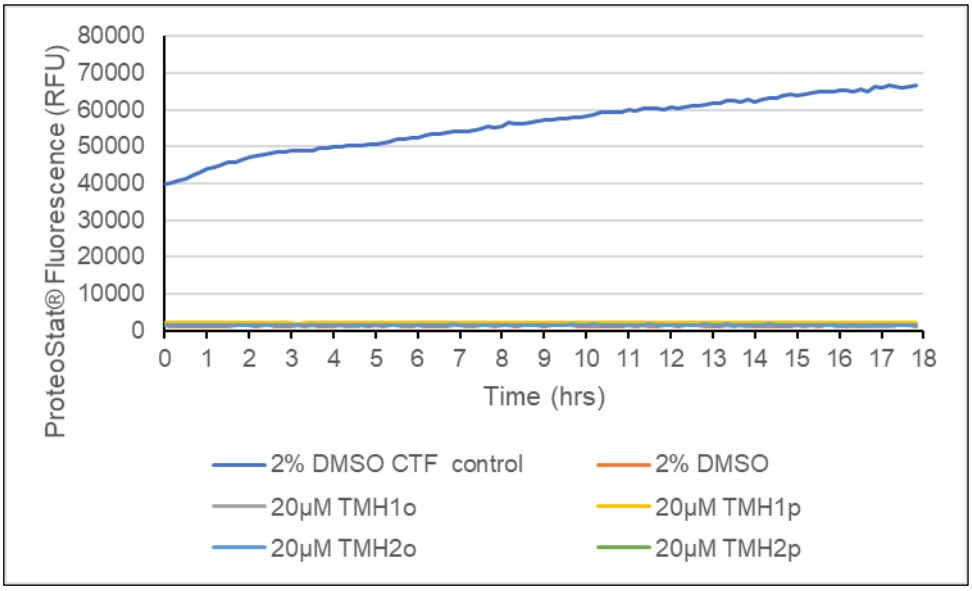


Figure 7.32. Self-aggregation testing of third-generation peptide inhibitors. ProteoStat® aggregation assay, showing ProteoStat® fluorescence with time for self-aggregation of third-generation peptide inhibitors. Aggregating TDP-43 CTF with 2% DMSO was used as a positive control, aggregation reaction buffer containing 2% DMSO was used as a negative. Inhibitors were diluted to 20 μM in aggregation buffer (2% DMSO final) in the absence of TDP-43 CTF.

The peptides were then tested as aggregation inhibitors, initially at equimolar concentration to the TDP-43 CTF. The results are shown in Fig. 7.33. All inhibitors led to a reduction in the ProteoStat® fluorescence across 18 hours, indicating a reduction in aggregation. Furthermore, they appeared to introduce a “lag” of approximately 3 hours before the fluorescence increase began following the initial burst in the dead time of the experiment. This pattern was seen repeatedly across multiple assays, but with some variation between which inhibitor led to the most aggregation reduction. As demonstrated by the error bars denoting standard deviation of the mean in Fig. 7.33, there was overlap between all four inhibitors, but not between any inhibitor and the untreated control after approximately 2 hours of aggregation.

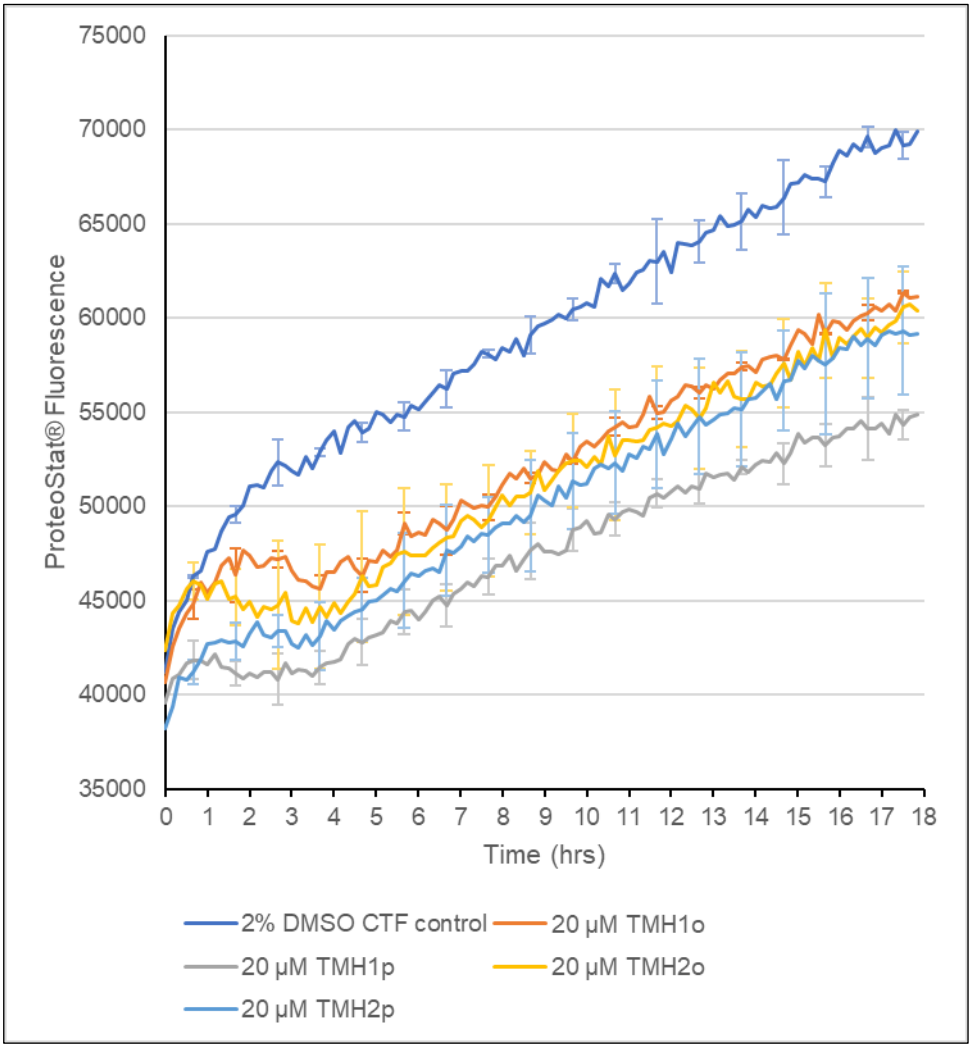


Figure 7.33. The effect of third-generation peptide inhibitors on the TDP-43 CTF aggregation assay. ProteoStat® aggregation assay, showing ProteoStat® fluorescence with time of aggregating 20 μM TDP-43 CTF with and without the addition of 20 μM third-generation peptide inhibitors. Lines represent mean fluorescence of triplicate reactions and error bars represent standard deviation of the means.

To identify if a higher inhibitor concentration could improve the reduction in aggregation, like AIM4 (Prasad *et al.*, 2016), inhibitors were tested against 20 μM TDP-43 CTF at 40 μM and 100 μM (2x and 5x molar excess respectively). For all inhibitors this marginally appeared to improve the reduction in aggregation compared to equimolar concentration. An example of this is shown for TMH2p in Fig. 7.34, with the error bar overlap demonstrating the minimal reduction with increased inhibitor concentration.

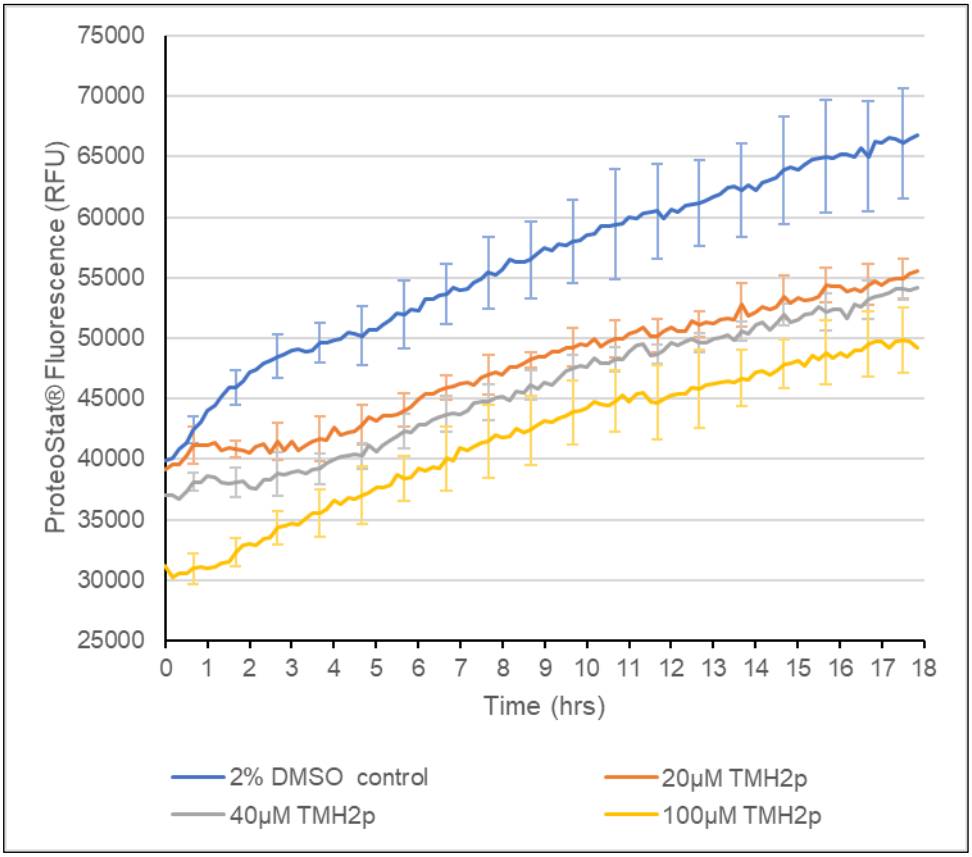


Figure 7.34. The effect of increased concentrations of a third-generation peptide inhibitor on the TDP-43 CTF aggregation assay. ProteoStat® aggregation assay, showing ProteoStat® fluorescence with time of aggregating 20 μM TDP-43 CTF with and without the addition of TMH2p inhibitor at 20 μM, 40 μM and 100 μM. Lines represent mean fluorescence of triplicate reactions and error bars represent standard deviation of the means.

Endpoints of aggregation reactions were then processed for TEM and imaged. Treatment with all four inhibitors appeared to produce aggregates with altered morphologies as demonstrated by representative images in Figs. 7.35, 7.36 and 7.37, including amorphous species, smaller granular species and aggregates appearing almost like the untreated CTF rod structures, but with more irregularity in their structure. The third-generation inhibitors therefore appear to reduce the aggregation of TDP-43 CTF and alter the morphologies of the aggregates that do form.

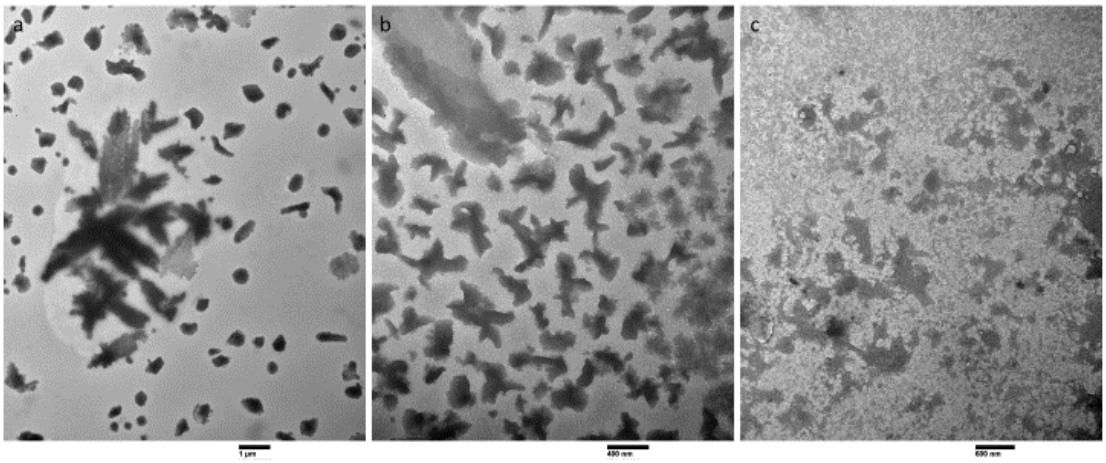


Figure 7.35. Representative TEM images of endpoint aggregation reactions for TDP-43 CTF treated with third-generation inhibitors, showing amorphous structures. a- endpoint sample of a TMH2p reaction, 12 000x magnification, scale bar represents 1 µm. b- endpoint sample of a TMH2p reaction, 40 000x magnification, scale bar represents 400 nm. c- endpoint sample of a TMH2o reaction, 25 000x magnification, scale bar represents 600 nm.

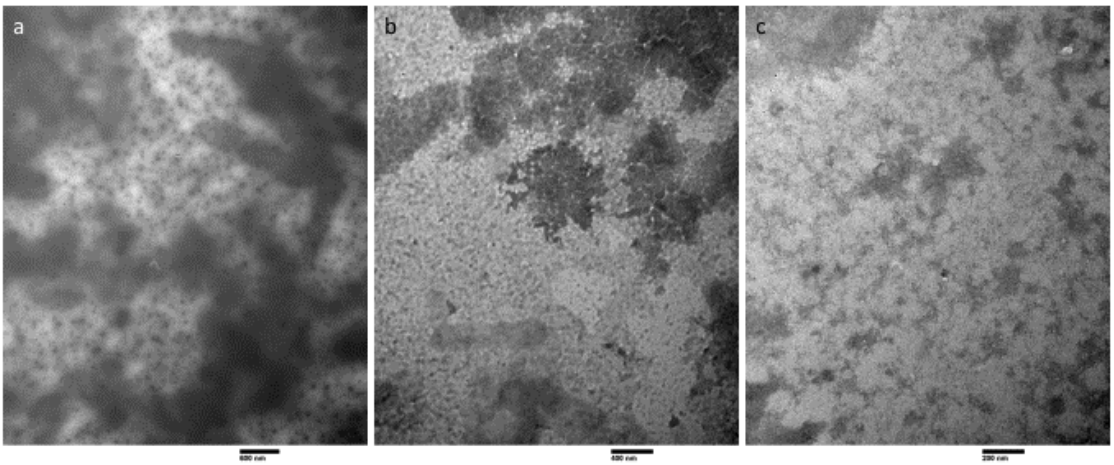


Figure 7.36. Representative TEM images of endpoint aggregation reactions for TDP-43 CTF treated with third-generation inhibitors showing granular morphology. a- endpoint sample of a TMH1o reaction, 25 000x magnification, scale bar represents 600 nm. b- endpoint sample of a TMH1p reaction, 40 000x magnification, scale bar represents 400 nm. c- endpoint sample of a TMH2o reaction, 80 000x magnification, scale bar represents 200 nm.

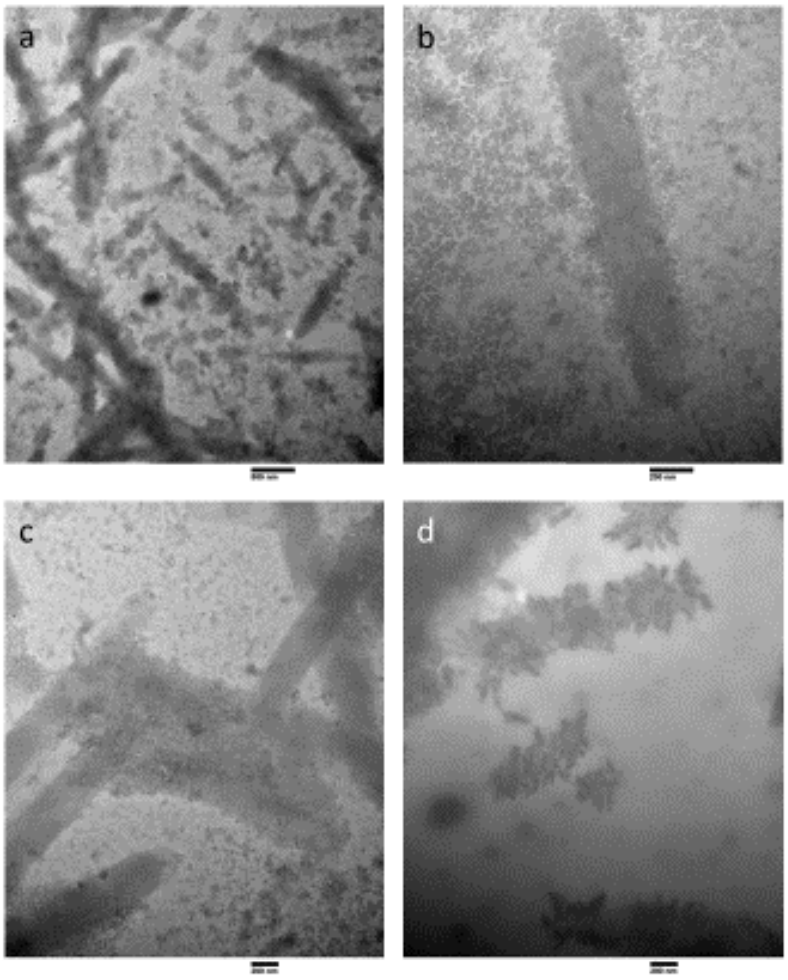


Figure 7.37. Representative TEM images of endpoint aggregation reactions for TDP-43 CTF treated with third-generation inhibitors showing altered rod-like morphology. a-endpoint sample of a TMH1p reaction, 20 000x magnification, scale bar represents 800 nm. b-endpoint sample of a TMH1p reaction, 80 000x magnification, scale bar represents 200 nm. c-endpoint sample of a TMH1p reaction, 50 000x magnification, scale bar represents 200 nm. d-endpoint sample of a TMH1o reaction, 80 000x magnification, scale bar represents 200 nm.

7.3.8 Investigation of the inhibition of aggregation of TDP-43 CTF by peptide inhibitor combinations

Given the altered morphologies of the aggregates observed by treatment with third-generation inhibitors, it was hypothesised that the aggregation could be occurring from a different aggregation sequence to that being targeted, if the inhibitors were indeed binding where they were designed to. The GAFSINPAMM sequence, that targeted by first generation inhibitors MT04, 05 and 06, is present in the TDP-43 CTF and as mentioned has been identified repeatedly to likely be involved in TDP-43 aggregation. Therefore, it was hypothesised that if the third-generation inhibitors were binding where designed and blocking aggregation from the AlphaFold predicted point of interaction, then another inhibitor could be used in tandem to target the GAFSINPAMM sequence, if the protein was now aggregating around this point. Due to time constraints of the project, another generation of inhibitors could not be designed and produced for this sequence, however the first generation MT04, 05 and 06 inhibitors were tested in tandem with TMH1o, 1p, 2o and 2p. Although apparently ineffective on their own, AlphaFold predicts them to bind in the space that would in theory be occupied by the third-generation inhibitor, and so they may now bind instead to their desired target sequence.

MT04, MT05 and MT06 were tested at equimolar concentrations to the TDP-43 CTF and TMH1o, 1p, 2o and 2p. In some cases, as for TMH1o shown in Fig. 7.38, on average, inclusion of the first-generation inhibitor slightly reduced the aggregation further than TMH1o alone, but as demonstrated by the error bar overlap, this could be due to just the variation in the experiment. In fact, for the other inhibitors, such as TMH1p shown in Fig. 7.39, inclusion of the first-generation inhibitors had the opposite effect. Across multiple assays, as found with the first-generation inhibitors alone, no consensus was determined on their effect, and no further alteration to aggregate morphology was observed by TEM. Given that the first-generation inhibitors were

generally ineffective by themselves, and in fact accelerated aggregation at higher concentrations, they were not used further.

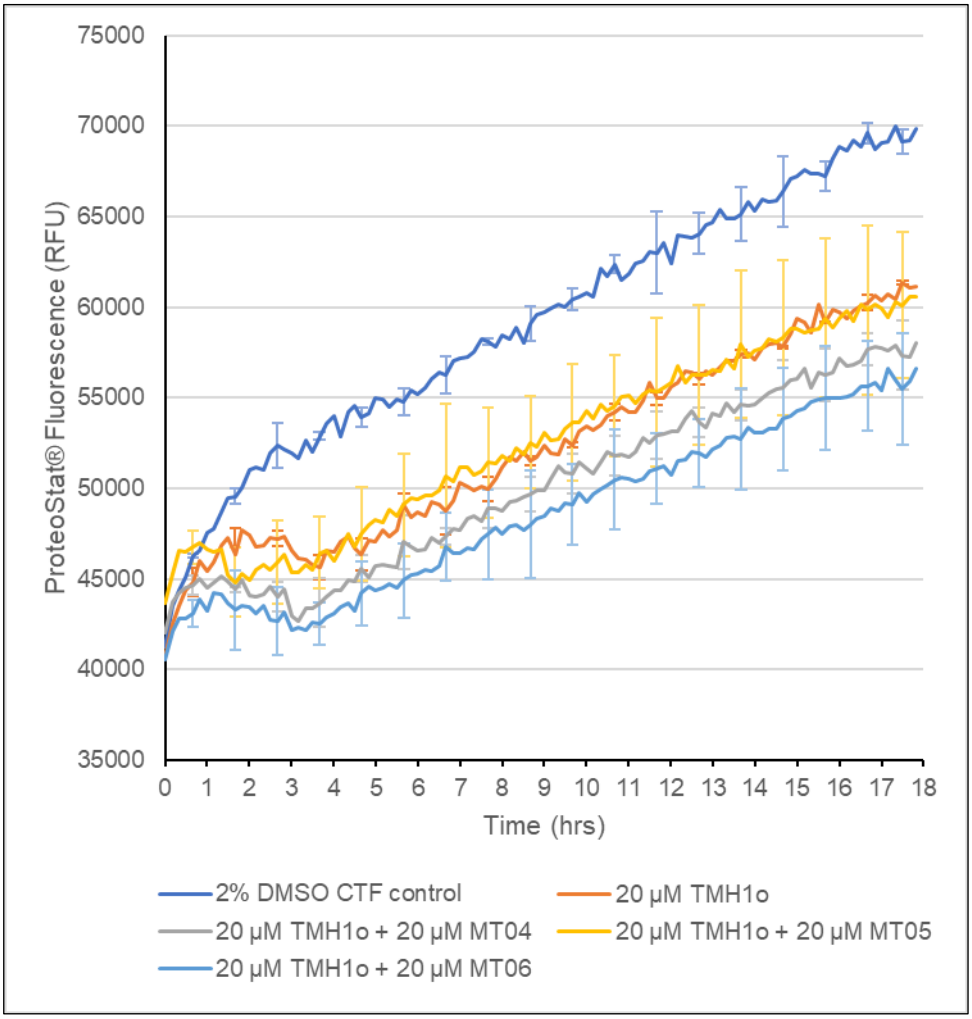


Figure 7.38. The effect of peptide inhibitor combinations on the TDP-43 CTF aggregation assay. ProteoStat® aggregation assay, showing ProteoStat® fluorescence with time of aggregating 20 μM TDP-43 CTF with and without the addition of TMH1o inhibitor at 20 μM, and 20 μM, of MT04, MT05 and MT06. Lines represent mean fluorescence of triplicate reactions and error bars represent standard deviation of the means.

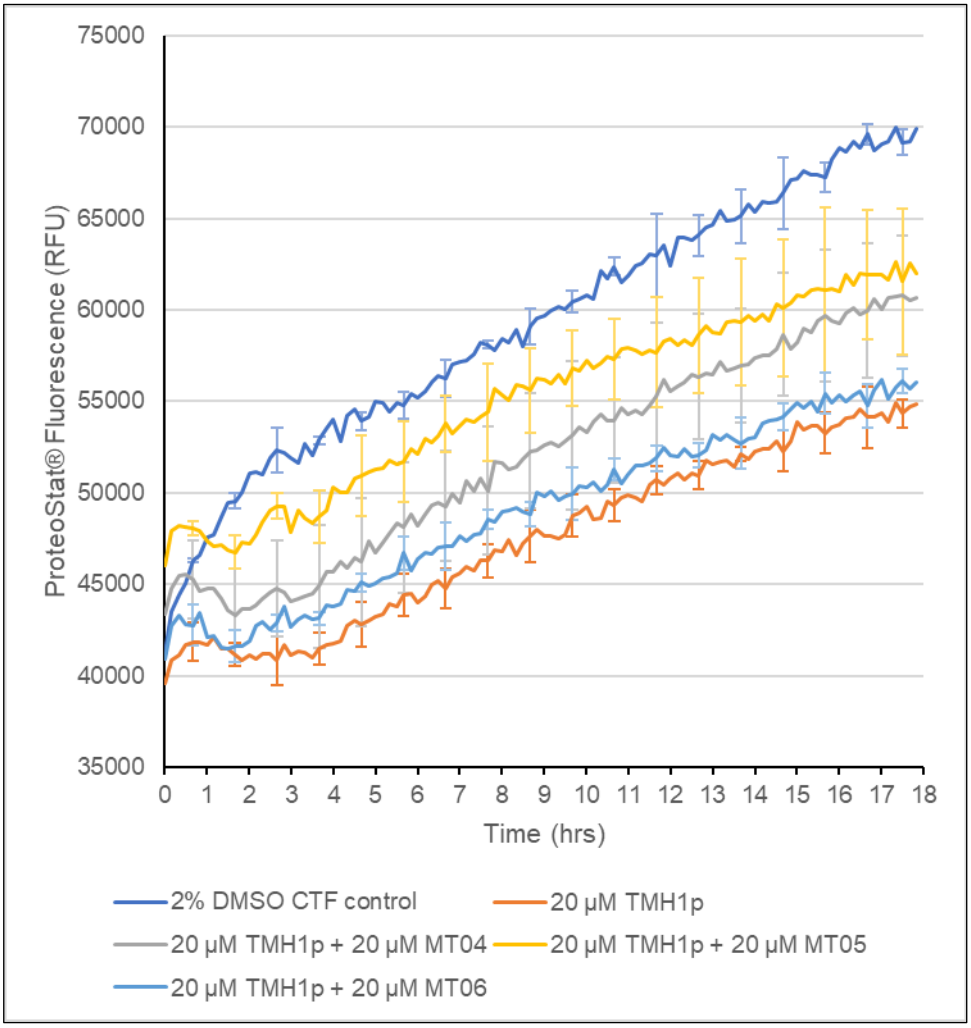


Figure 7.39. The effect of peptide inhibitor combinations on the TDP-43 CTF aggregation assay. ProteoStat® aggregation assay, showing ProteoStat® fluorescence with time of aggregating 20 µM TDP-43 CTF with and without the addition of TMH1p inhibitor at 20 µM, and 20 µM, of MT04, MT05 and MT06. Lines represent mean fluorescence of triplicate reactions and error bars represent standard deviation of the means.

7.4 Conclusion

This chapter presents an investigation into the repurposing of generic aggregation inhibitory molecules for TDP-43 and the design of specifically targeted peptide aggregation inhibitors. The TDP-43 CTF aggregation assay developed in the previous chapter was used to investigate potential inhibitory activity.

From the aggregation assay and corresponding transmission electron microscopy, most of the generic aggregation inhibitor molecules identified in the literature had minimal apparent effect when tested against the TDP-43 CTF, with some even accelerating aggregation. However, there was some evidence that rosmarinic acid may have some effect on the aggregation over a longer time period and at high concentrations compared to the TDP-43 CTF (10x molar excess). The lipid PEG2000PE appeared to have some effect in the aggregation assay, with lower concentrations reducing the fluorescence in the assay and leading to minimal or no increase after an initial burst of approximately 3-4 hours. TEM imaging did not however identify any changes to the aggregates produced.

The first two generations of specifically designed peptide aggregation inhibitors were unsuccessful at inhibiting aggregation, with higher concentrations of first-generation and equimolar concentrations of second-generation inhibitors clearly accelerating the TDP-43 CTF aggregation. Furthermore, the second-generation inhibitors led to morphological changes in the aggregates, some of which resembling amyloid fibrils.

The third-generation inhibitors designed using AlphaFold repeatedly led to a reduction in the endpoint fluorescence of the TDP-43 CTF aggregation reactions, with no increase in the fluorescence for the first 3 hours of the assay. TEM imaging identified the aggregates produced to have a variety of different morphologies, indicating the protein may aggregate differently in the presence of these peptides.

8. Yeast TDP-43 proteinopathy model

8.1 Introduction

Any potentially aggregation inhibitory molecules identified in *in vitro* aggregation assays would need further testing in cell-based models. This is required to determine to what extent the inhibitors may work in a cellular environment, and also to investigate whether they are toxic to cells, either by themselves or by conferring a toxic affect by their actions on TDP-43.

Several higher eukaryote and human cell models have been generated to investigate the relationship between TDP-43 aggregation and cellular toxicity, as reviewed in Hergesheimer *et al*., (2019), however not all show the formation of TDP-43 aggregates and not all show aggregates with corresponding toxicity. These models vary in the details of the TDP-43 recombinant protein expressed, including wild-type protein, proteins with ALS associated mutations, and ΔNLS constructs. For example, Kitamura *et al*. (2016) and Watanabe *et al*. (2012) both used a Neuro2A cell model that did not show aggregates but did show toxicity when expressing wild type TDP-43. This may be due to the overexpression of the TDP-43 being inherently toxic, and potentially causing toxicity before the formation of aggregates could occur. It may also reflect the other cellular changes that may occur in TDP-43 proteinopathies that may lead to aggregation not being present in these cell models.

Contrastingly to the human and rodent cell models generated, multiple yeast cell models expressing wild type TDP-43 have shown TDP-43 cytoplasmic aggregate formation, with an associated decrease in cell viability (Johnson *et al*., 2009; Armakola *et al*., 2011; Prasad *et al*., 2016; Liu *et al*., 2017). Of course, the intracellular environment of a yeast cell is unrepresentative of that of a neuron, and given that yeast do not have a TDP-43 homologue, it may be that the foreign protein may be toxic regardless of aggregation (Hergesheimer *et al*., 2019).

However, one clear example of the relevance of the yeast model to TDP-43 aggregation studies is that of the case of ataxin 2. Ataxin 2 is a polyglutamine repeat containing protein, whose yeast analogue, Pbp1 was found to be a modifier of TDP-43 aggregation related toxicity in yeast cells with Pdp1 loss of function supressing toxicity (Elden *et al*., 2010). Following on from this finding, it was identified that ataxin 2 modulated TDP-43 toxicity in *Drosophila*, (Elden *et al*., 2010) and therapeutic reduction of ataxin 2 reduces pathology and extends lifespan in mice expressing human TDP-43 (Becker *et al*., 2017).

This chapter presents the production of a “druggable” *Saccharomyces cerevisiae* cell model of TDP-43 aggregation, with distinct intracellular aggregation and growth reduction that can be used for the testing of TDP-43 aggregation inhibitor molecules.

8.2 Expression of fluorescently tagged TDP-43 in *S. cerevisiae*

Two yeast expression plasmids encoding fluorescently tagged TDP-43 were acquired from Addgene. pRS416Gal TDP43 WT YFP and pRS426Gal TDP43 WT GFP were a gift from Aaron Gitler (Addgene plasmid # 27447 and # 27467 respectively). These plasmids express C-terminally fluorescently labelled (YFP or GFP) wild-type human TDP-43 under the control of a galactose inducible GAL promoter. The two plasmids are maintained at differing copy numbers (pRS416Gal TDP43 WT YFP low copy plasmid, pRS426Gal TDP43 WT GFP high copy plasmid), resulting in two different levels of expression upon induction. Plasmids were purified from cultures grown from the purchased bacterial stabs, and their integrity verified by sequencing.

An *ERG6* deletion mutant (Δerg6), with uracil auxotrophy (allowing selection of transformed cells containing a plasmid conferring uracil prototrophy) (Horizon Discovery) was transformed with the purified plasmids. Considerations of a suitable *S. cerevisiae* strain had to be made due to the desire to use the yeast as a model to test

the activity of aggregation inhibitors. The lipid composition of the *S. cerevisiae* cell membrane limits its permeability to small compounds, but for inhibitors to be tested in the yeast model, they must be able to enter the cell. Deletion of the *ERG6* gene has been shown to increase membrane permeability to small molecules (Emter *et al*., 2002), by inhibiting ergosterol biosynthesis (Gaber *et al*., 1989; Kryndushkin and Shewmaker, 2011). *ERG6* deletion mutants can therefore be used for chemical drug testing in yeast proteinopathy models, as in the testing of AIM4 against TDP-43 aggregation (Prasad *et al*., 2016).

Transformants were selected by growth on Synthetic Complete media without Uracil (SC-U) plates, as the plasmids confer uracil prototrophy. An empty pYES2 plasmid vector was used as a control, providing the Δerg6 strain with uracil prototrophy in the same way. To screen for successful transformation with the plasmids expressing fluorescently tagged proteins, transformed colonies were picked from the selective plate then resuspended in sterile water. The resuspended colony was then spotted onto an SC-U plate containing glucose (for repression of TDP-43 expression), and onto a corresponding region on one containing galactose (for induction of TDP-43). Following incubation of the plates for 24 hours, the resulting colonies on galactose induction plates were imaged using the AxioZoomV.16 fluorescence microscope, to check for GFP or YFP fluorescence from the colonies. Examples of this fluorescence are shown in Fig. 8.1. Fluorescent colonies were selected and picked from the corresponding colony on the glucose plate to be grown up for liquid culture.

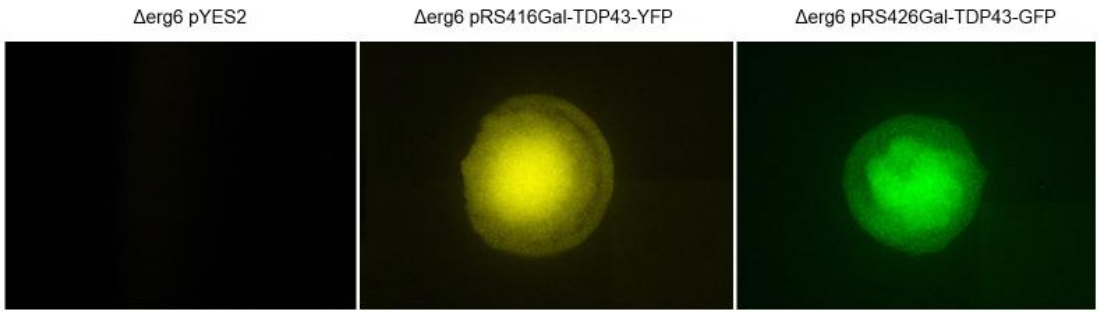


Figure 8.1. Colony fluorescence of Δerg6 S. cerevisiae grown on galactose induction plates. Transformed with pYES2 vector, pRS416Gal-TDP-YFP (YFP fluorescence) and pRS426Gal-TDP-GFP (GFP fluorescence) imaged using the AxioZoomV.16.

To further confirm expression of fluorescently tagged TDP-43, transformed Δerg6 cells were grown in liquid culture, and induced for 8 hours in SC-U media containing 2% galactose. Lysates were separated into soluble and insoluble fractions and subjected to SDS-PAGE and immunoblot, the results of which are shown in Fig 8.2. Although minimally visible in the SDS-PAGE gel of Fig. 8.2 (a), the presence of a band at approximately 72kDa found in the 8Hr insoluble fractions (lanes e and i), identified by the corresponding immunoblot, indicated the expression of fluorescently tagged TDP-43. Furthermore, as the band was only detected in the insoluble fractions, this indicated that TDP-43 may form insoluble inclusions in the yeast cells.

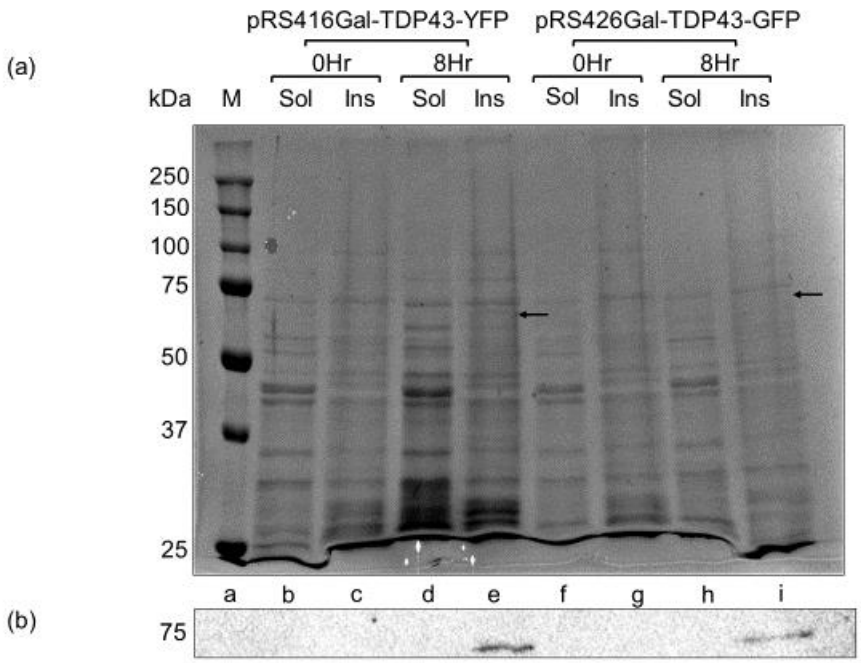


Figure 8.2. Expression and solubility of fluorescently tagged TDP-43. (**a**) SDS-PAGE gel of the soluble (Sol) and insoluble (Ins) cell fractions of Δerg6 S. cerevisiae transformed with pRS416Gal-TDP43-YFP and pRS426Gal-TDP43-GFP before and after 8Hr induction in 2% galactose medium. M (lane a) Precision Plus Protein™ All Blue protein standards (Bio-Rad); lanes b and c, soluble and insoluble fractions respectively of pRS416Gal-TDP43-YFP before induction (T=0); lanes d and e, soluble and insoluble fractions respectively of pRS416Gal-TDP43-YFP following 8 hours of induction; lanes f and g, soluble and insoluble fractions respectively of pRS426Gal-TDP43-GFP before induction (T=0); lanes h and i, soluble and insoluble fractions respectively of pRS426Gal-TDP43-GFP following 8 hours of induction. Arrows indicate ~72kDa band corresponding to fluorescently tagged TDP-43 (**b**) Corresponding immunoblot, probed with rabbit polyclonal anti-TDP-43 antibody (Invitrogen).

8.3 Expression of TDP-43 in *S. cerevisiae* results in intracellular aggregation

To identify the presence of any TDP-43 aggregation in the yeast cells, fluorescence microscopy was carried out, using the Leica Stellaris confocal microscope. Following 6-hour expression in galactose SC-U media, fixed cells, stained with DAPI, were imaged using a 100x objective, using GFP, YFP and DAPI filters. Representative images are shown in figures 8.3, 8.4 and 8.5. As shown in Fig. 8.3, Δerg6 *S. cerevisiae* transformed with a pYES2 vector (allowing growth in SC-U media), showed no apparent YFP fluorescence (Fig.8.3. (a)) but a small amount of GFP background fluorescence consistent with GFP autofluorescence of yeast cells (Fig. 8.3. (b) and (c)).

Cells transformed with the pRS416-Gal-TDP43-YFP low copy plasmid, following expression of YFP- tagged TDP-43 showed clear YFP fluorescence of intracellular foci (Fig. 8.4), with many cells demonstrating a single aggregate (Fig. 8.4 (c) and (d)), and some showing two or three distinct aggregates (Fig. 8.4 (a) and (b)). Expression of GFP tagged TDP-43 by Δerg6 *S. cerevisiae* transformed with the pRS426-Gal-TDP43-GFP high copy plasmid resulted in cells containing multiple distinct fluorescent aggregates as shown by Fig. 8.5. This series of images confirmed the presence of fluorescently tagged TDP-43 aggregates in the Δerg6 *S. cerevisiae* strain, following 6-hour protein expression.

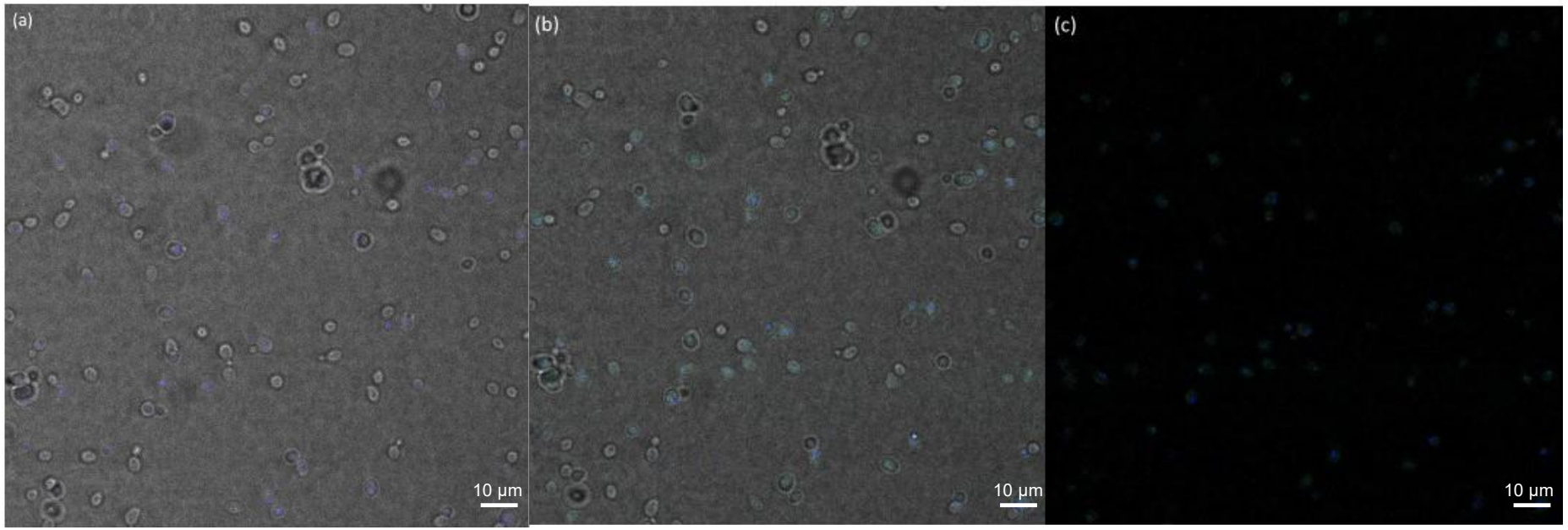


Figure 8.3. Composite fluorescence microscope images of Δerg6 S. cerevisiae cells transformed with a pYES2 vector. (a) brightfield, DAPI (blue) and YFP (green) channels, (b) brightfield, DAPI (blue) and GFP (green) channels and (c) DAPI (blue) and GFP (green) channels. All images taken using the Leica Stellaris confocal microscope using the 100x oil immersion objective. Images prepared using Fiji (ImageJ).

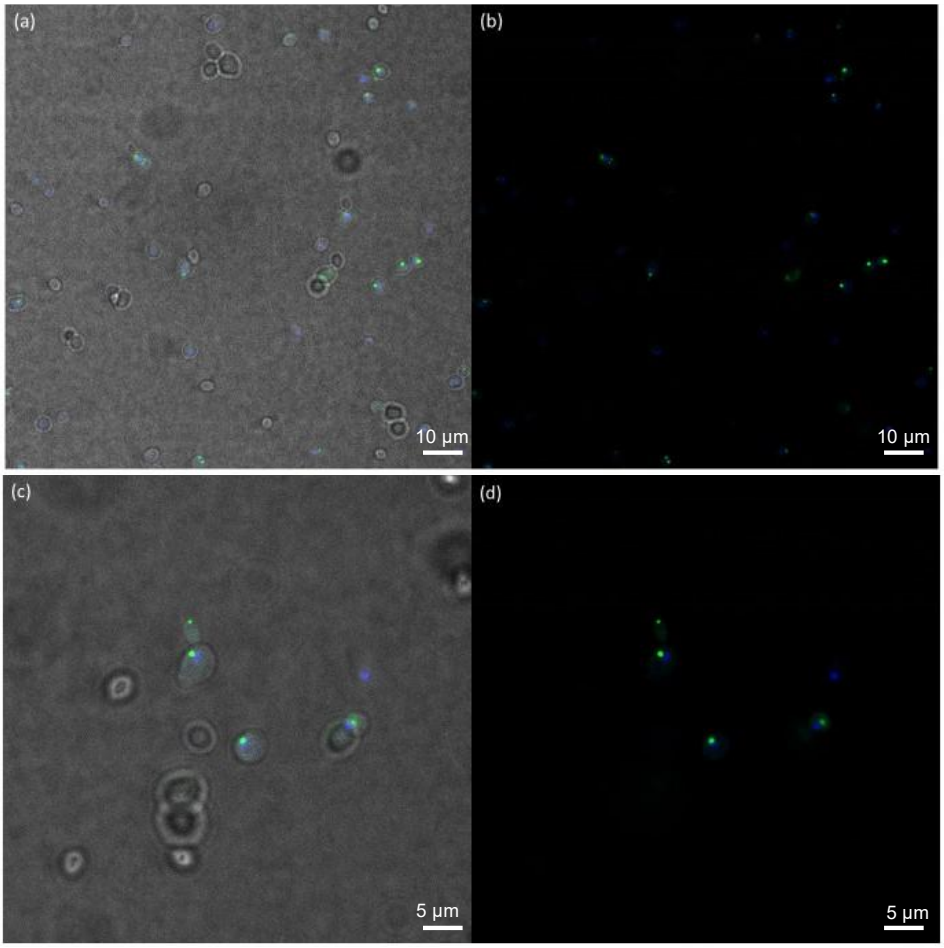


Figure 8.4. Composite fluorescence microscope images of Δerg6 S. cerevisiae transformed with pRS416-Gal-TDP-43-YFP following 6 hr expression in 2% galactose media. Composite images of (a) brightfield, DAPI (blue) and YFP (green) channels, (b) DAPI and YFP channels, (c) 2x Leica imaging software zoom using brightfield, DAPI and YFP channels, (d) 2x Leica imaging software zoom using DAPI and YFP channels. Images prepared using Fiji (ImageJ).

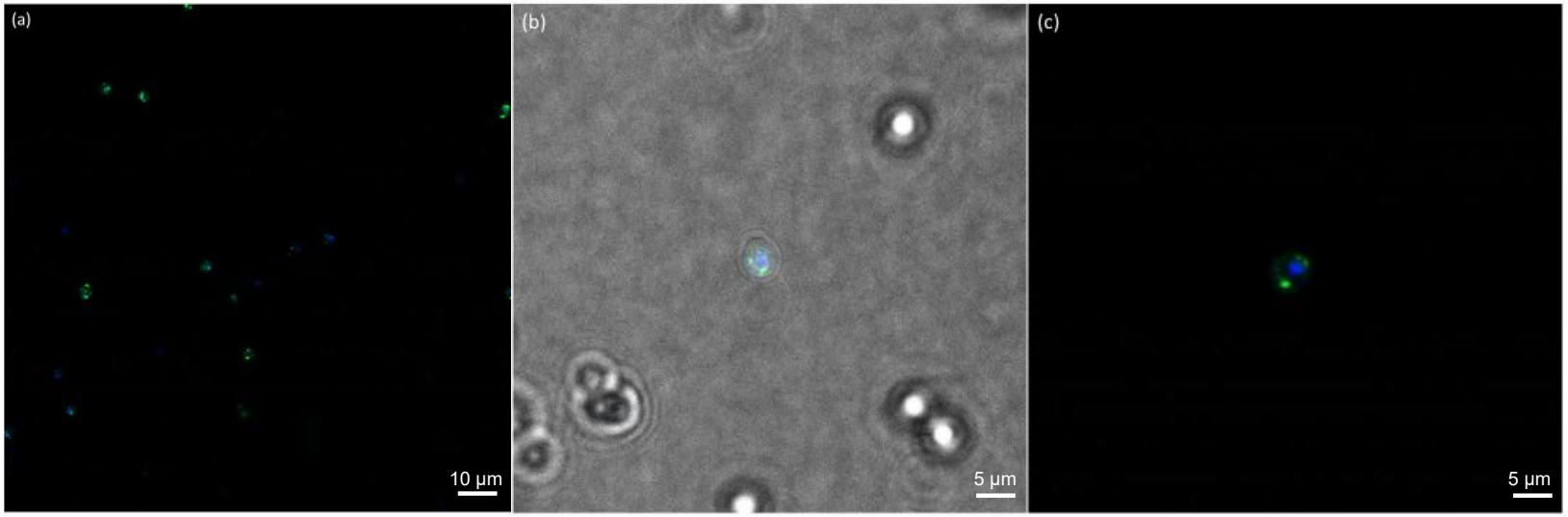


Figure 8.5. Composite fluorescence microscope images of Δerg6 S. cerevisiae transformed with pRS426-Gal-TDP-43-GFP following 6 hr expression in 2% galactose media. Composite images of (a) DAPI and GFP channels, (b) 2x Leica imaging software zoom using brightfield, DAPI and GFP channels, (c) 2x Leica imaging software zoom using DAPI and GFP channels. Images prepared using Fiji (ImageJ).

8.4 Growth reduction of *S. cerevisiae* following TDP-43 induction

To determine whether TDP-43 expression and aggregation correlated with a reduction in cell growth in this model, as demonstrated in the literature (Johnson *et al*., 2009) (Prasad *et al*., 2016), spotting assays and yeast growth curve assays were carried out. As shown in Fig. 8.6, the spotting assay demonstrated reduced growth of *S. cerevisiae* when grown on induction media plates, for both the pRS416Gal-TDP43-YFP low copy and the pRS426Gal-TDP43-GFP high copy plasmids, when compared to yeast transformed with an empty pYES2 plasmid control (allowing growth on the selective media). Repression of TDP-43 expression on glucose media resulted in a similar level of growth across all plasmid and strain combinations. In this assay, the growth of the Δerg6 strain was compared to another laboratory expression strain with a similar background and auxotrophic mutations, INVSc1 (Invitrogen). This was to determine whether the Δerg6 mutation resulted in any difference in growth on the different medias, with and without TDP-43 expression.

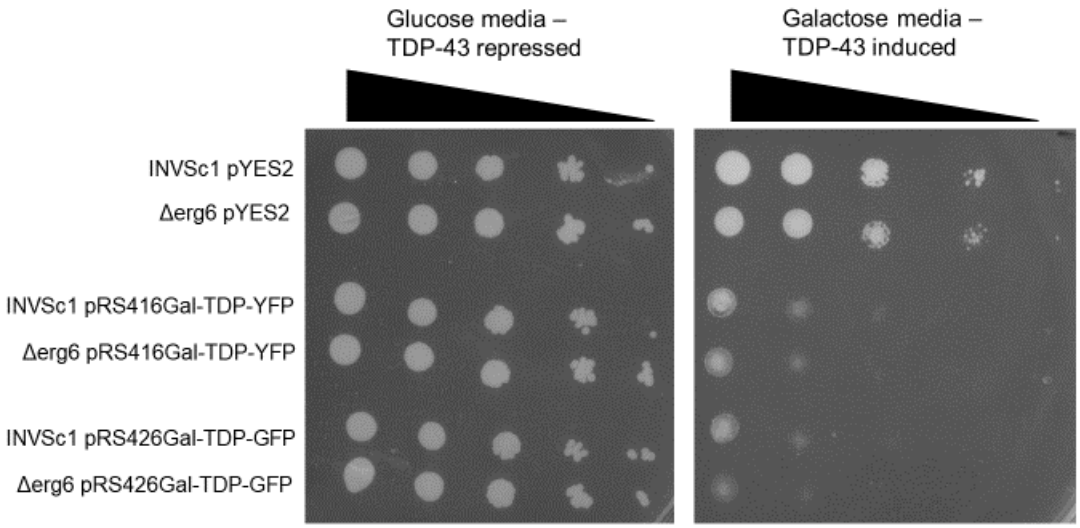


Figure 8.6. Spotting assay to identify growth reductions correlating to TDP-43 expression. Images show colony growth of 5µl of serial 1 in 10 dilutions of cultures normalised to an OD600 of 0.7, spotted onto either a repression or an induction media plate.

Figures 8.7, 8.8 and 8.9 show growth curves of Δerg6 *S. cerevisiae* transformed with either the control pYES2 plasmid, or the expression plasmids under various inducing and non-inducing conditions. The control pYES2 plasmid growth curves (Fig. 8.7) show the growth of a non-expression strain when provided media containing different sugar components. The varying raffinose/galactose concentrations were trialled in non-expression cells (pYES2 vector cells) due to the slower rate of growth of *S. cerevisiae* in galactose compared to raffinose and glucose media. As shown by the orange (0.1% galactose) and grey (0.5% galactose) curves, replacement of raffinose with galactose in the growth media for the non-expression strain appears to slow growth, however across 72 hours, the same culture end density is achieved. This delay to the growth is due to the need to activate genes required for galactose metabolism.

The growth curves for Δerg6 *S. cerevisiae* expressing fluorescently tagged TDP-43 from either a low copy (Fig. 8.8) or a high copy plasmid (Fig. 8.9) both show that

induction with both 0.1% and 0.5% galactose in the growth media reduces growth rate such that the end density of the cultures following 72 hours does not reach that of a non-induced culture grown in raffinose alone. Furthermore, the induced cultures do not appear to enter a distinct log growth phase as can be seen with the non-induced cultures and the pYES2 cultures. These growth curves did not appear to show any increased growth inhibition with expression from the pRS426Gal-TDP43-GFP high copy plasmid compared to the pRS416Gal-TDP43-YFP low copy plasmid, despite the marginal observable difference in the colony growth of these in the Fig. 8.6 spotting assay. However, as the two expression strains are grown from different starter cultures, which may have been in slightly different phases of their own growth curves when diluted to OD600 of 0.1, direct comparison between the strains is not necessarily a perfect indication of their viability under expression/non-expression conditions.

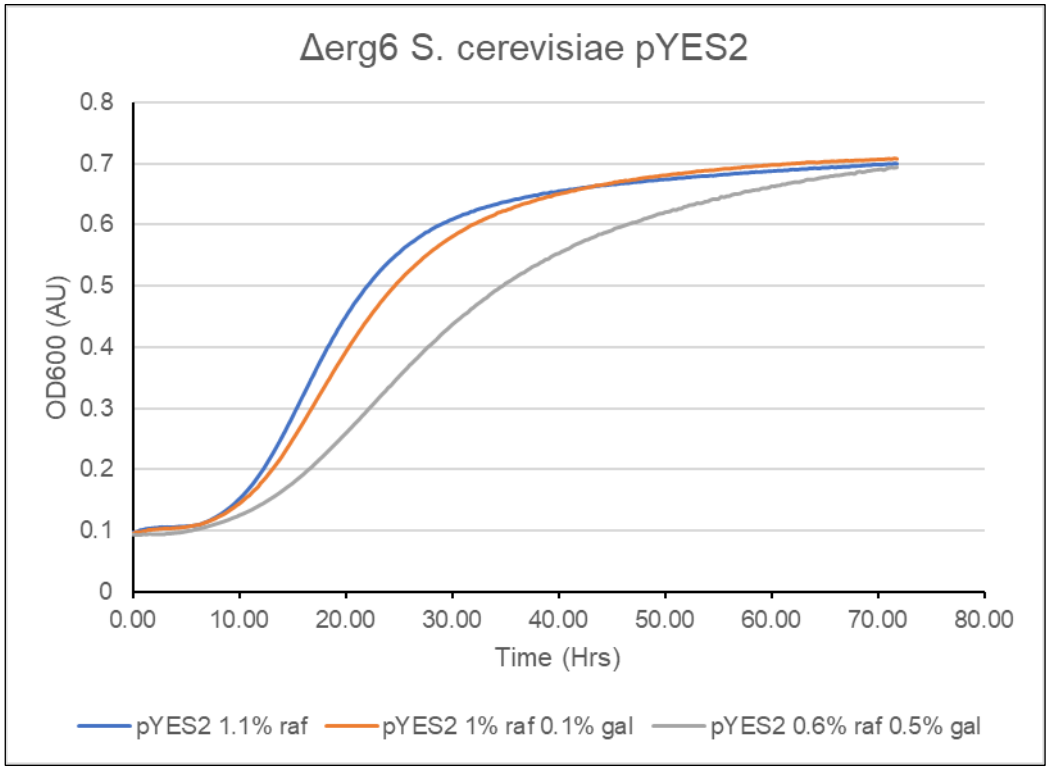


Figure 8.7. 72 hour growth curves of pYES2 control Δerg6 S. cerevisiae. Cells grown under three SC-U media conditions- 1.1% raffinose (blue), 1% raffinose with 0.1% galactose (orange), and 0.6% raffinose with 0.5% galactose (grey).

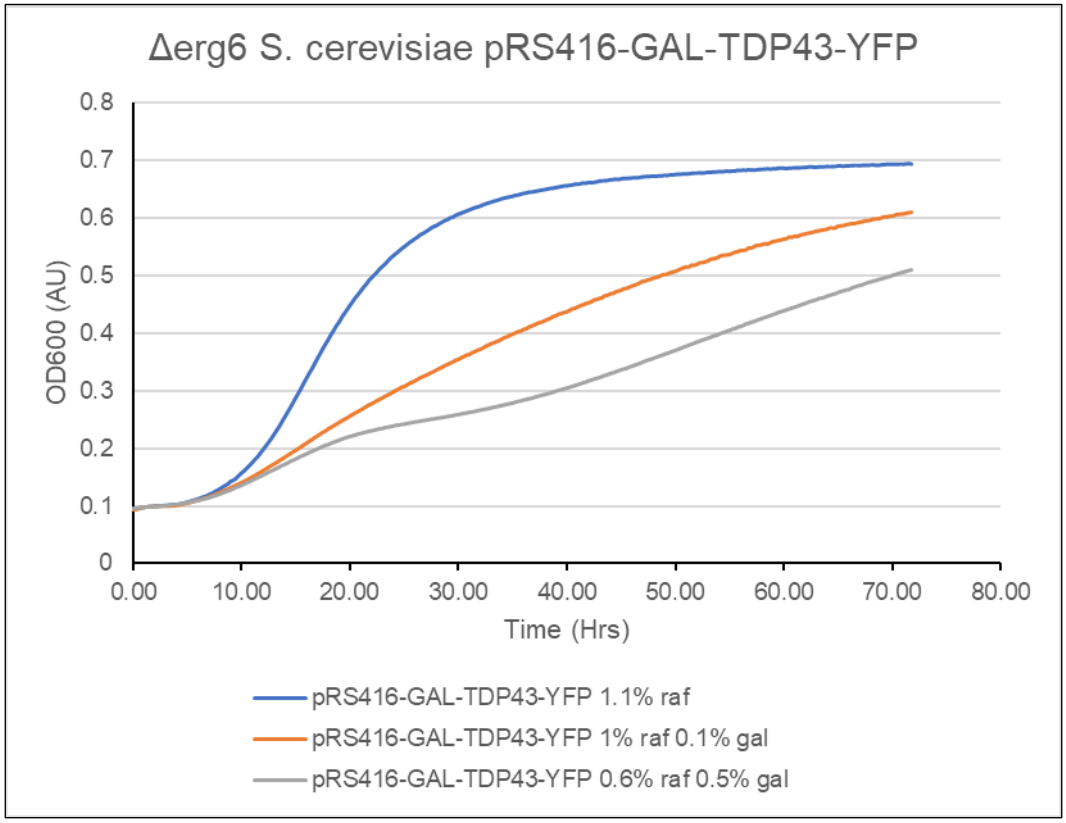


Figure 8.8. 72 hour growth curves of Δerg6 S. cerevisiae transformed with the pRS416Gal-TDP43-YFP plasmid. Cells grown under three SC-U media conditions- 1.1% raffinose (blue), 1% raffinose with 0.1% galactose (induction)(orange), and 0.6% raffinose with 0.5% galactose (induction)(grey).

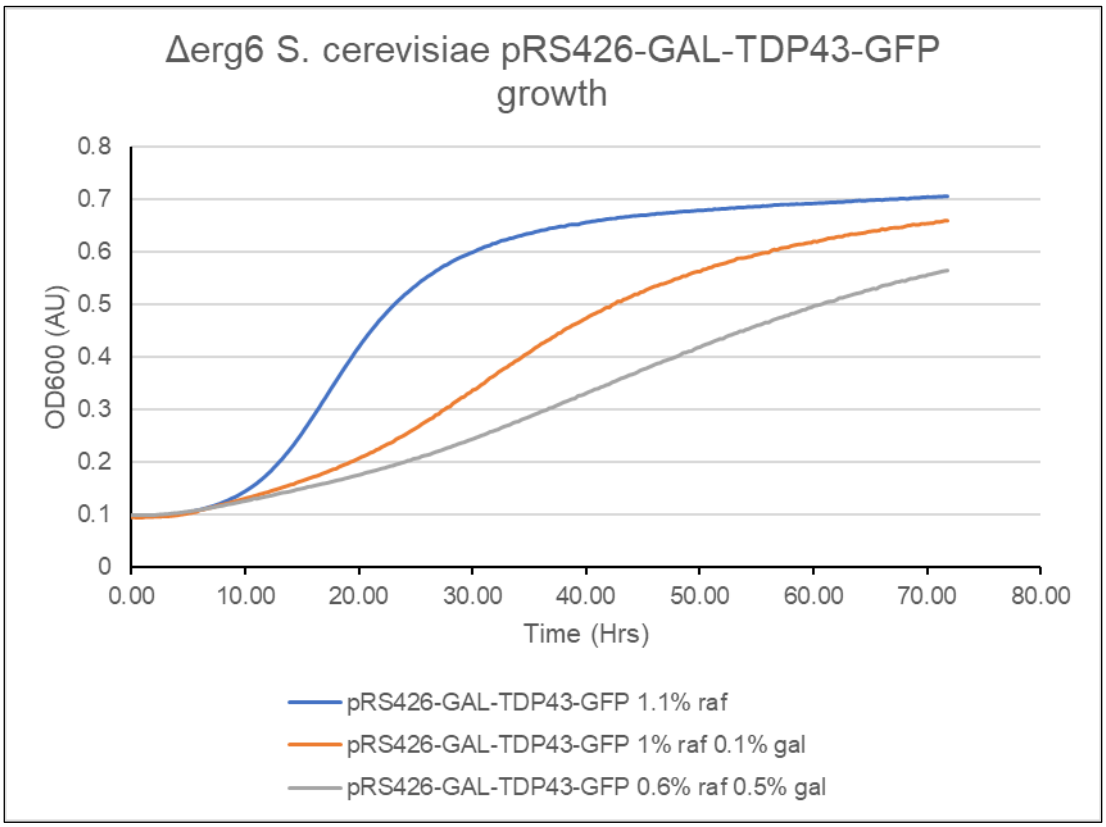


Figure 8.9. 72 hour growth curves of Δerg6 S. cerevisiae transformed with the pRS426Gal-TDP43-GFP plasmid. Cells were grown under three SC-U media conditions- 1.1% raffinose (blue), 1% raffinose with 0.1% galactose (induction) (orange), and 0.6% raffinose with 0.5% galactose (induction) (grey).

Further fluorescence microscopy was then carried out, to confirm that aggregation still occurred under the media conditions used in the growth curves. Figure 8.10 confirms that aggregation occurred in the low copy plasmid expression cells with both the (a) 1% raffinose 0.1% galactose or (b) 0.6% raffinose 0.5% galactose media.

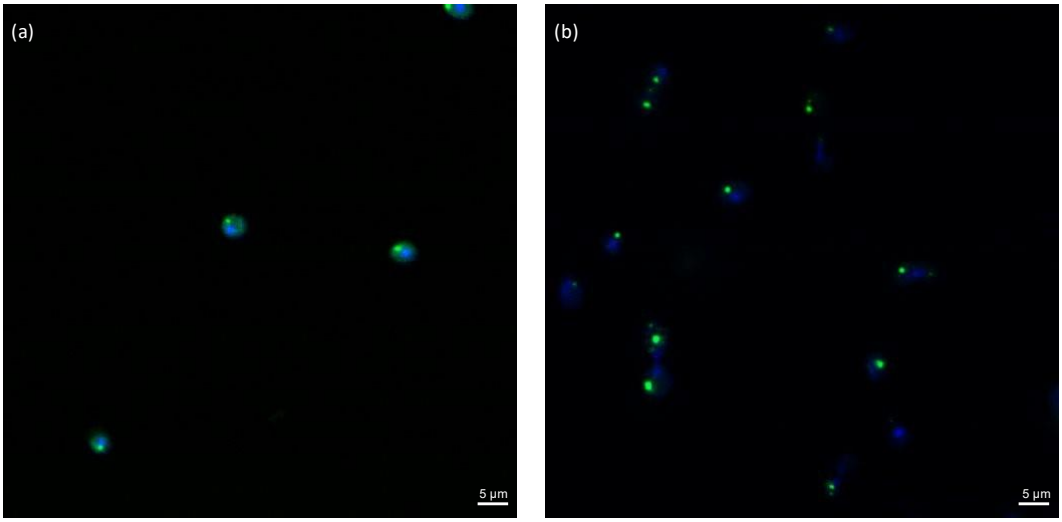


Figure 8.10. Composite fluorescence microscope images of Δerg6 S. cerevisiae transformed with pRS416-Gal-TDP-43-YFP following 6 hr expression. (a) 1% raffinose 0.1% galactose or (b) 0.6% raffinose 0.5% galactose media under DAPI (blue) and YFP (green) channels. Images taken using 2x Leica imaging software zoom and composites prepared from z-stacks using Fiji (ImageJ).

8.5 Treatment with polyphenols and third-generation peptide TDP-43 aggregation inhibitors does not prevent TDP-43 expression related growth inhibition

Following the model development, the yeast could be used for testing molecules for therapeutic potential. Modulation of TDP-43 aggregation for example by AIM4 (Prasad *et al*., 2016) and by *Pdb1* (yeast *ATXN2* homologue) suppression (Elden *et al*., 2010), has been shown to reduce the toxicity of TDP-43 expression.

The growth inhibition caused by TDP-43 expression can be utilised to screen drugs for therapeutic potential. One way of doing this is using the reduced growth on solid media as demonstrated by the spotting assay. Application of a successfully therapeutic drug to the surface of the media should encourage greater growth around that point than seen elsewhere on the plate. This method has been used to screen for small molecules potentially therapeutic against mammalian prions that cause toxicity in a yeast model (Bach *et al*., 2003).

For this assay, Δerg6 *S. cerevisiae* expressing YFP-tagged TDP-43 (pRS416-GAL-TDP43-YFP plasmid) were grown from a colony picked from a glucose SC-U plate overnight in SC-U media with 2% glucose. The culture was diluted appropriately from the spotting assay to give clear growth reduction on induction plates and was spread onto 2% glucose (repression) and 2% galactose (expression) SC-U plates. Sterile 6 mm filter paper were placed onto the agar surface and loaded with the compound to be tested. The compound can then diffuse into the agar surface surrounding the disc.

First, curcumin and rosmarinic acid were screened. Despite having minimal apparent effect on the aggregation of TDP-43 CTF in aggregation assays, the yeast model expresses full length TDP-43 which may aggregate from a differently folded/alternately stable starting point than the TDP-43 CTF maintained in urea. Furthermore, regardless of an effect on aggregation, curcumin and rosmarinic acids have been found to have

therapeutic antioxidant, anti-apoptotic and anti-inflammatory effects, (Adami & Bottai, 2020) (Novak *et al*., 2021) (Luo *et al*., 2020), and so it was considered if they could have any therapeutic effect on the growth of the yeast cells by other mechanisms.

Addition of the curcumin and rosmarinic acid to the filter discs did not lead to any noticeable increase in growth surrounding the disc, despite diffusion of the compounds into the media (particularly noticeable with curcumin due to colouration) (Fig. 8.11). Therefore, in this plate-based assay, neither compound appeared to rescue the growth reduction caused by TDP-43 expression. However due to non-uniform growth on the media plate, it is difficult to draw definite conclusions, particularly in the absence of a positive control that could rescue the growth of the yeast cells.

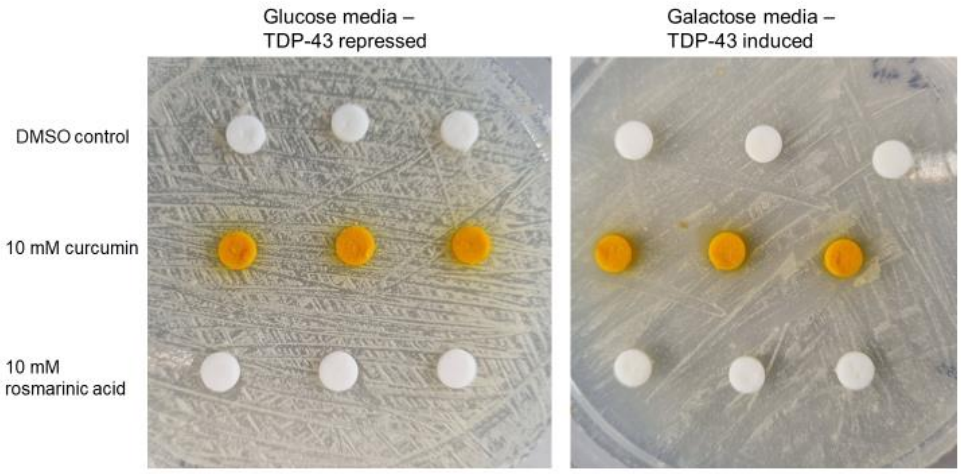


Figure 8.11. Disc diffusion yeast growth assay for polyphenol testing. Yeast plate growth assay showing growth of Δerg6 S. cerevisiae transformed with pRS416-Gal-TDP-43-YFP on glucose (repression) and galactose (TDP-43 induction) media with filter discs loaded with DMSO, 10 mM curcumin or 10 mM rosmarinic acid in triplicate.

Given that they had a noticeable effect on the CTF aggregation, the third-generation peptide inhibitors were also tested by this assay (Fig. 8.12). They, like rosmarinic acid

and curcumin, did not however appear to lead to any rescued growth on the induction plate.

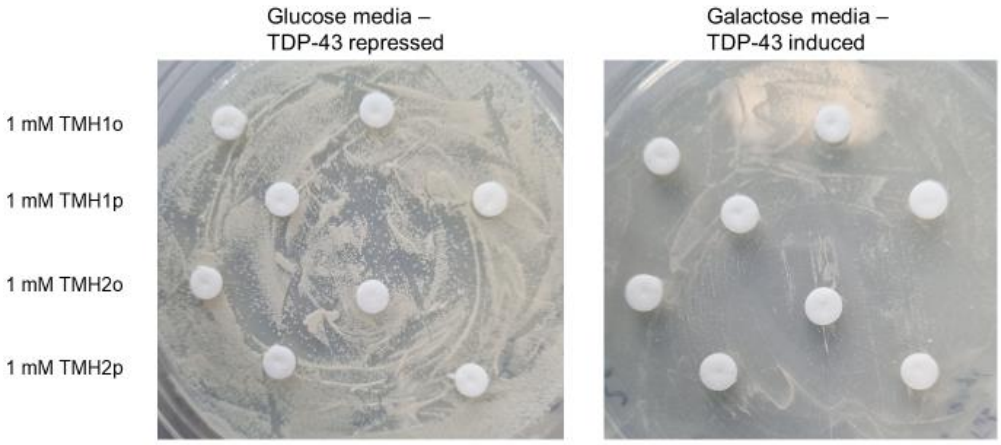


Figure 8.12. Disc diffusion yeast growth assay for peptide inhibitor testing. Yeast plate growth assay showing growth of Δerg6 S. cerevisiae transformed with pRS416-Gal-TDP-43-YFP on glucose (repression) and galactose (TDP-43 induction) media with filter discs loaded with 1 mM TMH1o, TMH1p, TMH2o and TMH2p in duplicate.

Growth curves were also attempted to confirm these findings, however no useable data was produced. There were several reasons for this- for curcumin, at a low enough DMSO concentration to not inhibit the growth of the yeast, curcumin became insoluble which affected the OD600 used to monitor yeast growth. Furthermore, due to an as yet undetermined cause, the majority of the yeast cultures, including non-expression and untreated controls, were not completing their growth curves and instead appeared to die off during the log phase. Troubleshooting could not be completed within the time constraints of the project, and so further work would be required to test any inhibitors in the growth curve assay.

8.6 Conclusion

This chapter presents results that show successful establishment of a *Saccharomyces cerevisiae* cell model of TDP-43 proteinopathy that can be used to investigate potential inhibitors of TDP-43 aggregation. A Δerg6 strain of yeast was selected to allow the cell model to be “druggable” with a cell membrane more penetrable to small molecules.

The yeast model produced, with both high copy and low copy expression plasmids demonstrate intracellular aggregates of fluorescently labelled TDP-43 following inducible expression. The expression of TDP-43 leads to reduced growth on both solid growth media and in liquid cultures.

Initial testing of polyphenols and third-generation peptide inhibitors did not lead to reduced toxicity from TDP-43 expression on solid media plates, but further work in liquid culture and fluorescence microscopy is required to confirm these findings.

9. Discussion

9.1 The solubility of recombinant untagged TDP-43

Bacteria are routinely used as a system in which to express recombinant proteins that can be then purified for further study. However, bacterial expression is not without its difficulties. One problem with over expressing recombinant proteins in bacterial systems is the accumulation of proteins into insoluble inclusion bodies. Due to the high level of expression driven by inducible promoters present in plasmid expression vectors such as the T7 promoter used in this project, the proteins produced may “overwhelm” the quality control system of bacteria, particularly at higher expression temperatures and with proteins, such as TDP-43, that are aggregation prone. The prokaryotic cells lack the intracellular compartments and organelles, and so some of the quality control systems, of eukaryotic cells, and so expression of eukaryotic proteins in prokaryotic cells can be a challenge (Schein, 1989). Therefore, it is no surprise that at higher expression temperatures used in this project, which encourage rapid growth and protein expression, TDP-43 was found in the insoluble fraction of the cell extract with the over-expressed TDP-43 accumulating into inclusion bodies. This occurred despite the use of SHuffle T7 Express *E. coli,* a strain engineered with an enhanced capacity to fold proteins, particularly those containing disulphide bonds. A reduction in expression temperature is often used to reduce inclusion body accumulation, and so this was tested. The low temperature slows the growth rate and metabolic activity of the bacteria, producing protein at a slower rate which may have more chance to fold and remain soluble rather than accumulate into inclusion bodies (Schein, 1989). This has successfully produced some soluble TDP-43 protein as previously reported in the literature, in combination with amino acid substitutions (Wright *et al*., 2020) or a reduction in IPTG concentration for induction (Carlomagno *et al*., 2014). Here however reduction to 16ºC did not result in more soluble TDP-43 being produced. This may be due to the pET expression system used being optimised for high level expression, but

is most likely due to the nature of TDP-43 itself, with its aggregation propensity and LLPS properties as discussed in the literature review.

Although not investigated in this project, the use of alternative expression hosts could be one method to generate soluble protein. Eukaryotic expression hosts could provide improved folding and post translational modifications that may aid solubility. One such host are yeast cells, including *Saccharomyces cerevisiae* and *Pichia pastoris*, though as shown in this project generating a TDP-43 aggregation cell model, expression in *S. cerevisiae* still results in the formation of inclusions that appear to be present in the cell insoluble fraction. Another expression host, capable of producing biologically active, folded and post-translationally modified mammalian proteins is the baculovirus infected insect cell system (Contreras-Gómez *et al*., 2013). In this system, baculovirus expression vectors engineered with the cDNA of the protein of interest are used to infect cultured insect cells, which in turn express the desired protein. Insect cells themselves can carry out most of the refolding and post-translational modifications of a mammalian cell, but can also be engineered to co-express chaperones and mammalian enzymes for specific post-translational modifications (Contreras-Gómez *et al*., 2013). There are currently no reports of successful purification of TDP-43 produced using this system.

Given the insolubility of TDP-43 in *E. coli*, new methods to purify the protein had to be considered. One potential approach considered was to express a TDP-43 construct with an attached solubility tag. Previous work showed that the ~55 kDa NusA solubility tag allowed production of soluble protein in a BL21 strain expression system (Humphreys, 2013). However, removal of the tag proved difficult, with a low efficiency of cleavage of the tag by HRV 3C protease and difficulty in purifying the cleaved protein from the protease and un-cleaved protein. The maltose binding protein (MBP) tag has been shown to also produce soluble TDP-43 at a 16ºC expression temperature, but cleavage of the tag triggers instant aggregation, indicating the TDP-43 section of the

fusion protein is still inherently unstable in this construct, and may not therefore be aggregating from a truly native start point (Wang *et al*., 2018). Recently, soluble TDP-43 has been expressed and purified using a SUMO tag, resulting in a monomeric TDP-43 protein with a high propensity to form oligomers and aggregates (Kumar et al., 2023). Due to the initial aim of using the protein in aggregation experiments to test aggregation inhibitors, we preferred to use pure TDP-43 protein with no residual tags, proteases or cleavage products that may affect aggregation. Therefore, an approach to purify untagged TDP-43 from inclusion bodies was required.

9.2 Purification of TDP-43 from bacterial inclusion bodies

9.2.1 TDP-43 purification under native conditions following inclusion body refolding

To purify TDP-43 from bacterial inclusion bodies, two approaches could be taken, both beginning with the solubilisation of the inclusion bodies by denaturation. The first approach was to solubilise the inclusion bodies, refold the protein mixture produced, and then purify TDP-43 in the absence of denaturants. This method was selected first, as it was hypothesised that the characteristics of the refolded protein could be exploited to purify it. These included its affinity for metal ions (Garnier *et al*., 2017), its affinity for nucleic acids, for which heparin chromatography can be used, and its affinity for hydroxyapatite (Humphreys, 2013). Furthermore, should it dimerise as reported in the literature (Vega *et al*., 2019), SEC could be used to separate it from any co-purifying fragments, or proteins with a similar size to monomeric TDP-43.

This approach however was unsuccessful. As demonstrated in chapter 3, all chromatographic methods tested following first step refolding had little effect on the purity of the sample. This appeared to be due to the formation of contaminated oligomeric species during the refolding method when attempted as an initial step. It

therefore appeared that for successful refolding of TDP-43, it had to be pure. However, as shown in chapter 4, refolding of pure TDP-43 produced by preparative SDS-PAGE by the shock dilution method presented in Vega *et al*, (2019) still resulted in the formation of large oligomers/aggregates. Therefore, this method and buffer combination may not be suitable for refolding of an untagged TDP-43 construct. Should a refolding protocol be developed that can maintain TDP-43 in a monomer/dimer/low order oligomer state following refolding of the inclusion body mixture, then the approach of purification by exploiting the innate properties of TDP-43 investigated here could be used in future.

9.2.2 TDP-43 purification under denaturing conditions

The second approach was to purify the protein under denaturing conditions and refold the resulting pure stock. This is more akin to the approaches by Vega *et al*., (2019) and Doke and Jha (2022). Vega *et al.,* (2019), utilised a Pel-B and His-tagged recombinant protein for IMAC purification under denaturing conditions, following which the protein was refolded and purification completed with an SEC polishing step. Doke and Jha, (2022), also used a His-tagged TDP-43 recombinant protein, purified by IMAC and anion exchange chromatography under denaturing conditions before refolding.

Due to the absence of an affinity tag, methods to purify TDP-43 under denaturing conditions were limited. Methods based on protein charge (anion exchange), metal ion chelation (IMAC) and size (SEC and preparative SDS-PAGE) were investigated. Anion exchange chromatography had minimal effect on the purity of TDP-43, with little separation of the protein from contaminants and causing a large dilution of the sample by elution occurring across the entire salt gradient applied. Furthermore, many contaminants remained in the sample. The predicted isoelectric point of TDP-43 is 5.85, with the average isoelectric point of the *E. coli* proteosome being 6.6 (Kozlowski, 2016). At the initial pH used of pH 8, based on net negative charge TDP-43 should

have a higher affinity for the anion exchange column than the average bacterial protein, however many of these will clearly still carry the net negative charge required to bind the column. A pH reduction to 7.5 was tested with no noticeable change in the results (not shown), but a further reduction could have been tested closer to the average bacterial protein pI which could have reduced the affinity of these contaminants to the column, resulting in a cleaner end product. The elution of TDP-43 across the entire salt gradient is consistent with the results shown for anion exchange of TDP-43 protein obtained previously (although this was carried out under native conditions), (Humphreys, 2013) which suggests varying levels of affinity for the anion exchange column among the TDP-43 in the sample.

Although metal chelation does not necessarily require a protein to be folded (demonstrated by the ability to use IMAC under denaturing conditions for His-tagged proteins), it does require enough chelating amino acids to be close enough together in sequence to coordinate the metal ion. As shown in chapter 3, TDP-43 could bind to Cu^{2+} and Ni^{2+} columns under denaturing conditions, but only weakly as elution began at the very start of the application of an imidazole gradient. It did not bind to a Zn^{2+} column under these conditions. An explanation for the lack of binding is that the chelating regions identified in RRM1 and RRM2 are thought to be spread across 8-12 amino acids, and are part of the folded structure of the RRM domains (Garnier *et al*., 2017; Golovin *et al*., 2020). Therefore, they may not be close enough together in the denatured structure to successfully and strongly chelate metal ions.

Following optimisation, preparative SDS-PAGE was found to be a method by which pure TDP-43 could be produced in a single step. Although the method used here with staining, excision and elution of the band is time consuming and is limited in the quantity of pure protein that can be produced in a given time, sufficient protein could be produced for refolding and initial characterisation studies. There is equipment available for larger scale, more automated preparative SDS-PAGE, such as the BioRad

Model 491 Prep Cell. This is a larger scale preparative SDS-PAGE system, with a greater loading capacity, where protein bands elute from the ends of a gel column and can be collected by a fraction collector, which would allow preparative SDS-PAGE of TDP-43 to be scaled up. Furthermore, elution from the gel would be 100%, whereas by passive diffusion used in this project, some protein will remain in the gel slice. Another alternative to passive elution is to use electroelution chambers to elute the protein from the gel slice. This again could have been more efficient than passive elution.

9.3 Refolding of TDP-43

Chapter 4 presents the refolding of TDP-43 purified by preparative SDS-PAGE. As discussed previously, two refolding methods for TDP-43 have been published in the literature (Vega *et al.*, 2019; Doke and Jha, 2022), but neither approach was successful for the stock of protein purified by SDS-PAGE despite use of the buffers published in these papers. One reason for this could be due to the upstream purification methods clearly differing, with sample from SDS-PAGE produced in this thesis potentially containing residual SDS or acetone. Another explanation may be the presence of tags and extra sequences. Vega *et al.*, (2019) purified a recombinant TDP-43 protein with an additional PelB sequence on the N-terminus as well as a C-terminal His-tag, whereas Doke and Jha, (2022) also used a C-terminal His-tag. There is potential that these could affect the folding and stability of the protein. A C-terminal histidine tag would add a stretch of six hydrophilic, polar residues to the aggregation prone, hydrophobic CTD region, while an N-terminal PelB sequence adds 22 amino acids to the N-terminus of the domain thought to be responsible for oligomerisation (NTD). To test this, 6-His and PelB tagged constructs could be purified and refolded the same way to identify any differences in the resulting protein.

From the testing of artificial chaperone assisted refolding using detergent and cyclodextrin, the findings from the SDS/MPD refolding method and the final method developed of refolding during SDS-PAGE elution, it is apparent that detergent is important to maintain the solubility and aid in the refolding of TDP-43. Stripping of detergent from TDP-43 by cyclodextrin resulted in protein precipitation. Furthermore, the buffers taken from the Doke and Jha, (2022) shock dilution refolding method contain no detergent, with refolding using these buffers also appearing to result in precipitation. The requirement of the presence of detergents in buffers to maintain the solubility of TDP-43 is consistent with reports in the literature. Sodium lauroyl sarcosinate (sarkosyl) was required to maintain the solubility of TDP-43 purified from the cell soluble fraction by Wright *et al*. (2020), Vega *et al*. (2019) utilised Triton X-100 in several buffers and Cascella *et al*., (2022) used a combination of lauryldimethylamine oxide (LDAO) and Octyl-β-D-Glucopyranoside (OG) detergents in their buffers to maintain protein solubility.

The purification and refolding method developed could repeatedly produce a stable 200-215 kDa species in the presence of Triton X-100, which was hypothesised to be a TDP-43 tetramer. Although the predicted molecular weight by amino acid composition of a TDP-43 monomer is 44.7 kDa, and so a tetramer would theoretically be 178.8 kDa, size exclusion chromatography separates protein molecules by their frictional coefficient/Stoke's radius rather than their molecular weight from amino acid sequence (Erickson, 2009).

Doke and Jha (2022) found a monomeric form of TDP-43 to be predicted at 55 kDa by size exclusion chromatography- suggesting a tetramer could exist as large as 220 kDa. This could be explained by the highly extended structure of TDP-43, particularly in its CTD, as shown in the TDP-43_{WtoA} structure (Wright *et al*., 2020), resulting in a larger structure in three dimensions than a more compact globular structure such as in transferrin and ovalbumin that were used to produce the SEC calibration curves. A

protein with an expanded, open structure, such as TDP-43, can elute from a size exclusion column far earlier than a compact globular protein of the same molecular weight, (even at a position twice the expected molecular weight), and so an SEC calibration curve generated with globular proteins will not give an accurate estimation of the molecular weight of a protein with expanded structure (Erickson, 2009).

The development of the refolding method as presented in chapter 4 demonstrates that the inclusion of Triton X-100 in the refolding buffer is required for the production of a tetramer (and potentially the other oligomeric states seen) rather than a species that could either be denatured TDP-43 or a soluble larger aggregate as determined from its elution from an SEC column. This therefore brings in to question the method by which the protein is refolding. The buffer was originally designed containing SDS and MPD, to act as detergent and modulating cosolvent as in Michaux *et al*., (2008) and Roussel *et al*., (2012). However, the necessity of Triton X-100 as a second detergent to be included in the buffer indicates that the refolding could instead be proceeding by detergent competition. In this method, an excess of a non-ionic, non-denaturing detergent outcompetes the protein for binding to SDS, releasing the protein from an SDS-denatured state and allowing refolding (Kaspersen *et al*., 2017). An almost identical method of preparative SDS-PAGE followed by single step elution and renaturation into a Triton X-100 containing refolding buffer has been presented for DNA-binding proteins in the literature (Ossipow *et al*., 1993). In this method it was the Triton X-100 detergent competition that resulted in the refolding of active DNA binding proteins.

9.4 Characterisation of TDP-43

Results presented in chapter 5 demonstrate characterisation of the TDP-43 purified and refolded by preparative SDS-PAGE and one step elution/refolding. Chemical crosslinking by glutaraldehyde initially demonstrated the presence of an apparent TDP-43 tetramer, crosslinking to a band of approximately 158 kDa on an SDS-PAGE gel. With glutaraldehyde crosslinked proteins sometimes demonstrating an increased rate of migration on SDS-PAGE gels than may be expected (Griffith, 1972), this could represent a tetramer with predicted molecular weight of 178.8 kDa. This band corresponded to the species eluting from analytical SEC at a position predicted to be a species 200-215 kDa. This therefore confirms that the analytical SEC carried out with a column calibrated with globular proteins was likely not accurate for predicting the molecular weight of a protein with expanded structure such as TDP-43, as discussed in 9.3. Later crosslinking experiments in alternative buffers originally designed for CD spectroscopy demonstrated the presence of dimers and monomers also in the sample.

This mixture of oligomeric states is contrasting to the PelB and His-tagged dimeric state purified by Vega *et al*., (2019), and the monomer by Doke and Jha (2022). To date, there are no reports in the literature of a tetramer of purified recombinant full-length TDP-43, however recombinant NTD (amino acids 1-102) has been shown to exist in a tetramer-dimer equilibrium (Herrera *et al*., 2023), or for a slightly shorter purified NTD protein (1-77), a monomer (Mompeán *et al*., 2016). At higher concentrations, N-terminal fragments (residues 1-261 and 1-89) have been shown to tetramerise (Jiang *et al*., 2017), as has RRM2 under low salt conditions (Kuo *et al*., 2009). Furthermore, a spectrum of oligomeric states, including monomers, dimers, trimers, tetramers and higher order states have been observed in human cells (Afroz *et al*., 2017).

The differences seen in this work compared with Vega *et al*., (2019) who purified a dimer and Doke and Jha (2022) who purified a monomer could again be due to the

absence of any tag, the alternative method of purification, or simply the different buffer conditions used. Doke and Jha (2022) themselves demonstrated how different buffer conditions could lead to the formation of a higher order (larger than tetrameric) oligomer, and Herrera *et al*., (2023), showed various oligomerisation states of the NTD depending on buffer conditions, including reducing environment, concentration and also the size of the NTD recombinant protein used.

Circular dichroism spectroscopy revealed indications of a protein structure containing both α helical and β sheet content, as would be expected from the solved folded structures of TDP-43. The spectrum produced is similar to those produced by Doke and Jha (2022) for their protein in both monomeric and oligomeric states, but differs from the CD spectrum of the PelB–TDP-43–His dimer produced by Vega *et al*. (2019), which appears to show more pronounced helical content with a prominent minimum at 208 nm. This may be due to the oligomerisation state being different, which could result in an altered secondary structure within the oligomer. The N-terminal PelB sequence of the dimer from Vega *et al*., (2019) could also have an effect on the spectrum, as the PelB peptide is predicted by AlphaFold to be helical.

The DNA binding assay presented in chapter 5 indicates that the protein produced may have greater affinity for a TG repeat ssDNA probe than an AC repeat probe which would be consistent with reports in the literature suggesting that TDP-43 has a high affinity for TG/UG rich sequences (Lukavsky *et al.*, 2013; Buratti and Buralle, 2001). This assay however requires further optimisation, as the immunoprecipitation step did not appear to be efficient in pulling down TDP-43. This may have been due to the buffer components. The fluorescence of the Cy3 DNA probe bound to the nylon membrane for the pellet fractions was fairly weak, requiring a longer exposure to reveal the fluorescence differences between the samples than for the supernatant fractions. This indicates there was much less Cy3 probe pulled down than remained in the supernatant. This could be due to the inefficiency of the IP step. It could also be related

to the tetramer being the predominant species in the sample. Lukavsky *et al*., (2013) presented a structure derived from NMR data whereby a nucleic acid strand is bound by a monomeric RRM1 and RRM2 construct, while Jiang *et al*., (2017), showed dimerization to be vital for splicing activity, and Kuo *et al*, (2009) suggested dimerization of RRM2 occurs during nucleic acid binding. Therefore, it may be that the low levels of dimer and monomer observed in the CD buffer could be the species responsible for DNA binding. Further investigation into the DNA binding ability of the TDP-43 produced in this project is required, and the presented results should be taken as preliminary. Other assays such as an EMSA (electrophoretic mobility shift assay) could be carried out, as in Vega *et al*., (2019), which suggested their dimer to have DNA binding ability.

9.5 Purification and aggregation of TDP-43 CTF

Chapter 6 presents the purification of TDP-43 CTF (208-414) and the development of an assay to monitor its aggregation. TDP-43 CTF was purified from the insoluble cell extract under denaturing conditions, using IMAC and SEC to produce a stock in 4 M urea. This is similar to the method used to purify TDP-43 CTF for use in aggregation assays by Prasad *et al*., (2016), but with an additional SEC step to improve purity. The protein could then be aggregated by dilution of the urea and incubation with shaking. A criticism of this method could be that the presence of urea could mean that the protein is not in a native state at the start point of the reaction, and so may aggregate differently than if it were folded. However, a large portion of this TDP-43 CTF has minimal structure, aside from the section of RRM2 at its N-terminus, so it was hypothesised that this method could still produce a useful assay for inhibitor testing.

The resulting aggregation assay reliably monitored the increase of aggregation with time using the aggregate binding dye ProteoStat®. Thioflavin T appeared to show a large amount of variation in its fluorescence when added to identical repeats of the

same aggregation reaction and furthermore required a very high plate reader gain setting to produce fluorescence, indicating the aggregates are only weakly ThT fluorescent. This is consistent with reports in the literature indicating TDP-43 aggregates are minimally ThT fluorescent (Johnson *et al*., 2009; Saini and Chauhan, 2011; Kumar *et al*., 2023). ProteoStat® has been used to monitor the aggregation of RRM domain constructs (Chen *et al*, 2019), and here it is also successful in monitoring the aggregation of the TDP-43 CTF containing a portion of RRM2. One potential limitation of the assay is the initial "burst" of fluorescence seen on urea dilution that did not seem to be avoidable, resulting in a starting fluorescence of the assay of around 30-40,000 RFU. It is possible that there could be some aggregates formed in the 4 M urea stock that rapidly bind to ProteoStat® once the urea is diluted, producing this burst phase. TEM imaging of the stock could help determine this, but this has not been investigated.

The aggregates formed with this TDP-43 CTF construct under these conditions differ morphologically to the "worm-like" aggregates shown for a CTF by Saini and Chauhan (2011), but are similar to aggregates of full length Q331K mutant TDP-43 (Johnson *et al*., 2009), and the "oval" morphological features observed with full length TDP-43 aggregated in the presence of an anti TDP-43 antibody as a potential aggregation inhibitor (Esposto and Martic, 2021). They also show some similarity to "short and sturdy filament" aggregates of a TDP-43 peptide of residues 279-360 seeded with a brain derived aggregate extract (Kumar *et al*., 2023). The aggregates are distinctly different to those produced by full length wild type TDP-43 (Johnson *et al*., 2009; Kumar *et al*., 2023), indicating that the TDP-43 CTF may aggregate differently.

9.6 Aggregation inhibitor testing

9.6.1 Generic aggregation inhibitor testing

Following development of a TDP-43 CTF aggregation assay, testing of potential inhibitor molecules was carried out. First, generic inhibitors, molecules reported in the literature to have inhibitory ability against other aggregating proteins, were tested. The polypeptides polyarginine, polyornithine and poly-glutamic acid appeared to have varying effects on the aggregation of the TDP-43 CTF, in some assays accelerating aggregation and slightly reducing aggregation in others, often with a large spread of data, particularly for the poly-cations. Furthermore, the poly-cations appeared to have a precipitation effect on the TDP-43 CTF, while the opposite was seen for polyglutamic acid, indicating that a charge effect may destabilise the TDP-43 CTF. TEM imaging did not demonstrate any changes to the morphology of the aggregates produced.

The polyphenols curcumin and rosmarinic acid also did not have a clear effect on the aggregation of the TDP-43 CTF. Despite an apparent reduction in the ProteoStat® fluorescence in the aggregation assay, it became clear that both polyphenols interfered with the assay in some way. Curcumin has been shown to interfere with the fluorescence of ThT in assays involving amyloid-beta (Hudson *et al*., 2009), and so a similar phenomenon could be occurring here for curcumin and rosmarinic acid interference with the ProteoStat® assay. There was minimal difference between the TEM images obtained from curcumin and rosmarinic acid treated samples and controls during the 18-hour assay. However, when extended to 48 hours, aggregates with a more granular morphology and a lack of large masses of aggregates were seen in a rosmarinic acid treated sample compared to controls, indicating that over longer time periods and at high concentrations, rosmarinic acid may have some effect on TDP-43 CTF aggregation. The lack of any effect of curcumin on the TDP-43 CTF aggregation may be consistent with results presented in the literature- only the curcumin derivatives dimethoxy curcumin and monocarbonyl dimethoxycurcumin C and not curcumin itself

reduced the aggregation of a TDP-43 fragment TDP-25 in NSC34 cells and reduced hyperexcitability induced by TDP-43 expression in a motoneuron-like cellular model of ALS (Duan *et al*., 2013; Dong *et al*., 2014). To date, no literature has reported the testing of rosmarinic acid as a TDP-43 aggregation inhibitor.

The lipid PEG200PE at lower concentrations than the TDP-43 CTF appeared to reduce aggregation, with no increase in ProteoStat® fluorescence being seen following the initial 3-4 hours of the assay. TEM did not demonstrate a change to the morphology of the aggregates. However, as this imaging is not quantitative, it could be that there was a reduction in the total number of aggregates being formed without a change to the morphology of those that are produced. PEG2000PE could therefore be tested further.

9.6.2 Targeted TDP-43 aggregation inhibitor peptide testing

Peptide inhibitors specifically targeting TDP-43 aggregation sequences were then designed and tested. A similar approach was initially taken to previously developed peptide inhibitors for amyloid-beta and tau (Austen *et al*., 2008; Taylor *et al*., 2010; Aggidis, 2019), that is, producing short peptides to bind to specific amino acid sequences in the protein target that have been identified as important in promoting aggregation. The six first-generation peptide inhibitors designed based on two aggregation sequences determined by Saini and Chauhan (2011) had minimal effect on the TDP-43 CTF aggregation at equimolar concentrations, with increased inhibitor concentrations accelerating aggregation. One explanation for this could be the inclusion of the peptide into the aggregate at an alternative binding site to its target as suggested by the AlphaFold multimer prediction in Fig. 7.27, where the peptide is included in the formation of a beta sheet in the folded region of the TDP-43 CTF. The second-generation inhibitors were then designed with an added octa-arginine sequence to the C-termini of the first-generation peptides. The rationale for this was that the introduction of a charged, hydrophilic stretch of sequence might improve

interference between two TDP-43 CTF molecules. In fact, this modification had the effect of rapidly accelerating the aggregation of the TDP-43 CTF to aggregates with altered morphology. The positively charged arginine sequence may have acted as electrostatically attractive to the target regions and the TDP-43 CTF as a whole which at the reaction conditions of pH 8 would be net negatively charged. As previously seen with polyarginine and polyornithine, the addition of poly-cations appears to destabilise the protein.

AlphaFold was then used to design potential aggregation inhibitors based on its predicted points of interactions between TDP-43 CTF molecules. A point of interaction around the RRM2 domain part of the TDP-43 CTF protein was predicted, part of the residue 246-258 sequence identified by Saini and Chauhan (2011). The 246-258 sequence EDLIIKGISVHIS was predicted to form an antiparallel sheet around the lysine residue with the hydrophobic GISVHIS sequence somewhat exposed. This GISVHIS sequence was predicted to be the point of interaction between TDP-43 CTF molecules and it was hypothesised that this point of interaction could be blocked to inhibit aggregation. A “hairpin” peptide was designed by sequentially modelling a binding sequence and hydrophilic stretch that could turn back on itself and be stapled into position by a di-bromo-xylene linker. Four versions of this peptide were produced, with two binding sequences and two lengths of linker, and tested against the aggregating TDP-43 CTF. The peptides repeatedly demonstrated a reduction in the aggregation of the TDP-43 CTF in the assay, with a noticeable 3-hour delay in the aggregation increase. TEM imaging demonstrated the aggregates produced to have an array of different morphologies, indicating that these peptides may lead to the TDP-43 CTF aggregating differently, in a way that maybe proceeds more slowly with the 3-hour delay.

Comparison with the effects of some aggregation inhibitors reported in the literature can be made. The generic aggregation inhibitor and acridine derivative AIM4 showed

a 37% reduction in aggregation of a TDP-43 CTF construct at equimolar concentrations of protein and inhibitor (Prasad *et al*., 2016). By normalising the data from Fig. 7.33 in the same way as Prasad *et al*., (2016), that is taking the end point of the untreated control as 100% fluorescence, and the lowest start point as 0% fluorescence, the best reduction seen here at equimolar concentrations for TMH1p is a 49.6% reduction in end point fluorescence. This normalisation represents the measured fluorescence increase across the assay time, negating the experimental dead-time "burst" of fluorescence seen. The inhibitor peptide recently designed by Kamagata *et al*., (2023) by using statistical binding energies between peptides and their target of the TDP-43 C-terminal domain, had minimal effect on the aggregation of TDP-43 at equimolar concentration (1 μM protein and peptide), but at a 50 times molar excess could reduce the aggregation of TDP-43 by 44%.

When assessing the effectiveness of the inhibitors, the limitations of the assay should be taken into account. Firstly, the presence of the 1.5 M urea, while preventing immediate precipitation of the TDP-43 CTF, could interfere with interactions between inhibitor molecules and their targets. On the other hand, it may be preventing more aggressive aggregation that the inhibitor may have a different effect on. Furthermore, with the 1.5 M urea present, there is a chance that the section of RRM2 in the TDP-43 CTF may not reliably fold, so the protein may not aggregate from a native start point, and in fact not self-associate in the way predicted by AlphaFold as used for third-generation peptide design. A second limitation is the concentration of protein required for detectable ProteoStat® fluorescence. The assay here was carried out at 20 μM protein concentration, whereas TDP-43 is thought to exist at up to 3 μM within a human cell (although localised concentrations in the cytoplasm during aggregation may be higher) (Wang *et al*., 2018). Phase separation and increase in turbidity of a solution of TDP-43 purified with a solubilising MBP tag can occur rapidly at even 2.5 μM following tag cleavage (Wang *et al*., 2018). The high concentration used in this assay is therefore

more than six-fold that which would occur in a cell for full-length protein. Furthermore, the TDP-43 CTF fragment may be even more aggregation prone than the full-length protein, given its lack of structure and the fact that most identified aggregation sequences lie in this region of the protein (Fig. 1.10). Another assay that could utilise lower protein concentrations and lower urea concentrations could be used to further investigate the inhibitors. An example of this is the sedimentation analysis used by Kamagata *et al*., (2023), utilising only 1 μM of aggregating protein.

9.7 Yeast cell model

Chapter 8 reports the generation of an *S. cerevisiae* cell model of TDP-43 aggregation. Like the model first reported by Johnson *et al*., (2008), this model demonstrates clear intracellular aggregation of fluorescently labelled TDP-43, and produces growth inhibition on both plate media and in liquid culture. This growth reduction is reported to be due to toxicity of the accumulating TDP-43 to the yeast cells rather than simply growth inhibition (Johnson *et al*., 2008). It has been shown that reduction of TDP-43 aggregation in yeast cells by Pbp1 (yeast ataxin 2 analogue) suppression or by chemical treatment with aggregation inhibitors such as AIM4 can rescue this inhibition, providing growth assays to test aggregation inhibitors (Elden *et al*., 2010; Prasad *et al*., 2016). Due to the requirements of using the model for drug testing, a version of the model was created using an *S. cerevisiae* strain with a deletion of the *ERG6* gene, which leads to an increased permeability of the yeast cell membrane to small molecules by inhibition of the ergosterol biosynthesis pathway (Emter *et al*., 2002; Gaber *et al*., 1989; Kryndushkin and Shewmaker, 2011). By use of this Δerg6 mutant strain, a model was produced which could be “druggable” for the testing of aggregation inhibitors.

For testing in the yeast model, candidates with inhibitory ability against purified protein needed to be identified first using the TDP-43 CTF aggregation assay. If a potential inhibitor was directly tested in the yeast model and proved to be non-functional it could be due to a number of reasons, including degradation by the yeast cell or lack of transport into the cell where it could be affective. Given that few candidates appeared to have potential against TDP-43 aggregation in the TDP-43 CTF aggregation assay, minimal drug testing was carried out in the yeast model. Furthermore, with the specifically designed inhibitors being peptides, and other molecules being ionic polypeptides, these would likely be degraded by yeast proteases, both those secreted into the growth media and those intracellularly (Ogrydziak, 1993). As shown by the testing of the third-generation peptides on the growth plate in the disc diffusion assay, the peptides that caused an apparent reduction in TDP-43 CTF aggregation had no rescuing effect on the growth reduction of the yeast cells following TDP-43 expression. Unfortunately, liquid culture growth curves could not be generated due to issues with culture survival in the plate readers, and troubleshooting is required to generate these for further evidence and inhibitor testing. Once the liquid culture issue is resolved, fluorescence imaging of the yeast cells can also be carried out to identify any aggregation inhibition in the cells. It is likely however that the peptides would need modification to allow them to be proteolytically stable, such as via retroinversion (reversed sequence in d-amino acid form) or by N-methylation. They may also need modification to allow them to penetrate the yeast cell, either with the addition of a cell penetrating peptide sequence or by addition to a liposome as a delivery system.

Curcumin and rosmarinic acid were also tested in the yeast model, with no effect in the disc diffusion assay. Again, no liquid culture data could be obtained. Despite having little apparent effect on TDP-43 CTF aggregation, these molecules were tested due to their therapeutic antioxidant, anti-apoptotic and anti-inflammatory effects, (Adami & Bottai, 2020; Novak *et al*., 2021; Luo *et al*., 2020), in case they could aid yeast survival

in the presence of TDP-43 expression. One reason for their ineffectiveness could be the poor bioavailability of the compounds, particularly curcumin, and this could be improved by addition to liposomes, with curcumin-decorated liposomes themselves having been shown to inhibit amyloid peptide aggregation (Taylor *et al*., 2011). However, only the curcumin derivatives dimethoxy curcumin and monocarbonyl dimethoxycurcumin C and not curcumin itself reduced the aggregation of a TDP-43 fragment TDP-25 in NSC34 cells and reduced hyperexcitability induced by TDP-43 expression in a motoneuron-like cellular model of ALS (Duan *et al*., 2013; Dong *et al*., 2014). Further work could involve the testing of curcumin and rosmarinic acid liposomes in the TDP-43 CTF aggregation assay and yeast model.

9.8 Conclusions and future work

This project has developed a method for the purification and refolding of recombinant full-length, untagged TDP-43 expressed in *E. coli.* The protein produced appears to exist predominantly as a tetramer, but dimeric and monomeric states may be observed under varying buffer conditions. CD spectroscopy demonstrates the protein to have secondary structural elements, likely with both α-helical and β-sheet structures. Preliminary DNA binding assays indicate the protein to have some affinity for a TG repeat ssDNA, but further work is required to confirm this. Future work on this area of the project would involve further characterisation of the nucleic acid binding ability of the protein, by optimisation of the co-immunoprecipitation assay or by use of an alternative such as EMSA. Further investigation into buffers that may alter the oligomeric state could be carried out, using alternative salt and crowding agent concentrations. The protein could also be aggregated, and TEM imaging carried out to identify the morphology of the aggregates produced, to see whether they may appear consistent with the literature.

An aggregation assay using TDP-43 CTF (208-414) was developed using the aggregate binding dye ProteoStat®. The majority of potential aggregation inhibitor molecules tested in this system were ineffective, however a series of peptides designed using structural predictions from the AI system AlphaFold reduced TDP-43 CTF aggregation and lead to altered morphologies of the aggregates produced. Further work could be carried out testing these peptides, potentially in combination with liposomes or with other molecules such as PEG2000PE which also led to a reduction in aggregation related fluorescence in the ProteoStat® assay. Alternative assays could be used that utilise protein concentrations closer to endogenous levels such as a sedimentation assay. Furthermore, they could also be tested against the full-length protein which may aggregate differently.

An *S. cerevisiae* TDP-43 aggregation model was produced, demonstrating intracellular aggregation and growth reduction, which by the selection of a Δerg6 mutant strain could be “druggable”. Limited testing of aggregation inhibitors in this model was carried out, and so future work could involve testing the third-generation inhibitors, potentially modified for proteolytic stability and cell penetration, as well as other molecules such as PEG2000PE and curcumin and rosmarinic acid liposomes. The growth assays, both on plate and in liquid culture, and fluorescence microscopy to image the presence or absence of aggregates could be used to determine any therapeutic effect of inhibitor molecules.

10. References

18:0 peg2000 PE (2023) *Avanti Polar Lipids (en-US)*. Available at: https://avantilipids.com/product/880120 (Accessed: April 28, 2023).

Abd. Wahab, N.A., H. Lajis, N., Abas, F., Othman, I., and Naidu, R. (2020). Mechanism of anti-cancer activity of curcumin on androgen-dependent and androgen-independent prostate cancer. *Nutrients*, 12(3), p.679.

Abhyankar, M., Urekar, C. and Reddi, P. (2007). A Novel CpG-free Vertebrate Insulator Silences the Testis-specificSP-10Gene in Somatic Tissues. *Journal of Biological Chemistry*, 282(50), pp.36143-36154.

Adami, R. and Bottai, D. (2020). Curcumin and neurological diseases. *Nutritional Neuroscience*, 25(3), pp.441–461.

Afroz, T., Hock, E., Ernst, P., Foglieni, C., Jambeau, M., Gilhespy, L., Laferriere, F., Maniecka, Z., Plückthun, A., Mittl, P., Paganetti, P., Allain, F. and Polymenidou, M. (2017). Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. *Nature Communications*, 8(45).

Aggidis, A.G.A., (2019) *The development of peptide-based inhibitors for Tau aggregation as a potential therapeutic for Alzheimer's disease.* PhD. Lancaster University.

Altman, T., Ionescu, A., Ibraheem, A., Priesmann, D., Gradus-Pery, T., Farberov, L., Alexandra, G., Shelestovich, N., Dafinca, R., Shomron, N., Rage, F., Talbot, K., Ward, M., Dori, A., Krüger, M. and Perlson, E., (2021). Axonal TDP-43 condensates drive neuromuscular junction disruption through inhibition of local synthesis of nuclear encoded mitochondrial proteins. Nature Communications, 12(1).

Anand, U., Jash, C. and Mukherjee, S. (2011). Protein unfolding and subsequent refolding: a spectroscopic investigation. *Physical Chemistry Chemical Physics*, 13(45), p.20418.

Armakola, M., Hart, M. and Gitler, A. (2011). TDP-43 toxicity in yeast. *Methods*, 53(3), pp.238-245.

Arseni, D., Hasegawa, M., Murzin, A., Kametani, F., Arai, M., Yoshida, M. and Ryskeldi-Falcon, B. (2021). Structure of pathological TDP-43 filaments from ALS with FTLD. *Nature*, 601(7891), pp.139-143.

Austen, B.M., Paleologou, K.E., Ali, S.A., Qureshi, M.M., Allsop, D., and El-Agnaf, O.M. (2008). Designing peptide inhibitors for oligomerization and toxicity of alzheimer's β-amyloid peptide. *Biochemistry*, 47(7), pp.1984–1992.

Ayala, Y., De Conti, L., Avendaño-Vázquez, S., Dhir, A., Romano, M., D'Ambrogio, A., Tollervey, J., Ule, J., Baralle, M., Buratti, E. and Baralle, F. (2010). TDP-43 regulates its mRNA levels through a negative feedback loop. *The EMBO Journal*, 30(2), pp.277-288.

Bach, S., Talarek, N., Andrieu, T., Vierfond, J.-M., Mettey, Y., Galons, H., Dormont, D., Meijer, L., Cullin, C., and Blondel, M. (2003). Isolation of drugs active against mammalian prions using a yeast-based screening assay. *Nature Biotechnology*, 21(9), pp.1075–1081.

Bahia, V. S., Takada, L. T., & Deramecourt, V. (2013). Neuropathology of frontotemporal lobar degeneration: a review. *Dementia & neuropsychologia*, *7*(1), 19–26.

Baker, M., Mackenzie, I., Pickering-Brown, S., Gass, J., Rademakers, R., Lindholm, C., Snowden, J., Adamson, J., Sadovnick, A., Rollinson, S., Cannon, A., Dwosh, E., Neary, D., Melquist, S., Richardson, A., Dickson, D., Berger, Z., Eriksen, J., Robinson, T., Zehr, C., Dickey, C., Crook, R., McGowan, E., Mann, D., Boeve, B., Feldman, H.

and Hutton, M. (2006). Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17. *Nature*, 442(7105), pp.916-919.

Baralle, M., Buratti, E. and Baralle, F. (2013). The role of TDP-43 in the pathogenesis of ALS and FTLD. *Biochemical Society Transactions*, 41(6), pp.1536-1540.

Becker, L.A., Huang, B., Bieri, G., Ma, R., Knowles, D.A., Jafar-Nejad, P., Messing, J., Kim, H.J., Soriano, A., Auburger, G., Pulst, S.M., Taylor, J.P., Rigo, F., and Gitler, A.D. (2017). Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice. *Nature*, 544(7650), pp.367–371.

Berning, B. and Walker, A. (2019). The Pathobiology of TDP-43 C-Terminal Fragments in ALS and FTLD. *Frontiers in Neuroscience*, 13(335).

Boxer, A.L., Gold, M., Huey, E., Gao, F., Burton, E.A., Chow, T., Kao, A., Leavitt, B.R., Lamb, B., Grether, M., Knopman, D., Cairns, N.J., Mackenzie, I.R., Mitic, L., Roberson, E.D., Van Kammen, D., Cantillon, M., Zahs, K., Salloway, S., Morris, J., Tong, G., Feldman, H., Fillit, H., Dickinson, S., Khachaturian, Z., Sutherland, M., Farese, R., Miller, B.L., and Cummings, J. (2012). Frontotemporal degeneration, the next therapeutic frontier: Molecules and animal models for frontotemporal degeneration drug development. *Alzheimer's & Dementia*, 9(2), pp.176–188.

Buratti, E. and Baralle, F.E. (2001). Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. *Journal of Biological Chemistry*, 276(39), pp.36337–36343.

Buratti, E. and Baralle, F. (2010). The multiple roles of TDP-43 in pre-mRNA processing and gene expression regulation. *RNA Biology*, 7(4), pp.420-429.

Buratti, E., De Conti, L., Stuani, C., Romano, M., Baralle, M. and Baralle, F. (2010). Nuclear factor TDP-43 can affect selected microRNA levels. *FEBS Journal*, 277(10), pp.2268-2281.

Carlomagno, Y., Zhang, Y., Davis, M., Lin, W.-L., Cook, C., Dunmore, J., Tay, W., Menkosky, K., Cao, X., Petrucelli, L., and DeTure, M. (2014). Casein kinase II induced polymerization of soluble TDP-43 into filaments is inhibited by heat shock proteins. *PLoS ONE*, 9(3).

Cascella, R., Banchelli, M., Abolghasem Ghadami, S., Ami, D., Gagliani, M. C., Bigi, A., Staderini, T., Tampellini, D., Cortese, K., Cecchi, C., Natalello, A., Adibi, H., Matteini, P., & Chiti, F. (2022). An in situ and in vitro investigation of cytoplasmic TDP-43 inclusions reveals the absence of a clear amyloid signature. *Annals of Medicine*, *55*(1), 72–88.

Cascella, R., Bigi, A., Riffert, D., Gagliani, M., Ermini, E., Moretti, M., Cortese, K., Cecchi, C. and Chiti, F. (2022). A quantitative biology approach correlates neuronal toxicity with the largest inclusions of TDP-43. *Science Advances*, 8(30).

Chen, H., Topp, S., Hui, H., Zacco, E., Katarya, M., McLoughlin, C., King, A., Smith, B., Troakes, C., Pastore, A. and Shaw, C. (2019) RRM adjacent TARDBP mutations disrupt RNA binding and enhance TDP-43 proteinopathy. *Brain*, 142(12), pp.3753-3770.

Chen, Y., Wang, H., Ying, Z. and Gao, Q. (2020). Ibudilast enhances the clearance of SOD1 and TDP-43 aggregates through TFEB-mediated autophagy and lysosomal biogenesis: The new molecular mechanism of ibudilast and its implication for neuroprotective therapy. *Biochemical and Biophysical Research Communications*, 526(1), pp.231-238.

Chen, Z., Chen, J., Keshamouni, V.G., and Kanapathipillai, M. (2017). Polyarginine and its analogues inhibit p53 mutant aggregation and cancer cell proliferation in vitro. *Biochemical and Biophysical Research Communications*, 489(2), pp.130–134.

Clavaguera, F., Bolmont, T., Crowther, R., Abramowski, D., Frank, S., Probst, A., Fraser, G., Stalder, A., Beibel, M., Staufenbiel, M., Jucker, M., Goedert, M. and Tolnay,

M. (2009). Transmission and spreading of tauopathy in transgenic mouse brain. *Nature Cell Biology*, 11(7), pp.909-913.

Cleary, J., Pattamatta, A. and Ranum, L. (2018). Repeat-associated non-ATG (RAN) translation. *Journal of Biological Chemistry*, 293(42), pp.16127-16141.

Cohen, T., Hwang, A., Restrepo, C., Yuan, C., Trojanowski, J. and Lee, V. (2015). An acetylation switch controls TDP-43 function and aggregation propensity. *Nature Communications*, 6(1).

Conchillo-Solé, O., de Groot, N., Avilés, F., Vendrell, J., Daura, X., Ventura, S. (2007). AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides, *BMC Bioinformatics*, 27(8)65.

Contreras-Gómez, A., Sánchez-Mirón, A., García-Camacho, F., Molina-Grima, E., and Chisti, Y. (2013). Protein production using the baculovirus-insect cell expression system. *Biotechnology Progress*, 30(1), pp.1–18.

Corcia, P. and Meininger, V. (2019). Grey Matter 150th anniversary of Charcot's description of amyotrophic lateral sclerosis. *Brain*, 142(10), pp.3306-3313.

Costa-Mattioli, M. and Walter, P. (2020). The Integrated Stress Response: From Mechanism to disease. *Science*, 368(6489).

Dammer, E., Fallini, C., Gozal, Y., Duong, D., Rossoll, W., Xu, P., Lah, J., Levey, A., Peng, J., Bassell, G. and Seyfried, N. (2012). Coaggregation of RNA-Binding Proteins in a Model of TDP-43 Proteinopathy with Selective RGG Motif Methylation and a Role for RRM1 Ubiquitination. *PLoS ONE*, 7(6), p.e38658.

Das, A. and Mukhopadhyay, C. (2009). Urea-Mediated Protein Denaturation: A Consensus View. *The Journal of Physical Chemistry B*, 113(38), pp.12816-12824.

Das, A.S. and Zou, W.-Q. (2016). Prions: Beyond a single protein. *Clinical Microbiology Reviews*, 29(3), pp.633–658.

DeJesus-Hernandez, M., Mackenzie, I., Boeve, B., Boxer, A., Baker, M., Rutherford, N., Nicholson, A., Finch, N., Flynn, H., Adamson, J., Kouri, N., Wojtas, A., Sengdy, P., Hsiung, G., Karydas, A., Seeley, W., Josephs, K., Coppola, G., Geschwind, D., Wszolek, Z., Feldman, H., Knopman, D., Petersen, R., Miller, B., Dickson, D., Boylan, K., Graff-Radford, N. and Rademakers, R. (2011). Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron*, 72(2), pp.245-256.

Deng, H., Chen, W., Hong, S., Boycott, K., Gorrie, G., Siddique, N., Yang, Y., Fecto, F., Shi, Y., Zhai, H., Jiang, H., Hirano, M., Rampersaud, E., Jansen, G., Donkervoort, S., Bigio, E., Brooks, B., Ajroud, K., Sufit, R., Haines, J., Mugnaini, E., Pericak-Vance, M. and Siddique, T. (2011). Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature*, 477(7363), pp.211-215.

Ding, X., Xiang, Z., Qin, C., Chen, Y., Tian, H., Meng, L., Xia, D., Liu, H., Song, J., Fu, J., Ma, M. and Wang, X., (2021). Spreading of TDP-43 pathology via pyramidal tract induces ALS-like phenotypes in TDP-43 transgenic mice. *Acta Neuropathologica Communications*, 9(1).

Doke, A. A., & Jha, S. K. (2022). Effect of *in vitro* solvation conditions on inter- and intramolecular assembly of full-length TDP-43. *The Journal of Physical Chemistry B, 126*(26), pp.4799-4813.

Donde, A., Sun, M., Ling, J., Braunstein, K., Pang, B., Wen, X., Cheng, X., Chen, L. and Wong, P. (2019). Splicing repression is a major function of TDP-43 in motor neurons. *Acta Neuropathologica*, 138(5), pp.813-826.

Dong, H., Xu, L., Wu, L., Wang, X., Duan, W., Li, H., and Li, C. (2014). Curcumin abolishes mutant TDP-43 induced excitability in a motoneuron-like cellular model of ALS. *Neuroscience*, 272, pp.141–153.

Doyle, S., Genest, O. and Wickner, S. (2013). Protein rescue from aggregates by powerful molecular chaperone machines. *Nature Reviews Molecular Cell Biology*, 14(10), pp.617-629.

Dreumont, N., Hardy, S., Behm-Ansmant, I., Kister, L., Branlant, C., Stévenin, J. and Bourgeois, C., (2009). Antagonistic factors control the unproductive splicing of SC35 terminal intron. *Nucleic Acids Research*, 38(4), pp.1353-1366.

Duan, L., Zaepfel, B., Aksenova, V., Dasso, M., Rothstein, J., Kalab, P. and Hayes, L. (2022). Nuclear RNA binding regulates TDP-43 nuclear localization and passive nuclear export. *Cell Reports*, 40(3), p.111106.

Duan, W., Guo, Y., Xiao, J., Chen, X., Li, Z., Han, H., and Li, C. (2013). Neuroprotection by monocarbonyl dimethoxycurcumin C: Ameliorating the toxicity of mutant TDP-43 via HO-1. *Molecular Neurobiology*, 49(1), pp.368–379.

Elden, A.C., Kim, H.-J., Hart, M.P., Chen-Plotkin, A.S., Johnson, B.S., Fang, X., Armakola, M., Geser, F., Greene, R., Lu, M.M., Padmanabhan, A., Clay-Falcone, D., McCluskey, L., Elman, L., Juhr, D., Gruber, P.J., Rüb, U., Auburger, G., Trojanowski, J.Q., Lee, V.M.-Y., Van Deerlin, V.M., Bonini, N.M., and Gitler, A.D. (2010). Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature*, 466(7310), pp.1069–1075.

Emter, R., Heese-Peck, A. and Kralli, A., (2002). ERG6 and PDR5 regulate small lipophilic drug accumulation in yeast cells via distinct mechanisms. *FEBS Letters*, 521(1-3), pp.57-61.

Erickson, H.P. (2009). Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biological Procedures Online*, 11(1), pp.32–51.

Esposto, J.C. and Martic, S. (2021). “Phosphorylated tar DNA-binding protein-43: Aggregation and antibody-based inhibition,” *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1867(12), p. 166234. Available at: https://doi.org/10.1016/j.bbadis.2021.166234.

Fadouloglou, V.E., Kokkinidis, M., and Glykos, N.M. (2008). Determination of protein oligomerization state: Two approaches based on glutaraldehyde crosslinking. *Analytical Biochemistry*, 373(2), pp.404–406.

Fang, X., Yang, T., Wang, L., Yu, J., Wei, X., Zhou, Y., Wang, C., and Liang, W. (2016). Nano-cage-mediated refolding of insulin by Peg-Pe Micelle. *Biomaterials*, 77, pp.139–148.

Fang, X., Yousaf, M., Huang, Q., Yang, Y., and Wang, C. (2018). Dual effect of PEG-pe micelle over the oligomerization and fibrillation of human islet amyloid polypeptide. *Scientific Reports*, 8(1).

Fang, Y.-S., Tsai, K.-J., Chang, Y.-J., Kao, P., Woods, R., Kuo, P.-H., Wu, C.-C., Liao, J.-Y., Chou, S.-C., Lin, V., Jin, L.-W., Yuan, H.S., Cheng, I.H., Tu, P.-H., and Chen, Y.-R. (2014). Full-length TDP-43 forms toxic amyloid oligomers that are present in frontotemporal lobar dementia-TDP patients. *Nature Communications*, 5(1), p.4824.

Fay, M., Anderson, P. and Ivanov, P. (2017). ALS/FTD-Associated C9ORF72 Repeat RNA Promotes Phase Transitions In Vitro and in Cells. *Cell Reports*, 21(12), pp.3573-3584.

Fecto, F., Yan, J., Vemula, S., Liu, E., Chen, W., Zheng, J., Shi, Y., Siddique, N., Arrat, H., Donkervoort, S., Ajroud-Driss, S., Sufit, R., Heller, S. and Deng, H. (2011). SQSTM1 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis. *Archives of Neurology*, 68(11), pp.1440-1446.

Finkbeiner, S. (2019). The autophagy lysosomal pathway and neurodegeneration. *Cold Spring Harbor Perspectives in Biology*, 12(3).

François-Moutal, L., Perez-Miller, S., Scott, D., Miranda, V., Mollasalehi, N. and Khanna, M. (2019). Structural Insights Into TDP-43 and Effects of Post-translational Modifications. *Frontiers in Molecular Neuroscience*, 12, p.301.

Furukawa, Y., Kaneko, K., Watanabe, S., Yamanaka, K., and Nukina, N. (2011). A seeding reaction recapitulates intracellular formation of Sarkosyl-insoluble transactivation response element (TAR) DNA-binding protein-43 inclusions. *Journal of Biological Chemistry*, 286(21), pp.18664–18672.

Gaber, R., Copple, D., Kennedy, B., Vidal, M. and Bard, M., (1989). The yeast gene ERG6 is required for normal membrane function but is not essential for biosynthesis of the cell-cycle-sparking sterol. *Molecular and Cellular Biology*, 9(8), pp.3447-3456.

Gan, L., Cookson, M., Petrucelli, L. and La Spada, A. (2018). Converging pathways in neurodegeneration, from genetics to mechanisms. *Nature Neuroscience*, 21(10), pp.1300-1309.

Gao, J., Wang, L., Gao, C., Arakawa, H., Perry, G. and Wang, X. (2020). TDP-43 inhibitory peptide alleviates neurodegeneration and memory loss in an APP transgenic mouse model for Alzheimer's disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1866(1), p.165580.

Gao, J., Wang, L., Yan, T., Perry, G. and Wang, X. (2019). TDP-43 proteinopathy and mitochondrial abnormalities in neurodegeneration. *Molecular and Cellular Neuroscience*, 100, p.103396.

Garnier, C., Devred, F., Byrne, D., Puppo, R., Roman, A.Y., Malesinski, S., Golovin, A.V., Lebrun, R., Ninkina, N.N. and Tsvetkov, P.O. (2017). Zinc binding to RNA

recognition motif of TDP-43 induces the formation of amyloid-like aggregates. *Scientific Reports*. 7(1).

Gautam, M., Xie, E., Kocak, N. and Ozdinler, P. (2019). Mitoautophagy: A Unique Self-Destructive Path Mitochondria of Upper Motor Neurons With TDP-43 Pathology Take, Very Early in ALS. *Frontiers in Cellular Neuroscience*, 13(489).

Goh, C., Lee, I., Sundaram, J., George, S., Yusoff, P., Brush, M., Sze, N. and Shenolikar, S. (2017). Chronic oxidative stress promotes GADD34-mediated phosphorylation of the TAR DNA-binding protein TDP-43, a modification linked to neurodegeneration. *Journal of Biological Chemistry*, 293(1), pp.163-176.

Golovin, A.V., Devred, F., Yatoui, D., Roman, A.Yu., Zalevsky, A.O., Puppo, R., Lebrun, R., Guerlesquin, F., and Tsvetkov, P.O. (2020). Zinc binds to RRM2 peptide of TDP-43. *International Journal of Molecular Sciences*, 21(23), p.9080.

Gorbunoff, M. and Timasheff, S. (1984). The interaction of proteins with hydroxyapatite. *Analytical Biochemistry*, 136(2), pp.440-445.

Griffith, I.P. (1972). The effect of cross-links on the mobility of proteins in dodecyl sulphate–polyacrylamide gels. *Biochemical Journal*, 126(3), pp.553–560.

Gruijs da Silva, L., Simonetti, F., Hutten, S., Riemenschneider, H., Sternburg, E., Pietrek, L., Gebel, J., Dötsch, V., Edbauer, D., Hummer, G., Stelzl, L. and Dormann, D., (2022). Disease-linked TDP-43 hyperphosphorylation suppresses TDP-43 condensation and aggregation. The EMBO Journal, e108443.

Guenther, E., Cao, Q., Trinh, H., Lu, J., Sawaya, M., Cascio, D., Boyer, D., Rodriguez, J., Hughes, M. and Eisenberg, D. (2018). Atomic structures of TDP-43 LCD segments and insights into reversible or pathogenic aggregation. *Nature Structural & Molecular Biology*, 25(6), pp.463-471.

Hans, F., Glasebach, H. and Kahle, P. (2020). Multiple distinct pathways lead to hyperubiquitylated insoluble TDP-43 protein independent of its translocation into stress granules. *Journal of Biological Chemistry*, 295(3), pp.673-689.

Hanson, PE. and Gellman SH .(1998) .Mechanistic comparison of artificial-chaperone-assisted and unassisted refolding of urea-denatured carbonic anhydrase B. *Folding and Design* 3(6): pp.457–468.

Hasegawa-Ogawa, M. and Okano, H., (2021). Characterization of the upstream and intron promoters of the gene encoding TAR DNA-binding protein. *Scientific Reports*, 11(1).

Hergesheimer, R., Chami, A., de Assis, D., Vourc'h, P., Andres, C., Corcia, P., Lanznaster, D. and Blasco, H. (2019). The debated toxic role of aggregated TDP-43 in amyotrophic lateral sclerosis: a resolution in sight?. *Brain*, 142(5), pp.1176-1194.

Herrera, M.G., Amundarain, M.J. and Santos, J. (2023). "Biophysical evaluation of the oligomerization and conformational properties of the N-terminal domain of TDP-43," *Archives of Biochemistry and Biophysics*, p. 109533. Available at: https://doi.org/10.1016/j.abb.2023.109533.

Hershko, A., Heller, H., Elias, S., and Ciechanover, A. (1983). Components of ubiquitin-protein ligase system. resolution, affinity purification, and role in protein breakdown. *Journal of Biological Chemistry*, 258(13), pp.8206–8214.

Hilbrig, F. and Freitag, R., (2011). Isolation and purification of recombinant proteins, antibodies and plasmid DNA with hydroxyapatite chromatography. *Biotechnology Journal*, 7(1), pp.90-102.

Hofmann, S., Kedersha, N., Anderson, P., and Ivanov, P. (2021). Molecular mechanisms of stress granule assembly and disassembly. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1868(1), p.118876.

Huang, C., Xia, P. and Zhou, H. (2010). Sustained Expression of TDP-43 and FUS in Motor Neurons in Rodent's Lifetime. *International Journal of Biological Sciences*, 6(4), pp.396-406.

Hudson, S.A., Ecroyd, H., Kee, T.W., and Carver, J.A. (2009). The thioflavin T fluorescence assay for amyloid fibril detection can be biased by the presence of exogenous compounds. *FEBS Journal*, 276(20), pp.5960–5972.

Humphreys, K., (2013). *Expression And Purification Of TDP43 And FUS, Proteins Implicated In Neurodegeneration, And Characterisation Of Novel Anti-TDP43 Monoclonal Antibodies*. PhD. Lancaster University.

Iguchi, Y., Eid, L., Parent, M., Soucy, G., Bareil, C., Riku, Y., Kawai, K., Takagi, S., Yoshida, M., Katsuno, M., Sobue, G. and Julien, J. (2016). Exosome secretion is a key pathway for clearance of pathological TDP-43. *Brain*, 139(12), pp.3187-3201.

Iyer, S., Subramanian, V. and Acharya, K. (2018). C9orf72, a protein associated with amyotrophic lateral sclerosis (ALS) is a guanine nucleotide exchange factor. *PeerJ*, 6, p.e5815.

Jaiswal, M. (2018). Riluzole and edaravone: A tale of two amyotrophic lateral sclerosis drugs. *Medicinal Research Reviews*, 39(2), pp.733-748.

Jiang, L.-L., Xue, W., Hong, J.-Y., Zhang, J.-T., Li, M.-J., Yu, S.-N., He, J.-H., and Hu, H.-Y. (2017). The N-terminal dimerization is required for TDP-43 splicing activity. *Scientific Reports*, 7(1).

Jiang, L., Zhao, J., Yin, X., He, W., Yang, H., Che, M. and Hu, H. (2016). Two mutations G335D and Q343R within the amyloidogenic core region of TDP-43 influence its aggregation and inclusion formation. *Scientific Reports*, 6(1).

Johnson, B.S., McCaffery, J.M., Lindquist, S., and Gitler, A.D. (2008). A yeast TDP-43 proteinopathy model: Exploring the molecular determinants of TDP-43 aggregation and

cellular toxicity. *Proceedings of the National Academy of Sciences*, 105(17), pp.6439–6444.

Johnson, B., Snead, D., Lee, J., McCaffery, J., Shorter, J. and Gitler, A. (2009). TDP-43 Is Intrinsically Aggregation-prone, and Amyotrophic Lateral Sclerosis-linked Mutations Accelerate Aggregation and Increase Toxicity. *Journal of Biological Chemistry*, 284(30), pp.20329-20339.

Johnson, J., Pioro, E., Boehringer, A., Chia, R., Feit, H., Renton, A., Pliner, H., Abramzon, Y., Marangi, G., Winborn, B., Gibbs, J., Nalls, M., Morgan, S., Shoai, M., Hardy, J., Pittman, A., Orrell, R., Malaspina, A., Sidle, K., Fratta, P., Harms, M., Baloh, R., Pestronk, A., Weihl, C., Rogaeva, E., Zinman, L., Drory, V., Borghero, G., Mora, G., Calvo, A., Rothstein, J., Drepper, C., Sendtner, M., Singleton, A., Taylor, J., Cookson, M., Restagno, G., Sabatelli, M., Bowser, R., Chiò, A. and Traynor, B. (2014). Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nature Neuroscience*, 17(5), pp.664-666.

Jucker, M., Walker, L. (2013). Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature*, 501, pp.45–51.

Kamagata, K., Kanbayashi, S., Koda, S., Kadotani, A., Ubukata, O., and Tashima, T. (2023). Suppression of TDP-43 aggregation by artificial peptide binder targeting to its low complexity domain. *Biochemical and Biophysical Research Communications*, 662, pp.119–125.

Kametani, F., Obi, T., Shishido, T., Akatsu, H., Murayama, S., Saito, Y., Yoshida, M. and Hasegawa, M. (2016). Mass spectrometric analysis of accumulated TDP-43 in amyotrophic lateral sclerosis brains. *Scientific Reports*, 6(1), p.23281.

Kaspersen, J.D., Søndergaard, A., Madsen, D.J., Otzen, D.E., and Pedersen, J.S. (2017). Refolding of SDS-unfolded proteins by nonionic surfactants. *Biophysical Journal*, 112(8), pp.1609–1620.

Khosravi, B., LaClair, K., Riemenschneider, H., Zhou, Q., Frottin, F., Mareljic, N., Czuppa, M., Farny, D., Hartmann, H., Michaelsen, M., Arzberger, T., Hartl, F., Hipp, M. and Edbauer, D., (2020). Cell-to-cell transmission of C9orf72 poly-(Gly-Ala) triggers key features of ALS / FTD. *The EMBO Journal*, 39(8).

Kiernan, M., Vucic, S., Cheah, B., Turner, M., Eisen, A., Hardiman, O., Burrell, J. and Zoing, M. (2011). Amyotrophic Lateral Sclerosis. *Lancet*, 377, pp.942-955.

Kim, H., Kim, N., Wang, Y., Scarborough, E., Moore, J., Diaz, Z., MacLea, K., Freibaum, B., Li, S., Molliex, A., Kanagaraj, A., Carter, R., Boylan, K., Wojtas, A., Rademakers, R., Pinkus, J., Greenberg, S., Trojanowski, J., Traynor, B., Smith, B., Topp, S., Gkazi, A., Miller, J., Shaw, C., Kottlors, M., Kirschner, J., Pestronk, A., Li, Y., Ford, A., Gitler, A., Benatar, M., King, O., Kimonis, V., Ross, E., Weihl, C., Shorter, J. and Taylor, J. (2013). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature*, 495(7442), pp.467-473.

King, I., Yartseva, V., Salas, D., Kumar, A., Heidersbach, A., Ando, D., Stallings, N., Elliott, J., Srivastava, D. and Ivey, K. (2014). The RNA-binding Protein TDP-43 Selectively Disrupts MicroRNA-1/206 Incorporation into the RNA-induced Silencing Complex. *Journal of Biological Chemistry*, 289(20), pp.14263-14271.

Kirby, J., Al Sultan, A., Waller, R. and Heath, P. (2016). The genetics of amyotrophic lateral sclerosis: current insights. *Degenerative Neurological and Neuromuscular Disease*, 6, pp.49-64.

Kitamura, A., Nakayama, Y., Shibasaki, A., Taki, A., Yuno, S., Takeda, K., Yahara, M., Tanabe, N. and Kinjo, M. (2016). Interaction of RNA with a C-terminal fragment of the amyotrophic lateral sclerosis-associated TDP43 reduces cytotoxicity. *Scientific Reports*, 6(1).

Koike, Y., Pickles, S., Estades Ayuso, V., Jansen-West, K., Qi, Y.A., Li, Z., Daughrity, L.M., Yue, M., Zhang, Y.-J., Cook, C.N., Dickson, D.W., Ward, M., Petrucelli, L., and

Prudencio, M. (2023). TDP-43 and other hnRNPs regulate cryptic exon inclusion of a key ALS/FTD risk gene, UNC13A. *PLOS Biology*, 21(3).

Koppers, M., Blokhuis, A., Westeneng, H., Terpstra, M., Zundel, C., Vieira de Sá, R., Schellevis, R., Waite, A., Blake, D., Veldink, J., van den Berg, L. and Pasterkamp, R. (2015). C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. *Annals of Neurology*, 78(3), pp.426-438.

Kozlowski, L.P. (2016). Proteome-PI: Proteome Isoelectric Point Database. *Nucleic Acids Research*, 45(D1).

Kryndushkin, D. and Shewmaker, F., (2011). Modeling ALS and FTLD proteinopathies in yeast: An efficient approach for studying protein aggregation and toxicity. *Prion*, 5(4), pp.250-257.

Kumar, S.T., Nazarov, S., Porta, S., Maharjan, N., Cendrowska, U., Kabani, M., Finamore, F., Xu, Y., Lee, V.M.-Y., and Lashuel, H.A. (2023). Seeding the aggregation of TDP-43 requires post-fibrillization proteolytic cleavage. *Nature Neuroscience*, https://doi.org/10.1038/s41593-023-01341-4

Kumar, V., Wahiduzzaman, Prakash, A., Tomar, A., Srivastava, A., Kundu, B., Lynn, A. and Imtaiyaz Hassan, M. (2019). Exploring the aggregation-prone regions from structural domains of human TDP-43. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1867(3), pp.286-296.

Kuo, P., Doudeva, L., Wang, Y., Shen, C. and Yuan, H. (2009). Structural insights into TDP-43 in nucleic-acid binding and domain interactions. *Nucleic Acids Research*, 37(6), pp.1799-1808.

Lackie, R., Maciejewski, A., Ostapchenko, V., Marques-Lopes, J., Choy, W., Duennwald, M., Prado, V. and Prado, M. (2017). The Hsp70/Hsp90 Chaperone Machinery in Neurodegenerative Diseases. *Frontiers in Neuroscience*, 11, Article 254

Lambert, M., Barlow, A., Chromy, B., Edwards, C., Freed, R., Liosatos, M., Morgan, T., Rozovsky, I., Trommer, B., Viola, K., Wals, P., Zhang, C., Finch, C., Krafft, G. and Klein, W. (1998). Diffusible, nonfibrillar ligands derived from A 1-42 are potent central nervous system neurotoxins. *Proceedings of the National Academy of Sciences*, 95(11), pp.6448-6453.

Laos, V., Do, T., Bishop, D., Jin, Y., Marsh, N., Quon, B., Fetters, M., Cantrell, K., Buratto, S. and Bowers, M. (2019). Characterizing TDP-43307–319 Oligomeric Assembly: Mechanistic and Structural Implications Involved in the Etiology of Amyotrophic Lateral Sclerosis. *ACS Chemical Neuroscience*, 10(9), pp.4112-4123.

Leibiger, C., Deisel, J., Aufschnaiter, A., Ambros, S., Tereshchenko, M., Verheijen, B., Büttner, S. and Braun, R. (2018). TDP-43 controls lysosomal pathways thereby determining its own clearance and cytotoxicity. *Human Molecular Genetics*, 27(9), pp.1593-1607.

Li, H., Chiang, W., Chou, P., Wang, W. and Huang, J. (2018). TAR DNA-binding protein 43 (TDP-43) liquid–liquid phase separation is mediated by just a few aromatic residues. *Journal of Biological Chemistry*, 293(16), pp.6090-6098.

Ling, J., Pletnikova, O., Troncoso, J. and Wong, P. (2015). TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science*, 349(6248), pp.650-655.

Liu, G., Coyne, A., Pei, F., Vaughan, S., Chaung, M., Zarnescu, D. and Buchan, J. (2017). Endocytosis regulates TDP-43 toxicity and turnover. *Nature Communications*, 8(1).

Liu, R., Yang, G., Nonaka, T., Arai, T., Jia, W. and Cynader, M., (2013). Reducing TDP-43 aggregation does not prevent its cytotoxicity. Acta Neuropathologica Communications, 1(1).

Liu-Yesucevitz, L., Bilgutay, A., Zhang, Y., Vanderwyde, T., Citro, A., Mehta, T., Zaarur, N., McKee, A., Bowser, R., Sherman, M., Petrucelli, L. and Wolozin, B. (2010). Tar DNA Binding Protein-43 (TDP-43) Associates with Stress Granules: Analysis of Cultured Cells and Pathological Brain Tissue. *PLoS ONE*, 5(10), p.e13250.

Lobstein, J., Emrich, C., Jeans, C., Faulkner, M., Riggs, P. and Berkmen, M. (2012). SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microbial Cell Factories*, 11(1).

Luan, W., Wright, A.L., Brown-Wright, H., Le, S., San Gil, R., Madrid San Martin, L., Ling, K., Jafar-Nejad, P., Rigo, F., and Walker, A.K. (2023). Early activation of cellular stress and death pathways caused by cytoplasmic TDP-43 in the RNLS8 mouse model of ALS and FTD. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-023-02036-9

Lukavsky, P., Daujotyte, D., Tollervey, J., Ule, J., Stuani, C., Buratti, E., Baralle, F., Damberger, F. and Allain, F. (2013). Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nature Structural & Molecular Biology*, 20(12), pp.1443-1449.

Luo, C., Zou, L., Sun, H., Peng, J., Gao, C., Bao, L., Ji, R., Jin, Y., and Sun, S. (2020). A review of the anti-inflammatory effects of rosmarinic acid on inflammatory diseases. *Frontiers in Pharmacology*, 11.

Mackenzie, I., Arzberger, T., Kremmer, E., Troost, D., Lorenzl, S., Mori, K., Weng, S., Haass, C., Kretzschmar, H., Edbauer, D. and Neumann, M. (2013). Dipeptide repeat protein pathology in C9ORF72 mutation cases: clinico-pathological correlations. *Acta Neuropathologica*, 126(6), pp.859-879.

Maruyama, H., Morino, H., Ito, H., Izumi, Y., Kato, H., Watanabe, Y., Kinoshita, Y., Kamada, M., Nodera, H., Suzuki, H., Komure, O., Matsuura, S., Kobatake, K., Morimoto, N., Abe, K., Suzuki, N., Aoki, M., Kawata, A., Hirai, T., Kato, T., Ogasawara, K., Hirano, A., Takumi, T., Kusaka, H., Hagiwara, K., Kaji, R. and Kawakami, H. (2010).

Mutations of optineurin in amyotrophic lateral sclerosis. *Nature*, 465(7295), pp.223-226.

Mathis, S., Goizet, C., Soulages, A., Vallat, J.-M., and Masson, G.L. (2019). Genetics of Amyotrophic Lateral sclerosis: A Review. *Journal of the Neurological Sciences*, 399, pp.217–226.

Mayes, J., Tinker-Mill, C., Kolosov, O., Zhang, H., Tabner, B.J., and Allsop, D. (2014). B-amyloid fibrils in alzheimer disease are not inert when bound to copper ions but can degrade hydrogen peroxide and generate reactive oxygen species. *Journal of Biological Chemistry*, 289(17), pp.12052–12062.

Mejzini, R., Flynn, L., Pitout, I., Fletcher, S., Wilton, S. and Akkari, P. (2019). ALS Genetics, Mechanisms, and Therapeutics: Where Are We Now?. *Frontiers in Neuroscience*, 13(1310).

Meneses, A., Koga, S., O'Leary, J., Dickson, D.W., Bu, G., and Zhao, N. (2021). TDP-43 pathology in Alzheimer's disease. *Molecular Neurodegeneration*, 16(1).

Michaux, C., Pomroy, N.C., and Privé, G.G. (2008). Refolding SDS-denatured proteins by the addition of amphipathic cosolvents. *Journal of Molecular Biology*, 375(5), pp.1477–1488.

Milstead, R. A., Link, C. D., Xu, Z., & Hoeffer, C. A. (2022). TDP-43 knockdown in mouse model of ALS leads to dsrna deposition, gliosis, and neurodegeneration in the spinal cord. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhac461

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). Colabfold - making protein folding accessible to all. *Nature Methods*, 19, pp.679–682.

Mitra, J., Guerrero, E., Hegde, P., Liachko, N., Wang, H., Vasquez, V., Gao, J., Pandey, A., Taylor, J., Kraemer, B., Wu, P., Boldogh, I., Garruto, R., Mitra, S., Rao, K. and

Hegde, M. (2019). Motor neuron disease-associated loss of nuclear TDP-43 is linked to DNA double-strand break repair defects. *Proceedings of the National Academy of Sciences*, 116(10), pp.4696-4705.

Mitra, J. and Hegde, M. (2019). A Commentary on TDP-43 and DNA Damage Response in Amyotrophic Lateral Sclerosis. *Journal of Experimental Neuroscience*, 13, pp.1-5.

Moiola, M., Memeo, M.G., and Quadrelli, P. (2019). Stapled peptides—a useful improvement for peptide-based drugs. *Molecules*, 24(20), p.3654.

Mompeán, M., Buratti, E., Guarnaccia, C., Brito, R., Chakrabartty, A., Baralle, F. and Laurents, D. (2014). "Structural characterization of the minimal segment of TDP-43 competent for aggregation". *Archives of Biochemistry and Biophysics*, 545, pp.53-62.

Mompeán, M., Ramírez de Mingo, D., Hervás, R., Fernández-Ramírez, M., Carrión-Vázquez, M. and Laurents, D. (2019). Molecular mechanism of the inhibition of TDP-43 amyloidogenesis by QBP1. *Archives of Biochemistry and Biophysics*, 675, p.108113.

Mompeán, M., Romano, V., Pantoja-Uceda, D., Stuani, C., Baralle, F., Buratti, E. and Laurents, D. (2016). The TDP-43 N-terminal domain structure at high resolution. *The FEBS Journal*, 283(7), pp.1242-1260.

Morera, A., Ahmed, N. and Schwartz, J. (2019). TDP-43 regulates transcription at protein-coding genes and Alu retrotransposons. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1862(10), p.194434.

Moretti, M., Marzi, I., Cantarutti, C., Vivoli Vega, M., Mandaliti, W., Mimmi, M., Bemporad, F., Corazza, A. and Chiti, F. (2022). Conversion of the Native N-Terminal Domain of TDP-43 into a Monomeric Alternative Fold with Lower Aggregation Propensity. *Molecules*, 27(13), p.4309.

Nadimidla, K., Ismail, T., and Kanapathipillai, M. (2017). Tau peptides and tau mutant protein aggregation inhibition by cationic polyethyleneimine and polyarginine. *Biopolymers*, 107(9).

Nakashima-Yasuda, H., Uryu, K., Robinson, J., Xie, S.X., Hurtig, H., Duda, J.E., Arnold, S.E., Siderowf, A., Grossman, M., Leverenz, J.B., Woltjer, R., Lopez, O.L., Hamilton, R., Tsuang, D.W., Galasko, D., Masliah, E., Kaye, J., Clark, C.M., Montine, T.J., Lee, V.M., and Trojanowski, J.Q. (2007). Co-morbidity of TDP-43 proteinopathy in lewy body related diseases. *Acta Neuropathologica*, 114(3), pp.221–229.

Narayanan, R., Panwar, A., Butler, T., Cutrupi, A., Kennerson, M., Vucic, S., Balasubramaniem, A., Mangelsdorf, M., and Wallace, R. (2023). Transgenic mice overexpressing mutant TDP-43 show aberrant splicing of autism associated gene *zmynd11* prior to onset of motor symptoms. *MicroPublication Biology*, p.10.17912/micropub.biology.000777.

Navarro, S. and Ventura, S. (2014). Fluorescent dye ProteoStat® to detect and discriminate intracellular amyloid-like aggregates in*escherichia coli*. *Biotechnology Journal*, 9(10), pp.1259–1266.

Nelson, P., Dickson, D., Trojanowski, J., Jack, C., Boyle, P., Arfanakis, K., Rademakers, R., Alafuzoff, I., Attems, J., Brayne, C., Coyle-Gilchrist, I., Chui, H., Fardo, D., Flanagan, M., Halliday, G., Hokkanen, S., Hunter, S., Jicha, G., Katsumata, Y., Kawas, C., Keene, C., Kovacs, G., Kukull, W., Levey, A., Makkinejad, N., Montine, T., Murayama, S., Murray, M., Nag, S., Rissman, R., Seeley, W., Sperling, R., White III, C., Yu, L. and Schneider, J. (2019). Limbic-predominant age-related TDP-43 encephalopathy (LATE): consensus working group report. *Brain*, 142(6), pp.1503-1527.

Neumann, M., Kwong, L., Lee, E., Kremmer, E., Flatley, A., Xu, Y., Forman, M., Troost, D., Kretzschmar, H., Trojanowski, J. and Lee, V. (2009). Phosphorylation of S409/410

of TDP-43 is a consistent feature in all sporadic and familial forms of TDP-43 proteinopathies. *Acta Neuropathologica*, 117(2), pp.137-149.

Neumann, M., Rademakers, R., Roeber, S., Baker, M., Kretzschmar, H. and Mackenzie, I. (2009). A new subtype of frontotemporal lobar degeneration with FUS pathology. *Brain*, 132(11), pp.2922-2931.

Neumann, M., Sampathu, D., Kwong, L., Truax, A., Micsenyi, M., Chou, T., Bruce, J., Schuck, T., Grossman, M., Clark, C., McCluskey, L., Miller, B., Masliah, E., Mackenzie, I., Feldman, H., Feiden, W., Kretzschmar, H., Trojanowski, J. and Lee, V. (2006). Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Science*, 314(5796), pp.130-133.

Ni, J., Ren, Y., Su, T., Zhou, J., Fu, C., Lu, Y., Li, D., Zhao, J., Li, Y., Zhang, Y., Fang, Y., Liu, N., Geng, Y. and Chen, Y., 2021. Loss of TDP-43 function underlies hippocampal and cortical synaptic deficits in TDP-43 proteinopathies. *Molecular Psychiatry*.

Nishimura, A., Župunski, V., Troakes, C., Kathe, C., Fratta, P., Howell, M., Gallo, J., Hortobágyi, T., Shaw, C. and Rogelj, B. (2010). Nuclear import impairment causes cytoplasmic trans-activation response DNA-binding protein accumulation and is associated with frontotemporal lobar degeneration. *Brain*, 133(6), pp.1763-1771.

Nolan, M., Talbot, K., and Ansorge, O. (2016). Pathogenesis of FUS-associated ALS and FTD: Insights from rodent models. *Acta Neuropathologica Communications*, 4(1).

Nonaka, T. and Hasegawa, M. (2012). Intracellular seeded aggregation of TDP-43. *Rinsho Shinkeigaku*, 52(11), pp.1056-1058.

Nonaka, T. and Hasegawa, M. (2020). Prion-like properties of assembled TDP-43. *Current Opinion in Neurobiology*, 61, pp.23-28.

Nonaka, T., Masuda-Suzukake, M., Arai, T., Hasegawa, Y., Akatsu, H., Obi, T., Yoshida, M., Murayama, S., Mann, D., Akiyama, H. and Hasegawa, M., (2013). Prion-like Properties of Pathological TDP-43 Aggregates from Diseased Brains. *Cell Reports*, 4(1), pp.124-134.

Novak, V., Rogelj, B., and Župunski, V. (2021). Therapeutic potential of polyphenols in amyotrophic lateral sclerosis and frontotemporal dementia. *Antioxidants*, 10(8), p.1328.

Ogrydziak, D.M. (1993). Yeast extracellular proteases. *Critical Reviews in Biotechnology*, 13(1), pp.1–55.

Oshinbolu, S., Shah, R., Finka, G., Molloy, M., Uden, M., and Bracewell, D.G. (2018). Evaluation of fluorescent dyes to measure protein aggregation within mammalian cell culture supernatants. *Journal of Chemical Technology & Biotechnology*, 93(3), pp.909–917.

Ossipow, V., Laemmlii, U.K., and Schibler, U. (1993). A simple method to renature DNA-binding proteins separated by SDS-polyacrylamide gel electrophoresis. *Nucleic Acids Research*, 21(25), pp.6040–6041.

Ou, S., Wu, F., Harrich, D., García-Martínez, L. and Gaynor, R. (1995). Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *Journal of virology*, 69(6), pp.3584-3596.

Parker, S., Meyerowitz, J., James, J., Liddell, J., Nonaka, T., Hasegawa, M., Kanninen, K., Lim, S., Paterson, B., Donnelly, P., Crouch, P. and White, A. (2012). Inhibition of TDP-43 Accumulation by Bis(thiosemicarbazonato)-Copper Complexes. *PLoS ONE*, 7(8), p.e42277.

Perez, M., Santa-María, I., Tortosa, E., Cuadros, R., Valle, M., Hernández, F., Moreno, F. and Avila, J. (2007). The role of the VQIVYK peptide in tau protein phosphorylation. *Journal of Neurochemistry*, 103(4), pp.1447-1460.

Petkova, A., Ishii, Y., Balbach, J., Antzutkin, O., Leapman, R., Delaglio, F. and Tycko, R. (2002). A structural model for Alzheimer's -amyloid fibrils based on experimental constraints from solid state NMR. *Proceedings of the National Academy of Sciences*, 99(26), pp.16742-16747.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF chimera-a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), pp.1605–1612.

Pirie, E., Oh, C., Zhang, X., Han, X., Cieplak, P., Scott, H., Deal, A., Ghatak, S., Martinez, F., Yeo, G., Yates, J., Nakamura, T. and Lipton, S., (2021). S-nitrosylated TDP-43 triggers aggregation, cell-to-cell spread, and neurotoxicity in hiPSCs and in vivo models of ALS/FTD. *Proceedings of the National Academy of Sciences*, 118(11), p.e2021368118.

Pottier, C., Bieniek, K., Finch, N., van de Vorst, M., Baker, M., Perkersen, R., Brown, P., Ravenscroft, T., van Blitterswijk, M., Nicholson, A., DeTure, M., Knopman, D., Josephs, K., Parisi, J., Petersen, R., Boylan, K., Boeve, B., Graff-Radford, N., Veltman, J., Gilissen, C., Murray, M., Dickson, D. and Rademakers, R. (2015). Whole-genome sequencing reveals important role for TBK1 and OPTN mutations in frontotemporal lobar degeneration without motor neuron disease. *Acta Neuropathologica*, 130(1), pp.77-92.

Pandey, N., Strider, J., Nolan, W.C., Yan, S.X., and Galvin, J.E. (2008). Curcumin inhibits aggregation of α-synuclein. *Acta Neuropathologica*, 115(4), pp.479–489.

Prasad, A., Raju, G., Sivalingam, V., Girdhar, A., Verma, M., Vats, A., Taneja, V., Prabusankar, G. and Patel, B., (2016). An acridine derivative, [4,5-bis{(N-carboxy

methyl imidazolium)methyl}acridine] dibromide, shows anti-TDP-43 aggregation effect in ALS disease models. *Scientific Reports*, 6(1).

Ramachandran, G. and Udgaonkar, J. (2011). Understanding the Kinetic Roles of the Inducer Heparin and of Rod-like Protofibrils during Amyloid Fibril Formation by Tau Protein. *Journal of Biological Chemistry*, 286(45), pp.38948-38959.

Rane, J.S., Bhaumik, P., and Panda, D. (2017). Curcumin inhibits tau aggregation and disintegrates preformed tau filaments in vitro. *Journal of Alzheimer's Disease*, 60(3), pp.999–1014.

Ravits, J., Paul, P. and Jorg, C. (2007). Focality of upper and lower motor neuron degeneration at the clinical onset of ALS. *Neurology*, 68(19), pp.1571-1575.

Reale, L.A., Dyer, M.S., Perry, S.E., Young, K.M., Dickson, T.C., Woodhouse, A., and Blizzard, C.A. (2023). Pathologically mislocalised TDP-43 in upper motor neurons causes a die-forward spread of als-like pathogenic changes throughout the Mouse Corticomotor System. *Progress in Neurobiology*, p.102449.

Ross, C. and Poirier, M. (2004). Protein aggregation and neurodegenerative disease. *Nature Medicine*, 10(S7), pp.S10-S17.

Ross, J., Olson, L. and Coppotelli, G. (2015). Mitochondrial and Ubiquitin Proteasome System Dysfunction in Ageing and Disease: Two Sides of the Same Coin?. *International Journal of Molecular Sciences*, 16(8), pp.19458-19476.

Roussel, G., Perpète, E.A., Matagne, A., Tinti, E., and Michaux, C. (2012). Towards a universal method for protein refolding: The trimeric beta barrel membrane omp2a as a test case. *Biotechnology and Bioengineering*, 110(2), pp.417–423.

Roussel, G., Rouse, S.L., Sansom, M.S.P., Michaux, C., and Perpète, E.A. (2014). The role of 2-methyl-2, 4-pentanediol in sodium dodecyl sulfate micelle dissociation

unveiled by dynamic light scattering and molecular dynamics simulations. *Colloids and Surfaces B: Biointerfaces*, 114, pp.357–362.

Rozema, D. and Gellman, S.H. (1995). Artificial chaperones: Protein refolding via sequential use of detergent and Cyclodextrin. *Journal of the American Chemical Society*, 117(8), pp.2373–2374.

Rozema, D. and Gellman, S.H. (1996). Artificial Chaperone-assisted refolding of carbonic anhydrase B. *Journal of Biological Chemistry*, 271(7), pp.3478–3487.

Rozema, D. and Gellman, S.H. (1996). Artificial Chaperone-assisted refolding of denatured-reduced lysozyme: modulation of the competition between Renaturation and aggregation. *Biochemistry*, 35(49), pp.15760–15771.

Rubino, E., Rainero, I., Chio, A., Rogaeva, E., Galimberti, D., Fenoglio, P., Grinberg, Y., Isaia, G., Calvo, A., Gentile, S., Bruni, A., St. George-Hyslop, P., Scarpini, E., Gallone, S. and Pinessi, L. (2012). SQSTM1 mutations in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Neurology*, 79(15), pp.1556-1562.

Saini, A. and Chauhan, V. (2011). Delineation of the Core Aggregation Sequences of TDP-43 C-Terminal Fragment. *ChemBioChem*, 12(16), pp.2495-2501.

Schein, C. (1989). Production of Soluble Recombinant Proteins in Bacteria. *Nature Biotechnology*, 7(11), pp.1141-1149.

Schmidt, N., Mishra, A., Lai, G.H., and Wong, G.C.L. (2009). Arginine-rich cell-penetrating peptides. *FEBS Letters*, 584(9), pp.1806–1813.

Scotter, E., Chen, H. and Shaw, C. (2015). TDP-43 Proteinopathy and ALS: Insights into Disease Mechanisms and Therapeutic Targets. *Neurotherapeutics*, 12(2), pp.352-363.

Seyfried, N., Gozal, Y., Dammer, E., Xia, Q., Duong, D., Cheng, D., Lah, J., Levey, A. and Peng, J. (2010). Multiplex SILAC Analysis of a Cellular TDP-43 Proteinopathy

Model Reveals Protein Inclusions Associated with SUMOylation and Diverse Polyubiquitin Chains. *Molecular & Cellular Proteomics*, 9(4), pp.705-718.

Shenouda, M., Xiao, S., MacNair, L., Lau, A. and Robertson, J. (2022). A C-Terminally Truncated TDP-43 Splice Isoform Exhibits Neuronal Specific Cytoplasmic Aggregation and Contributes to TDP-43 Pathology in ALS. *Frontiers in Neuroscience*, 16, p.868556.

Shimonaka, S., Nonaka, T., Suzuki, G., Hisanaga, S. and Hasegawa, M. (2016). Templated Aggregation of TAR DNA-binding Protein of 43 kDa (TDP-43) by Seeding with TDP-43 Peptide Fibrils. *Journal of Biological Chemistry*, 291(17), pp.8896-8907.

Skibinski, G., Parkinson, N., Brown, J., Chakrabarti, L., Lloyd, S., Hummerich, H., Nielsen, J., Hodges, J., Spillantini, M., Thusgaard, T., Brandner, S., Brun, A., Rossor, M., Gade, A., Johannsen, P., Sørensen, S., Gydesen, S., Fisher, E. and Collinge, J. (2005). Mutations in the endosomal ESCRTIII-complex subunit CHMP2B in frontotemporal dementia. *Nature Genetics*, 37(8), pp.806-808.

Sormanni, P., Aprile, F. and Vendruscolo, M. (2015). The CamSol Method of Rational Design of Protein Mutants with Enhanced Solubility. *Journal of Molecular Biology*, 427(2), pp.478-490.

Stefani M. (2010) Protein Aggregation Diseases: Toxicity of Soluble Prefibrillar Aggregates and Their Clinical Significance. In: Bross P., Gregersen N. (eds) Protein Misfolding and Cellular Stress in Disease and Aging. Methods in Molecular Biology (Methods and Protocols), vol 648. Humana Press, Totowa, NJ

Sun, J., Jiang, G., and Shigemori, H. (2019). Inhibitory activity on amyloid aggregation of rosmarinic acid and its substructures from *isodon japonicus*. *Natural Product Communications*, 14(5).

Sun, X., Song, J., Huang, H., Chen, H. and Qian, K. (2018). Modeling hallmark pathology using motor neurons derived from the family and sporadic amyotrophic lateral sclerosis patient-specific iPS cells. *Stem Cell Research & Therapy*, 9(1).

Tabner, B., El-Agnaf, O., Turnbull, S., German, M., Paleologou, K., Hayashi, Y., Cooper, L., Fullwood, N. and Allsop, D. (2005). Hydrogen Peroxide Is Generated during the Very Early Stages of Aggregation of the Amyloid Peptides Implicated in Alzheimer Disease and Familial British Dementia. *Journal of Biological Chemistry*, 280(43), pp.35789-35792.

Taguchi, R., Hatayama, K., Takahashi, T., Hayashi, T., Sato, Y., Sato, D., Ohta, K., Nakano, H., Seki, C., Endo, Y., Tokuraku, K., and Uwai, K. (2017). Structure–activity relations of rosmarinic acid derivatives for the amyloid β aggregation inhibition and antioxidant properties. *European Journal of Medicinal Chemistry*, 138, pp.1066–1075.

Taylor, J., Hardy, J., Fischbeck, K. (2002). Toxic Proteins in Neurodegenerative Disease. *Science*, 296(5575), pp.1991-1995.

Taylor, M., Moore, S., Mayes, J., Parkin, E., Beeg, M., Canovi, M., Gobbi, M., Mann, D.M., and Allsop, D. (2010). Development of a proteolytically stable retro-inverso peptide inhibitor of β-amyloid oligomerization as a potential novel treatment for alzheimer's disease. *Biochemistry*, 49(15), pp.3261–3272.

Taylor, M., Moore, S., Mourtas, S., Niarakis, A., Re, F., Zona, C., Ferla, B.L., Nicotra, F., Masserini, M., Antimisiaris, S.G., Gregori, M., and Allsop, D. (2011). Effect of curcumin-associated and lipid ligand-functionalized nanoliposomes on aggregation of the Alzheimer's AB peptide. *Nanomedicine: Nanotechnology, Biology and Medicine*, 7(5), pp.541–550.

Thakur, A. and Wetzel, R. (2002). Mutational analysis of the structural organization of polyglutamine aggregates. *Proceedings of the National Academy of Sciences*, 99(26), pp.17014-17019.

Tiselius, A., Hjerten, S., Levin, Ö. (1956). Protein chromatography on calcium phosphate columns. *Arch. Biochem. Biophys.* 65, pp.132–155.

Trageser, K., Smith, C., Herman, F., Ono, K. and Pasinetti, G. (2019). Mechanisms of Immune Activation by c9orf72-Expansions in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. *Frontiers in Neuroscience*, 13(1289).

Tseng, Y.-L., Lu, P.-C., Lee, C.-C., He, R.-Y., Huang, Y.-A., Tseng, Y.-C., Cheng, T.-J.R., Huang, J.J.-T., and Fang, J.-M. (2023). Degradation of neurodegenerative disease-associated TDP-43 aggregates and oligomers via a proteolysis-targeting Chimera. *Journal of Biomedical Science*, 30(1).

Un.org. (2019). *World Population Ageing 2019: Highlights*. [online] Available at: https://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2019-Highlights.pdf [Accessed 20 Jan. 2020].

Vance, C., Rogelj, B., Hortobagyi, T., De Vos, K., Nishimura, A., Sreedharan, J., Hu, X., Smith, B., Ruddy, D., Wright, P., Ganesalingam, J., Williams, K., Tripathi, V., Al-Saraj, S., Al-Chalabi, A., Leigh, P., Blair, I., Nicholson, G., de Belleroche, J., Gallo, J., Miller, C. and Shaw, C. (2009). Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. *Science*, 323(5918), pp.1208-1211.

Vega, M., Nigro, A., Luti, S., Capitini, C., Fani, G., Gonnelli, L., Boscaro, F. and Chiti, F. (2019). Isolation and characterization of soluble human full-length TDP-43 associated with neurodegeneration. *The FASEB Journal*, 33(10), pp.10780-10793.

Voráčková, I., Suchanová, Š., Ulbrich, P., Diehl, W.E. and Ruml, T. (2011). Purification of proteins containing zinc finger domains using immobilized metal ion affinity chromatography. *Protein Expression and Purification*. 79(1),pp.88–95.

Walsh, D., Klyubin, I., Fadeeva, J., Cullen, W., Anwyl, R., Wolfe, M., Rowan, M. and Selkoe, D. (2002). Naturally secreted oligomers of amyloid β protein potently inhibit hippocampal long-term potentiation in vivo. *Nature*, 416(6880), pp.535-539.

Wang, A., Conicella, A.E., Schmidt, H.B., Martin, E.W., Rhoads, S.N., Reeb, A.N., Nourse, A., Ramirez Montero, D., Ryan, V.H., Rohatgi, R., Shewmaker, F., Naik, M.T., Mittag, T., Ayala, Y.M., and Fawzi, N.L. (2018). A single n-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. *The EMBO Journal*, 37(5).

Wang, W., Wang, L., Lu, J., Siedlak, S., Fujioka, H., Liang, J., Jiang, S., Ma, X., Jiang, Z., da Rocha, E., Sheng, M., Choi, H., Lerou, P., Li, H. and Wang, X., (2016). The inhibition of TDP-43 mitochondrial localization blocks its neuronal toxicity. *Nature Medicine*, 22(8), pp.869-878.

Wang, X., Fan, H., Ying, Z., Li, B., Wang, H. and Wang, G. (2010). Degradation of TDP-43 and its pathogenic form by autophagy and the ubiquitin-proteasome system. *Neuroscience Letters*, 469(1), pp.112-116.

Wang, Y.T., Kuo, P.H., Chiang, C.H., Liang, J.R., Chen, Y.R., Wang, S., Shen, J.C., Yuan, H.S. (2013). The truncated C-terminal RNA recognition motif of TDP-43 protein plays a key role in forming proteinaceous aggregates, *Journal of Biological Chemistry*, 288, pp.9049–9057.

Watanabe, S., Kaneko, K. and Yamanaka, K. (2012). Accelerated Disease Onset with Stabilized Familial Amyotrophic Lateral Sclerosis (ALS)-linked Mutant TDP-43 Proteins. *Journal of Biological Chemistry*, 288(5), pp.3641-3654.

Watts, G., Wymer, J., Kovach, M., Mehta, S., Mumm, S., Darvish, D., Pestronk, A., Whyte, M. and Kimonis, V. (2004). Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nature Genetics*, 36(4), pp.377-381.

Wei, Y., Thyparambil, A.A., and Latour, R.A. (2014). Protein helical structure determination using CD spectroscopy for solutions with strong background absorbance from 190 to 230nm. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1844(12), pp.2331–2337.

Weskamp, K., Tank, E., Miguez, R., McBride, J., Gómez, N., White, M., Lin, Z., Gonzalez, C., Serio, A., Sreedharan, J. and Barmada, S. (2020). Shortened TDP43 isoforms upregulated by neuronal hyperactivity drive TDP43 pathology in ALS. *Journal of Clinical Investigation*, 130(3), pp.1139-1155.

Wördehoff, M. and Hoyer, W. (2018). A-synuclein aggregation monitored by thioflavin T fluorescence assay. *BIO-PROTOCOL*, 8(14), p.e2941.

Wright, G., Watanabe, T., Amporndanai, K., Plotkin, S., Cashman, N., Antonyuk, S. and Hasnain, S., (2020). Purification and Structural Characterization of Aggregation-Prone Human TDP-43 Involved in Neurodegenerative Diseases. iScience, 23(6), p.101159.

Xue, C., Lin, T., Chang, D. and Guo, Z. (2017). Thioflavin T as an amyloid dye: fibril quantification, optimal concentration and effect on aggregation. *Royal Society Open Science*, 4(1), p.160696.

Yamashita, M., Nonaka, T., Hirai, S., Miwa, A., Okado, H., Arai, T., Hosokawa, M., Akiyama, H. and Hasegawa, M. (2014). Distinct pathways leading to TDP-43-induced cellular dysfunctions. *Human Molecular Genetics*, 23(16), pp.4345-4356.

Yang, F., Lim, G.P., Begum, A.N., Ubeda, O.J., Simmons, M.R., Ambegaokar, S.S., Chen, P.P., Kayed, R., Glabe, C.G., Frautschy, S.A., and Cole, G.M. (2005). Curcumin inhibits formation of amyloid β oligomers and fibrils, binds plaques, and reduces amyloid in vivo. *Journal of Biological Chemistry*, 280(7), pp.5892–5901.

Zarei, S., Carr, K., Reiley, L., Diaz, K., Guerra, O., Altamirano, PF., Pagani, W., Lodin, D., Orozco, G., Chinea, A. (2015) A comprehensive review of amyotrophic lateral sclerosis. Surgical Neurology International, 6(171) doi: 10.4103/2152-7806.169561

Zou, Z., Zhou, Z., Che, C., Liu, C., He, R. and Huang, H. (2017). Genetic epidemiology of amyotrophic lateral sclerosis: a systematic review and meta-analysis. *Journal of Neurology, Neurosurgery & Psychiatry*, 88(7), pp.540-549.

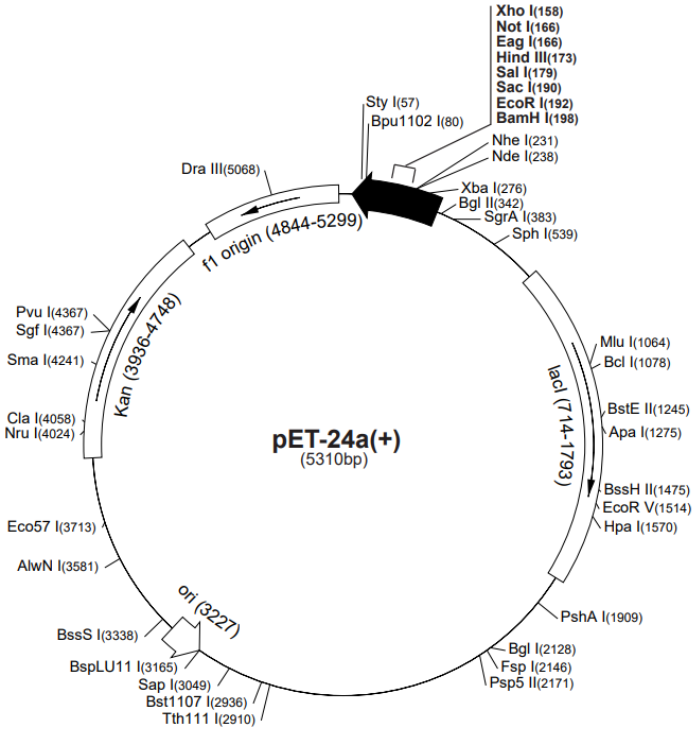
Zhang, L., Chen, Y., Liu, M., Wang, Y. and Peng, G. (2020). TDP-43 and Limbic-Predominant Age-Related TDP-43 Encephalopathy. *Frontiers in Aging Neuroscience*, 11(376).

Zhang, Y., Xu, Y., Cook, C., Gendron, T., Roettges, P., Link, C., Lin, W., Tong, J., Castanedes-Casey, M., Ash, P., Gass, J., Rangachari, V., Buratti, E., Baralle, F., Golde, T., Dickson, D. and Petrucelli, L. (2009). Aberrant cleavage of TDP-43 enhances aggregation and cellular toxicity. *Proceedings of the National Academy of Sciences*, 106(18), pp.7607-7612.

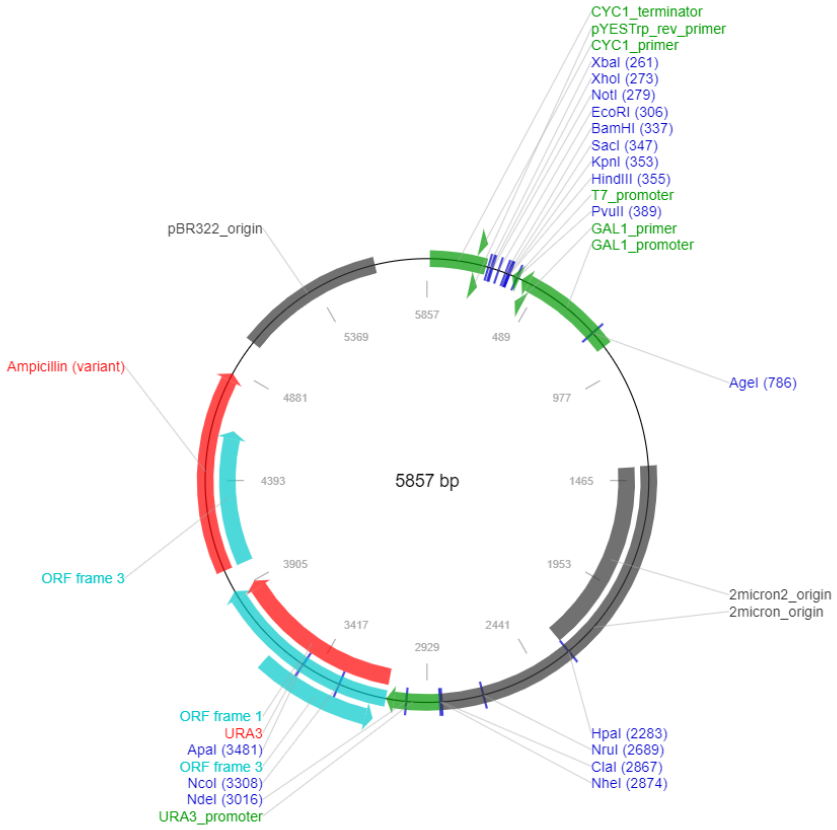
Zhao, M., Yao, X., Wei, P., Zhao, C., Cheng, M., Zhang, D., Xue, W., He, W., Xue, W., Zuo, X., Jiang, L., Luo, Z., Song, J., Shu, W., Yuan, H., Liang, Y., Sun, H., Zhou, Y., Zhou, Y., Zheng, L., Hu, H., Wang, J. and Du, H., (2021). O-GlcNAcylation of TDP-43 suppresses proteinopathies and promotes TDP-43's mRNA splicing activity. *EMBO reports*, e51649.

11. Appendices

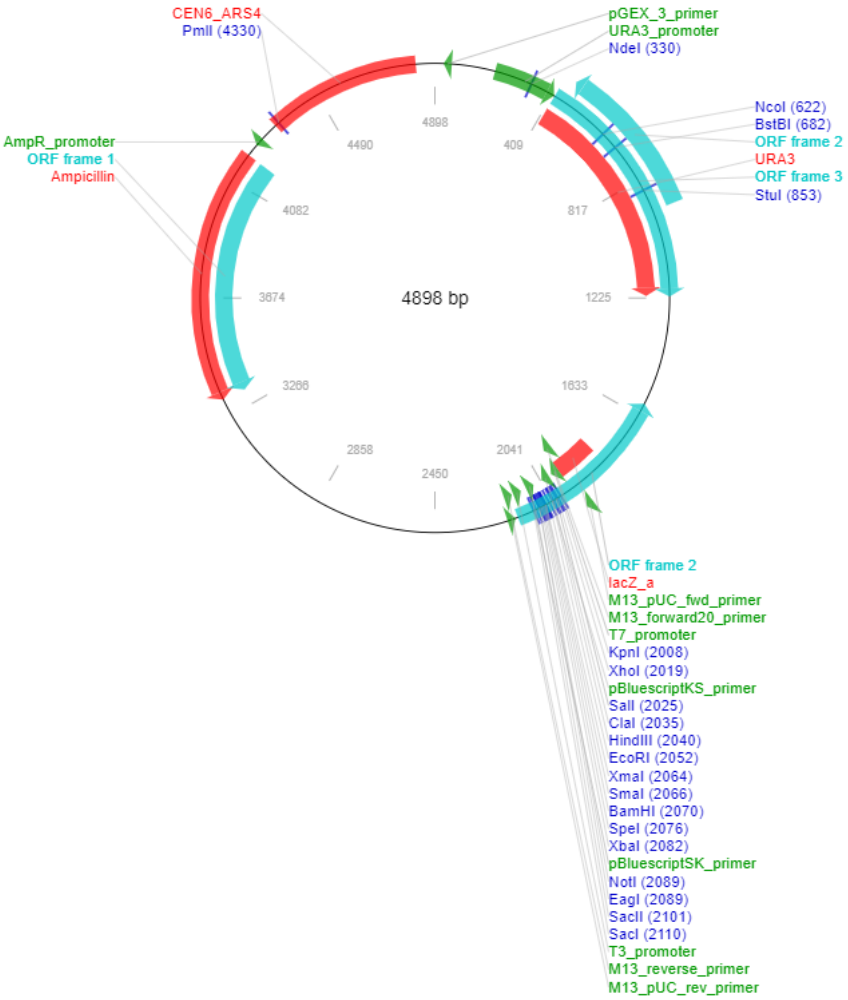
Appendix A: pET24a plasmid map



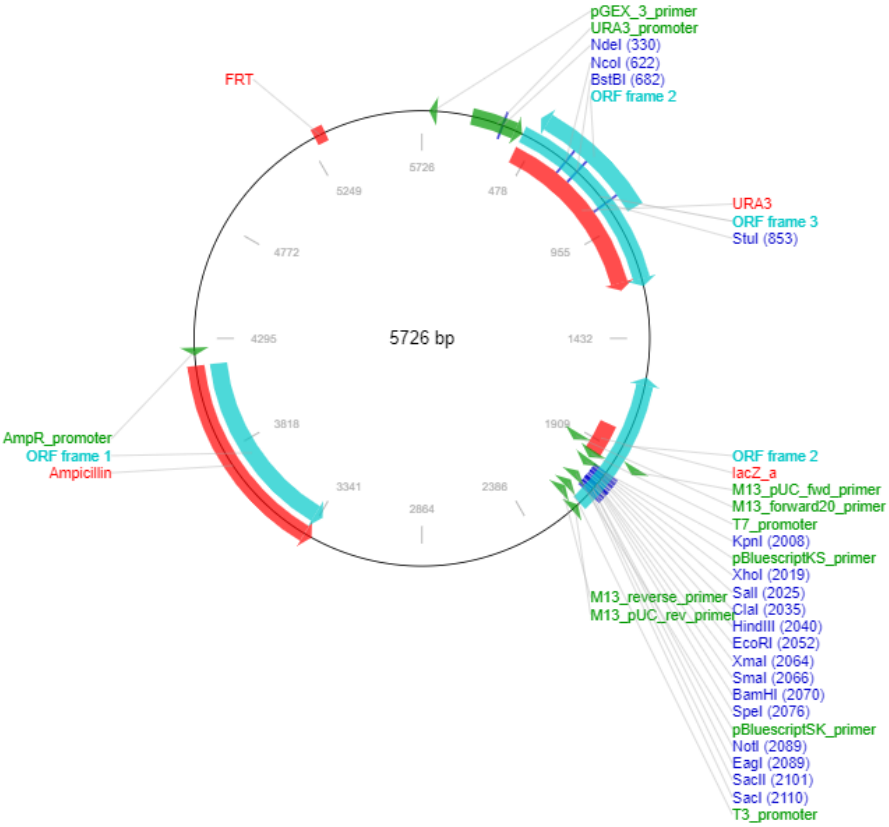
Appendix B: pYES2 plasmid map



Appendix C: pRS416 plasmid map



Appendix D: pRS426 plasmid map



Appendix E: Full length TDP-43 amino acid sequence

MSEYIRVTEDENDEPIEIPSEDDGTVLLSTVTAQFPGACGLRYRNPVSQCMRGVRLVEGILHA

PDAGWGNLVYVVNYPKDNKRKMDETDASSAVKVKRAVQKTSDLIVLGLPWKTTEQDLKEYFST

FGEVLMVQVKKDLKTGHSKGFGFVRFTEYETQVKVMSQRHMIDGRWCDCKLPNSKQSQDEPLR

SRKVFVGRCTEDMTEDELREFFSQYGDVMDVFIPKPFRAFAFVTFADDQIAQSLCGEDLIIKG

ISVHISNAEPKHNSNRQLERSGRFGGNPGGFGNQGGFGNSRGGGAGLGNNQGSNMGGGMNFGA

FSINPAMMAAAQAALQSSWGMMGMLASQQNQSGPSGNNQNQGNMQREPNQAFGSGNNSYSGSN

SGAAIGWGSASNAGSGSGFNGGFGSSMDSKSSGWGM

Appendix F: TDP-43 CTF (208-414 C-terminal 6xHis) amino acid sequence

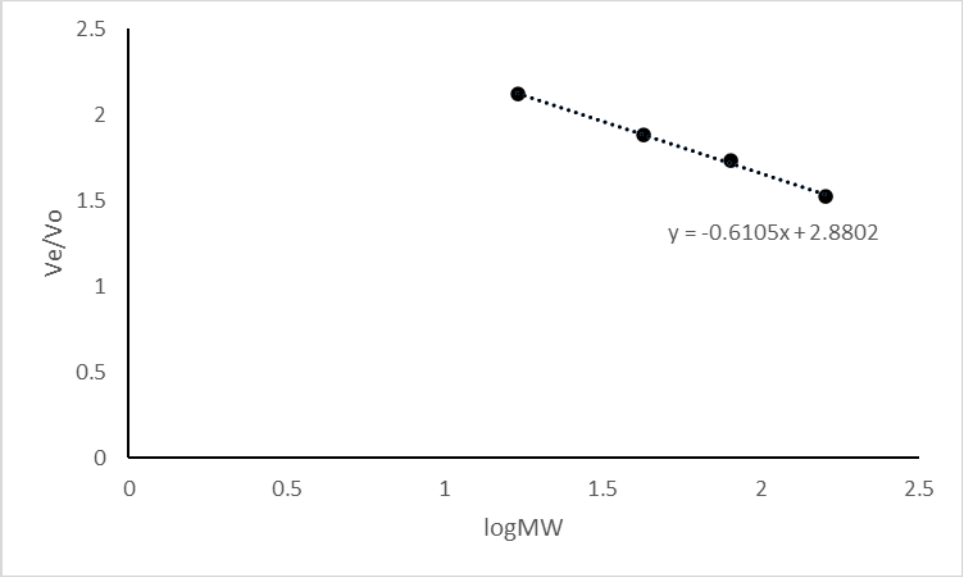
MREFFSQYGDVMDVFIPKPFRAXAFVTFADDQIAQSLCGEDLIIKGISVHISNAEPKHNSNRQ

LERSGRFGGNPGGFGNQGGFGNSRGGGAGLGNNQGSNMGGGINFGAFSINPAMMAAAQAALQS

SWGMMGMLASQQNQSGPSGNNQNQGNMQREPNQAFGSGNNSYSGSNSGAAIGWGSASNAGSGS

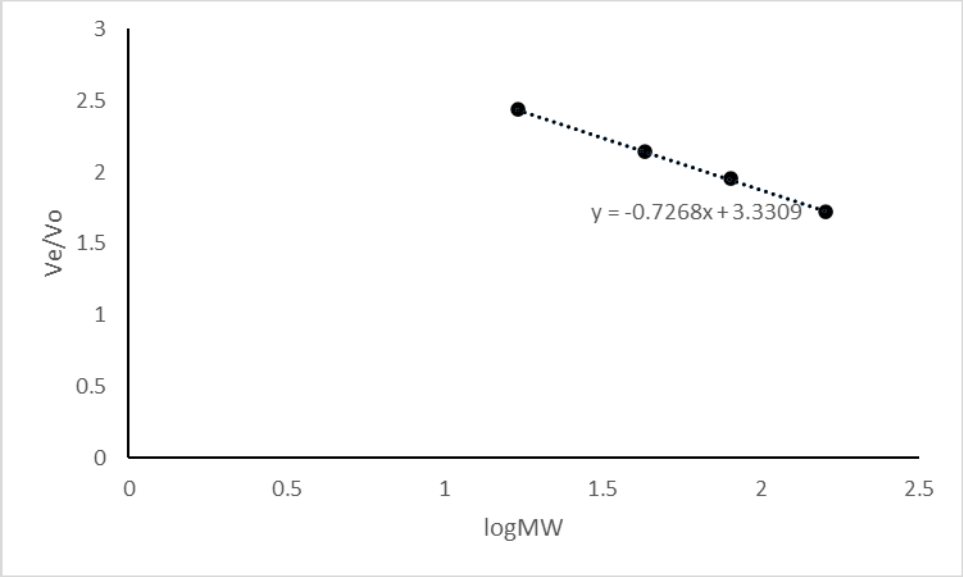
GFNGGFGSSMDSKSSGWGMLEHHHHHH

Appendix G: HiPrep 26/60 Sephacryl S-300 HR calibration curve



Appendix G: HiPrep 26/60 Sephacryl S-300 HR calibration curve. Calibration of a HiPrep 26/60 Sephacryl S-300 HR column using four standard proteins; Aldolase, Transferrin, Ovalbumin and Myoglobin. V_e/V_o indicates elution volume/void volume for each protein, plotted against the log of the protein's molecular weight. Equation of the line used to estimate the molecular weight of proteins applied to the column based on their elution volume.

Appendix H: Superdex 200 10/300 gl calibration curve



Appendix H: Superdex 200 10/300 gl calibration curve. Calibration of a Superdex 200 10/300 gl column using four standard proteins; Aldolase, Transferrin, Ovalbumin and Myoglobin. V_e/V_o indicates elution volume/void volume for each protein, plotted against the log of the protein's molecular weight. Equation of the line used to estimate the molecular weight of proteins applied to the column based on their elution volume.