2020

Characterization of the Dimerization and Salt Dependent Aggregation of the ALS-Linked Protein TDP-43

Elizabeth Ehrhardt
University of Denver

Recommended Citation

Ehrhardt, Elizabeth, "Characterization of the Dimerization and Salt Dependent Aggregation of the ALS-Linked Protein TDP-43" (2020). *Electronic Theses and Dissertations*. 1748.
https://digitalcommons.du.edu/etd/1748

Characterization of the Dimerization and Salt Dependent Aggregation of the ALS-Linked Protein TDP-43

Abstract

Trans-active response (TAR) DNA-binding protein 43 (TDP-43) is essential for RNA processing but can also form toxic cytoplasmic inclusions in neurons of patients with amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD). RNA-binding has been shown to have the potential to decrease or inhibit the aggregation of TDP-43, indicating that supplying RNA therapeutics may be a solution to treat these neurodegenerative disorders. However, the mechanism of aggregation, transitioning from TDP-43's native dimeric state to small oligomers to toxic aggregates, is still relatively unknown. This needs to be established before determining how to target and disrupt this aggregation. Using GFP-TDP-43, we first sought to further classify the native state of TDP-43 by characterizing its dimer interfaces by structure-guided mutagenesis. Combined, our results indicate that TDP-43 can form head-to-tail dimerization between two N-terminal domains (NTD) and head-to-head dimerization between two RNA-recognition motifs 2 (RRM2). When both interfaces are disrupted, we isolate a tentatively assigned monomeric species. The second aim of our study was to determine the propensity for aggregation of the same family of TDP-43 mutants. Using a denaturing purification method and low ionic strength buffers, we obtain a metastable intermediate of GFP-TDP-43 that undergoes inducible aggregation when electrolytes are added back to solution. By monitoring protein aggregation using turbidity assays and dynamic light scattering (DLS) measurements we observe that the addition of smaller divalent cations induces faster aggregation and produces larger aggregates than corresponding amounts of larger monovalent cations. These data support a model in which different electrolytes alter the zeta potential of TDP-43. Finally, we show that aggregation can be mitigated by the addition of DNA oligomers with some degree of sequence specificity. Excess TG12 DNA significantly reduces the rate of aggregation and size of aggregated species of GFP-TDP-43 whereas excess AC12 DNA has no significant difference in aggregation from GFP-TDP-43 without DNA. This result supports the hypothesis that UG-rich RNA should be considered for development of a RNA therapeutic drug to help patients with ALS.

Document Type

Thesis

Degree Name

M. S.

Department

Chemistry and Biochemistry

First Advisor

Erich G. Chapman

Second Advisor

Martin Margittai

Third Advisor

Scott Horowitz

Keywords

Aggregation, Dimer, Electrolytes, TDP-43

Characterization of the Dimerization and Salt Dependent Aggregation of the ALS-Linked Protein TDP-43

A Thesis
Presented to
the Faculty of the College of Natural Sciences and Mathematics
University of Denver

In Partial Fulfillment
of the Requirements for the Degree
Master of Science

---

by
Elizabeth Ehrhardt
June 2020
Advisor: Erich G. Chapman

Author: Elizabeth Ehrhardt
Title: Characterization of the Dimerization and Salt Dependent Aggregation of the ALS-Linked Protein TDP-43
Advisor: Erich G. Chapman
Degree Date: June 2020

ABSTRACT

Trans-active response (TAR) DNA-binding protein 43 (TDP-43) is essential for RNA processing but can also form toxic cytoplasmic inclusions in neurons of patients with amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD). RNA-binding has been shown to have the potential to decrease or inhibit the aggregation of TDP-43, indicating that supplying RNA therapeutics may be a solution to treat these neurodegenerative disorders. However, the mechanism of aggregation, transitioning from TDP-43's native dimeric state to small oligomers to toxic aggregates, is still relatively unknown. This needs to be established before determining how to target and disrupt this aggregation. Using GFP-TDP-43, we first sought to further classify the native state of TDP-43 by characterizing its dimer interfaces by structure-guided mutagenesis. Combined, our results indicate that TDP-43 can form head-to-tail dimerization between two N-terminal domains (NTD) and head-to-head dimerization between two RNA-recognition motifs 2 (RRM2). When both interfaces are disrupted, we isolate a tentatively assigned monomeric species. The second aim of our study was to determine the propensity for aggregation of the same family of TDP-43 mutants. Using a denaturing purification method and low ionic strength buffers, we obtain a metastable intermediate of GFP-TDP-43 that undergoes inducible aggregation when electrolytes are added back to solution. By monitoring protein aggregation using turbidity assays and dynamic light scattering (DLS) measurements we observe that the addition of smaller divalent cations induces faster aggregation and produces larger aggregates than corresponding amounts of larger monovalent cations. These data support a model in which different electrolytes alter the zeta potential of TDP-43. Finally, we show that aggregation can be mitigated by the addition of DNA oligomers with some degree of sequence specificity. Excess TG12 DNA significantly reduces the rate of aggregation and size of aggregated species of GFP-TDP-43 whereas excess AC12 DNA has no significant difference in aggregation from GFP-TDP-43 without DNA. This result supports the hypothesis that UG-rich RNA should be considered for development of a RNA therapeutic drug to help patients with ALS.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

CHAPTER 3

APPENDIX

CHAPTER 1: INTRODUCTION

1.1 Pathological Implications of TDP-43

1.1.1 Protein Aggregation and Neurodegenerative Diseases

There are several proteins that have been linked to neurodegenerative disorders. In patients with Alzheimer's Disease (AD), amyloid-β (Aβ) is present in large plaques seeding the aggregation of τ-protein fibrils and these toxic intermediates cause mis-localization of the protein oligomers in the cells (Thal et al. 2015). Similarly, alpha-synuclein aggregates are found in the neurons of patients with Parkinson's Disease (PD) that seed the formation of toxic Lewy bodies in the cytoplasm of neurons (Fields et al. 2015). Huntingtin protein (Htt) has also been shown to trigger Huntington's disease (HD) by aggregating into inclusion bodies (IBs) along the cerebral cortex and spinal cord causing interference of the synaptic transmissions required for normal body function (Arrasate et al. 2012). Lastly, cytoplasmic inclusions of TDP-43 are seen in approximately 97% of patients with amyotrophic lateral sclerosis (ALS) and 45% of patients with frontotemporal lobar degeneration (FTLD) (Prasad 2019).

Misfolding of each of these proteins is a key component in the aggregation pathway that leads to these neurodegenerative diseases (McEwan et al. 2016; Fields 2015; Arrasate 2012). This can be triggered by mutations in the gene of each protein or cellular stress in the neuron (Fig. 1A) (Thal 2015; Fields 2015; Arrasate 2012). Proteins predisposed to aggregation begin to form small reactive intermediates that lead to the formation of small oligomers (Fig. 1). These intermediates can cause neuronal damage and can eventually form large irreversible and toxic polymers in the cytoplasm of the neurons (Fig. 1) (Morbidelli et al. 2016). The intermediates, small oligomers, are seen to be reversible in many of these neurodegenerative diseases; giving a good target for therapeutics. However, each pathway has a different rate of nucleation depending on the total mass of the fibrils and fragmentation and can be hard to target since frequently the point of irreversible toxicity is unknown (Fig. 1B) (Morbidelli 2016). In addition, TDP-43 poses an even more complex mechanism to target since the full-length 3D structure is still not completely known. The contributions from the N-terminal domain and C-terminal domain to the aggregation pathway have been derived from NMR crystal structure analysis of these domains. However, due to the flexible linkers present between them and the dynamics of the low-complexity domain, the full length has not been able to be resolved (Francois-Moutal et al. 2019). Since large aggregates of several proteins are seen to be a pathological hallmark in AD, PD and HD, targeting the toxic but reversible intermediate species has been the main approach of researchers in establishing therapeutics for these still untreatable diseases (Fields 2015; McEwan 2016).

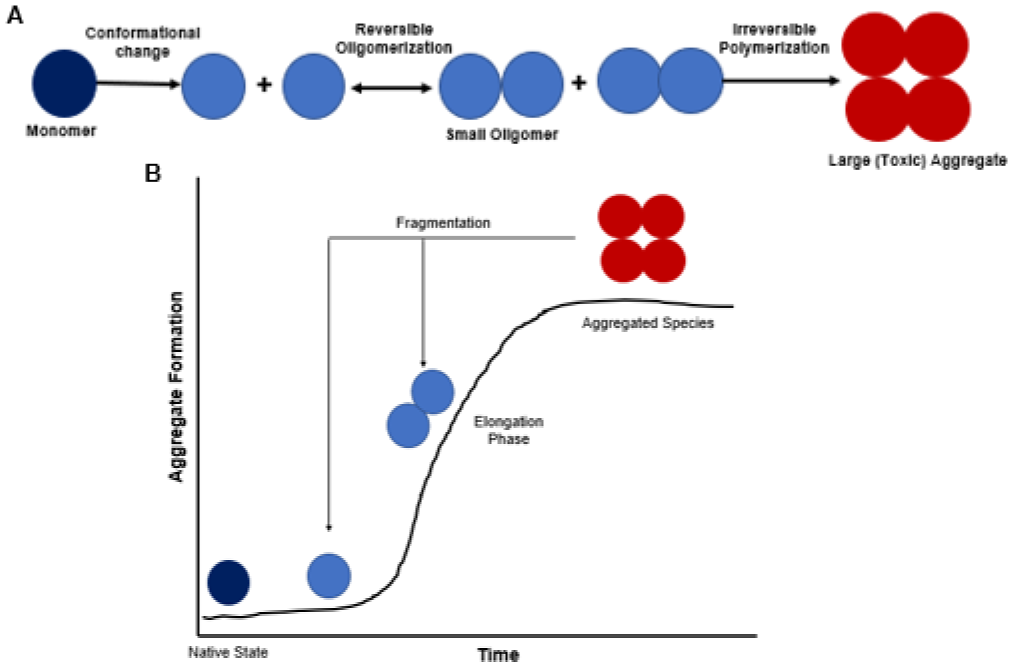


Figure 1. (A) Aggregation pathway for proteins involved in neurodegenerative diseases. Monomers undergo a conformational change due to cellular stress or mutations. These new monomers are then prone come together through reversible oligomerization. Under continuing stress, the small oligomers can polymerize into irreversible, toxic aggregates characteristic of the conformation found in patients with AD, PD, Htt, FTLD and ALS. (B) Kinetic model of the aggregation pathway. Small oligomers are formed during the elongation phase and continues oligomerizing until it reaches its final toxic aggregate species where the graph plateaus. That aggregated species can fragment to create more elongated oligomers or conformationally changed monomers, initiating the aggregation pathway to form more toxic aggregates.

1.1.2 Amyotrophic Lateral Sclerosis (ALS)

ALS is a neurodegenerative disorder that affects motor neurons in the cerebral cortex and spinal cord causing loss of motor function in the arms and legs, as well as, difficulty in speech, swallowing and breathing (Cooper-Knock et al. 2012; Morgan et al. 2016). It affects approximately 6 in every 100,000 people in the population (Cooper-Knock 2012), with 90% of the cases being sporadic and 10% familial (Morgan 2016). With no cure currently known, death from this disease commonly occurs within 3-5 years of symptomatic onset (Morgan 2016). Although the exact mechanism for the pathogenesis of this disease is unknown, there are several mutated genes that are frequently seen in patients with ALS: *C9orf72 (C9)*, Superoxide dismutase 1 *(SOD1)*, TAR-DNA binding protein *(TARDBP)*, and Fused in Sarcoma *(FUS)* (Fig. 2) (Cooper-Knock 2012; Morgan 2016; Coyne 2017; Prasad 2019).

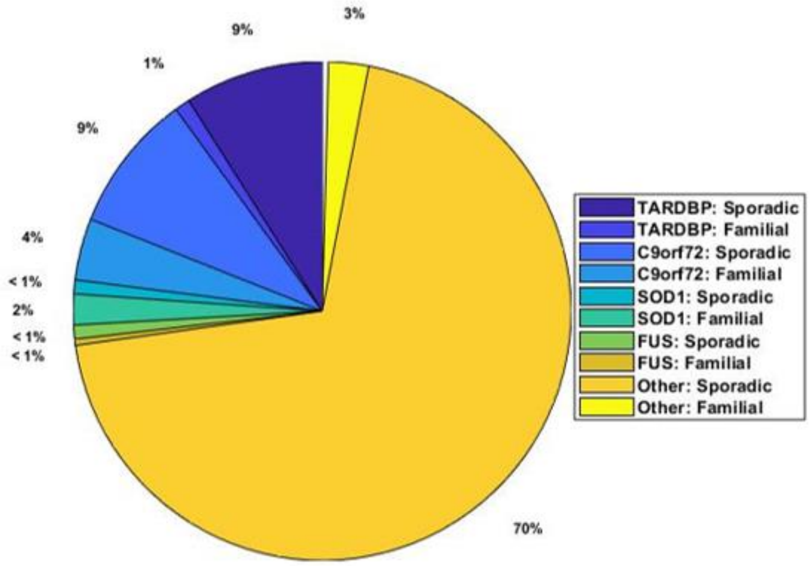


Figure 2. Breakdown of the prevalence of common genes (*C9orf72, SOD1, TARDBP*) that when mutated are seen in familial/sporadic cases of ALS.

Repeat expansion of GGGGCC (G_4C_2) region of *C9orf72* is normal up to about 20 repeats; however, in patients with ALS this expansion can continue for hundreds of repeats (Morgan 2016; Coyne 2017). This expansion can alter RNA metabolism and cause the formation of RNA stress granules (SGs) through altered splicing events and has even been linked to the redistribution of TDP-43 in the neuron; both factors are seen in the patients with ALS (Coyne 2017). *SOD1* was the first gene found to be associated with ALS; mutations prevalent in approximately 20% of familial cases and 1% of sporadic cases (Morgan 2016). *SOD1* interacts with TDP-43 to form detergent-insoluble fractions that alter mRNA stability in the neurons (Prasad 2019). Similarly, *FUS* co-localizes with TDP-43 and forms SGs that reduce the expression of some mRNA necessary for cellular function; several mutations in the gene cause a higher affinity for interaction between the two proteins (Prasad 2019; Morgan 2016). *TARDBP* is found on chromosome 1 and codes for TDP-43. Mutations in exon 6, which codes for the disordered C-terminal domain (CTD), are prevalent in 5-10% of sporadic and familial cases of ALS (Fig. 2) (Prasad 2019). Like *C9orf72*, these mutations can also cause mis-localization of TDP-43 into the cytoplasm, producing cytoplasmic ubiquitinated protein inclusions characteristic of ALS, to form (Prasad 2019; Cooper-Knock 2012; Morgan 2016). With a decrease of *TARDBP* in the nuclease there is a loss of regulation of mRNA metabolism and with overexpression of *TARDBP* improper splicing tends to occur from exon skipping, both yielding loss of normal biological function (Morgan 2016; Prasad 2019).

1.1.3 Mutations, Mis-localization, and Cytoplasmic Inclusions of TDP-43

TDP-43, along with several other neurodegenerative disease-linked proteins, can undergo liquid-liquid phase separation (LLPS) through its low complexity domain (LCD), which produces small reversible oligomers through charge screening and hydrophobic interactions (Fig. 3) (Babinchak et al. 2019). Current models propose that over time these oligomers become larger and less dynamic and eventually reach a point where they are toxic and irreversible; these are the cytoplasmic inclusions seen in the neurons of patients with ALS (Fig. 3) (Babinchak 2019; Sun et al. 2017). There are a multitude of mutations affiliated with ALS in the CTD of TDP-43 that increase the likelihood of the protein to undergo LLPS, such as, Q331K (causes LLPS through hydrophobic and electrostatic interactions) and M337V (alters dynamics of the liquid droplets) (Table 1) (Sun 2017; Li et al. 2018). However, there are also several mutations that occur earlier in the sequence towards the end of the N-terminal domain (NTD) in the nuclear localization sequence (NLS) region (Table 1). Deletion or mutation of this sequence can cause TDP-43 to relocalize to the cytoplasm rather than the nucleus and at high enough concentrations this can cause TDP-43 to aggregate into cytoplasmic inclusions (Prasad 2019; Huang 2013).

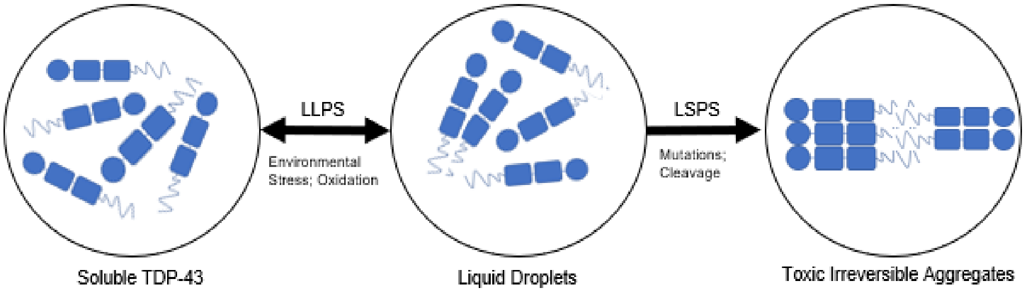
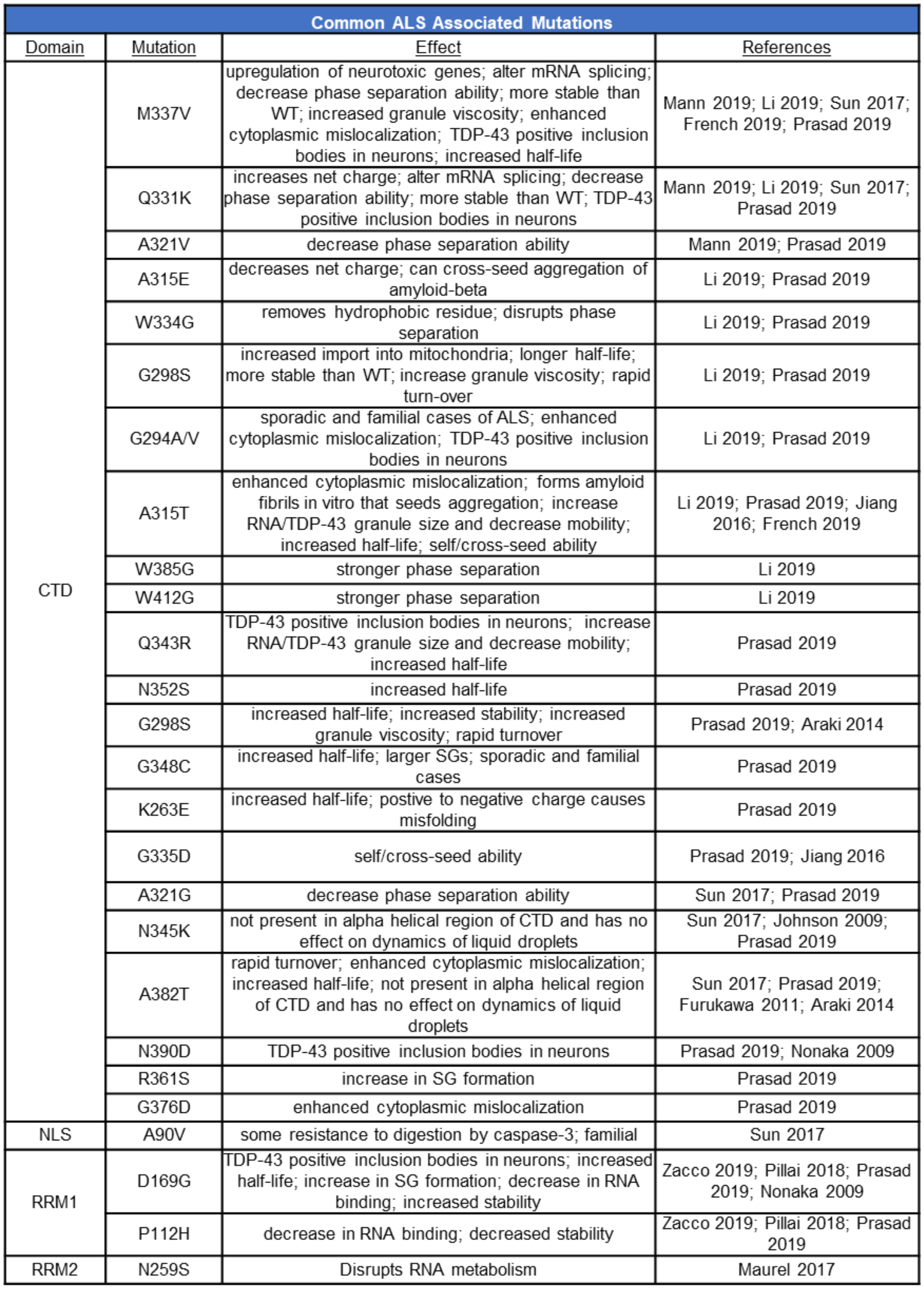


Figure 3. Schematic of the aggregation pathway of TDP-43 in the cytoplasm of a neuron. TDP-43 in the monomeric or dimeric state is soluble under normal physiological conditions. Environmental or oxidative stress to the cell can cause it to under-go liquid-liquid phase separation (LLPS) and form small, reversible, liquid droplets. Under continuous stress, cleavage or presence of mutations, TDP-43 will under-go liquid-solid phase separation and form toxic irreversible aggregates that are characteristic of the cytoplasmic inclusions found in patients with ALS

Using ProteoStat emission fluorescence assays, Zacco and co-workers (Zacco 2019) showed that RNA binding can mitigate this aggregation. Using two ssRNAs, one that has high binding affinity to TDP-43 and one with low binding affinity, they showed a decrease in measured PreoteoStat emission fluorescence in the sample containing the high-affinity ssRNA compared to the control with no ssRNA (Zacco 2019). The sample with the low-affinity ssRNA showed no significant effect on the aggregation indicating that RNA-binding is likely causing aggregation to decrease (Zacco 2019). Mutations in the RNA recognition motifs (RRMs), however, can disrupt this RNA-binding. These mutations can be found on the loops between the β-sheets and α-helices of the RRM1 domain with either direct disruption of RNA binding or causing cleavage of TDP-43 that indirectly prevents the RNA binding necessary for aggregation inhibition (Sun 2017).



Common ALS Associated Mutations

Domain	Mutation	Effect	References
CTD	M337V	upregulation of neurotoxic genes; alter mRNA splicing; decrease phase separation ability; more stable than WT; increased granule viscosity; enhanced cytoplasmic mislocalization; TDP-43 positive inclusion bodies in neurons; increased half-life	Mann 2019; Li 2019; Sun 2017; French 2019; Prasad 2019
	Q331K	increases net charge; alter mRNA splicing; decrease phase separation ability; more stable than WT; TDP-43 positive inclusion bodies in neurons	Mann 2019; Li 2019; Sun 2017; Prasad 2019
	A321V	decrease phase separation ability	Mann 2019; Prasad 2019
	A315E	decreases net charge; can cross-seed aggregation of amyloid-beta	Li 2019; Prasad 2019
	W334G	removes hydrophobic residue; disrupts phase separation	Li 2019; Prasad 2019
	G298S	increased import into mitochondria; longer half-life; more stable than WT; increase granule viscosity; rapid turn-over	Li 2019; Prasad 2019
	G294A/V	sporadic and familial cases of ALS; enhanced cytoplasmic mislocalization; TDP-43 positive inclusion bodies in neurons	Li 2019; Prasad 2019
	A315T	enhanced cytoplasmic mislocalization; forms amyloid fibrils in vitro that seeds aggregation; increase RNA/TDP-43 granule size and decrease mobility; increased half-life; self/cross-seed ability	Li 2019; Prasad 2019; Jiang 2016; French 2019
	W385G	stronger phase separation	Li 2019
	W412G	stronger phase separation	Li 2019
	Q343R	TDP-43 positive inclusion bodies in neurons; increase RNA/TDP-43 granule size and decrease mobility; increased half-life	Prasad 2019
	N352S	increased half-life	Prasad 2019
	G298S	increased half-life; increased stability; increased granule viscosity; rapid turnover	Prasad 2019; Araki 2014
	G348C	increased half-life; larger SGs; sporadic and familial cases	Prasad 2019
	K263E	increased half-life; postive to negative charge causes misfolding	Prasad 2019
	G335D	self/cross-seed ability	Prasad 2019; Jiang 2016
	A321G	decrease phase separation ability	Sun 2017; Prasad 2019
	N345K	not present in alpha helical region of CTD and has no effect on dynamics of liquid droplets	Sun 2017; Johnson 2009; Prasad 2019
	A382T	rapid turnover; enhanced cytoplasmic mislocalization; increased half-life; not present in alpha helical region of CTD and has no effect on dynamics of liquid droplets	Sun 2017; Prasad 2019; Furukawa 2011; Araki 2014
	N390D	TDP-43 positive inclusion bodies in neurons	Prasad 2019; Nonaka 2009
	R361S	increase in SG formation	Prasad 2019
	G376D	enhanced cytoplasmic mislocalization	Prasad 2019
NLS	A90V	some resistance to digestion by caspase-3; familial	Sun 2017
RRM1	D169G	TDP-43 positive inclusion bodies in neurons; increased half-life; increase in SG formation; decrease in RNA binding; increased stability	Zacco 2019; Pillai 2018; Prasad 2019; Nonaka 2009
	P112H	decrease in RNA binding; decreased stability	Zacco 2019; Pillai 2018; Prasad 2019
RRM2	N259S	Disrupts RNA metabolism	Maurel 2017

Table 1. Common ALS-associated mutations of TDP-43 found in sporadic and familial cases.

1.2 Physiological Importance of TDP-43

The protein TDP-43 plays several roles in RNA metabolism: transcription, translation, mRNA transport, mRNA splicing, and long non-coding RNA (lncRNA) processing (Fig. 4) (Prasad et al. 2019). TDP-43 has a high-affinity to bind to RNA with UG-rich sequences found in approximately one-third of the human transcriptome, (Prasad 2019 & Weskamp et al. 2018), which allows it to play a role in thousands of transcripts essential to the biological function of neurons. TDP-43 is primarily found in areas of the nucleus essential for mRNA transcription and splicing. It is typically in a dimer state with RNA bound to it (Fig. 4) (Coyne et al. 2017). TDP-43 also plays a role in mRNA transport, by joining with RNA to form ribonucleoproteins (RNPs) which transport mRNA from the nucleus into the cytoplasm for translation (Fig. 4) (Prasad 2019). Under normal conditions, TDP-43 would then move back into the nucleus after dissociating from the RNP. However, under cellular stress events, TDP-43 associated with RNPs can form into stress granules (SGs) which inhibits the translation of mRNA except for those necessary for cellular repair (Fig. 2) (Prasad 2019; Coyne 2017). SG formation happens directly through protein-protein and RNA interactions mediated by the low-complexity domain (LCD) of TDP-43 through liquid-liquid phase separation (LLPS) (Mann et al. 2019; Wang et al. 2018; Li 2019). These SGs are reversible and dissociate once the cellular stress has diminished, which increases the concentration of TDP-43 in the cell (Prasad 2019; Coyne 2017).

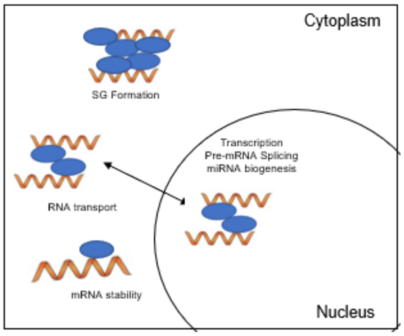


Figure 4. The physiological functions of TDP-43 in the neuron. TDP-43 (blue ovals) is found in a dimeric state bound to RNA (orange) to facilitate RNA processes. In the nucleus TDP-43 participates in transcription, pre-mRNA splicing and miRNA biogenesis. It than transports RNA into the cytoplasm. In the cytoplasm it can either dissociate from the RNA and shuttle back into the nucleus or facilitate mRNA stability or coagulate into stress granules (SG) to decrease environmental stress occurring in the neuron.

Due to its essential function in the body, TDP-43's concentration and location needs to be highly regulated within the neuron. For short term regulation, it can be controlled by cellular stress, indicating that it has a biological function helping maintain the health of the cell (Weskamp 2018). It has an autoregulation function through a negative feedback loop by binding with its own 3'-UTR region, causing reduced TDP-43 expression due to mRNA destabilization and RNA degradation (Weskamp 2018; Coyne 2017). This autoregulation allows for the protein to co-localize with SGs in the cell to help with cellular stress and biological function of the neuron (Coyne 2017). It also provides indirect regulation of gene expression through maintaining RNA transcript stability. Over-expression of TDP-43 causes a decrease in *tau* gene expression, which has been seen in patients with frontotemporal lobar degeneration (FTLD) (Coyne 2017; Papegaey et al. 2016). A decrease in expression of TDP-43 causes a decrease in *beta-adducin 2* gene expression, which would cause a loss of control over motor coordination and memory process (Coyne 2017; Kiang *et al.* 2018). The regulation also plays a role in CFTR splicing implicated in many diseases but has a normal biological function of regulating salt ion channels (Coyne 2017 & Buratti et al. 2004). Thus, physiological function of TDP-43 facilitates RNA processes in neurons that play a vital role in cell survival, however without proper regulation, TDP-43 can also promote several neurodegenerative disorders.

1.3 Structure of TDP-43

1.3.1 N-Terminal Domain (NTD)

TDP-43 contains 414 amino acids (aa) residues; the first 102 aa residues make up the NTD (Fig. 5) (Prasad 2019; Shimonaka et al. 2016; Huang 2013; Sasaguri et al. 2016). The NTD is a structured region that contains the nuclear localization signal (NLS) responsible for cytoplasmic shuttling during RNA processing (Zacco et al. 2018; Prasad 2019; McGurk et al. 2018; Huang 2013; Sasaguri 2016). There are several dimer interfaces proposed in the NTD, and the homodimerization within this domain allows for more efficient RNA binding and regulates biological activity (Sasaguri 2016; Wang et al. 2018; Jang et al. 2017). The NTD also contains two regions, one within an ubiquitin-like fold and a disordered lysine and arginine region, that support RNA binding (Jiang 2017; Wang 2018). These regions allow for weak dispersion interactions between the aromatic rings on the protein's backbone and in the nucleic acids or allow for weak electrostatic interactions. This interaction would indicate that the NTD is essential to prevent the aggregation of TDP-43 due to the role in RNA binding (Wang 2018). However, the first 10 aa residues on the NTD proves to be essential to the aggregation of TDP-43 because when deleted (TDP-$43_{10\text{-}414}$) the protein has a weaker propensity to aggregate and does not develop neurodegeneration in mice (Sasaguri 2016). TDP-$43_{10\text{-}414}$ showed a loss of biological function regarding mRNA processing implying that the reason for the decrease in cytoplasmic inclusions present for this mutant is due to TDP-$43_{10\text{-}414}$ not participating in cytoplasmic shuttling during RNA transport (Sasaguri 2016). The NTD also contains two cystine residues that makes it vulnerable to oxidative stress, disulfide formation and seeding the formation of SGs and aggregates (Wang 2018; Jiang 2017).

1.3.2 RNA-Recognition Motifs (RRMs)

TDP-43 has two RRMs; RRM1 is positively charged and unstable whereas RRM2 is negatively charged and the only stable domain of the protein (Fig. 5) (Zacco 2019; Chang et al.

2013; Mompean et al. 2016). Both RRMs consist of two β-sheets between two α-helices (RRM1 = β1α1β2β3α2β4β5) (Sun 2017). The RRM1 spans from residue 106 to 177 and has been shown by electrophoretic mobility shift assay (EMSA) to have a higher affinity for DNA and RNA binding (K_d = 65.2 nM for RRM1; K_d = 379 nM for RRM2) (Buratti et al. 2001; Chang et al. 2012; Sun 2017; Flores et al. 2019). It is connected to RRM2 through a 15-aa linker that is flexible in the absence of DNA or RNA (Sun 2017; Mompean et al. 2017; Zacco 2018). RRM2 spans from residue 192-259 and contains an extra β-strand which supports protein-protein interactions and makes this domain more stable than the rest of the protein. When the first β-strand is deleted the domain loses its secondary structure and becomes highly aggregation prone (Mompean 2016; Wei et al. 2017). There is a proposed dimer interface within RRM2 that may stabilize the protein when bound to RNA (Flores et al. 2019; Prasad 2019). In addition, both RRMs each contain two cysteine residues (C173/175 and C198/244) that under oxidative stress can form disulfide bonds that cause the protein to self-aggregate indicating that it may play a role in the pathological mechanism of TDP-43 (Prasad 2019). However, Chang *et al.* 2013, showed by oxidizing each domain individually, that only the oxidation of the C173 and C175 in RRM1 caused an increase in turbidity after H_2O_2 treatment indicating the presence of higher order oligomeric species. This is because the conformational change that occurs in RRM1 during this mechanism causes a cross-linking interaction that forms an insoluble dimer that blocks nucleic acid binding (Chang 2013). This result supports a model in which RRM1 is the most essential domain for TDP-43 RNA binding interactions, and RRM2 stabilizes this interaction by binding to RNA once it is already bound to RRM1.

1.3.3 C-Terminal Domain (CTD)

Residues 270 through 414 make up the intrinsically disordered region of the protein, the CTD, consisting of two prion-like regions connected by a hydrophobic fragment (Fig. 5) (Wang 2018; French et al. 2019; Sun 2017). The CTD contains most of the ALS-linked mutations and is highly prone to aggregation due to a glycine rich sequence (331-369) that has the tendency to self-assemble through hydrogen bonds (Li 2019; Weskamp 2018; Sun 2017; Mompean 2016). There is an alpha-helical structure present that contains a prion like, Q/N rich, region that can undergo LLPS in the presence of bound nucleic acids (Li 2019; Sun 2017; Prasad 2019). Over a 24-hour period, these liquid droplets can stabilize into irreversible aggregates that are toxic to the neuron (Sun 2017). Cleavage of TDP-43 into TDP-35 is found in the cytoplasm of neurons in some patients with ALS indicating another pathway to aggregation (Kitamura et al. 2015; Neumann et al. 2006; Wei 2017; Heyburn et al. 2019). This cleavage removes the beginning of the NTD to leave residues 90-414 still present, which has a higher aggregation propensity than full-length TDP-43 (Kitamura 2015; Heyburn 2019). Even further cleavage has been carried out *in vitro* to form TDP-25 (residues 220-414), yielding a more aggregation prone protein than TDP-43 and TDP-35; indicating the effect that the other domains may have on mitigating the CTDs aggregation tendencies (Kitamura 2015; Heyburn 2019).

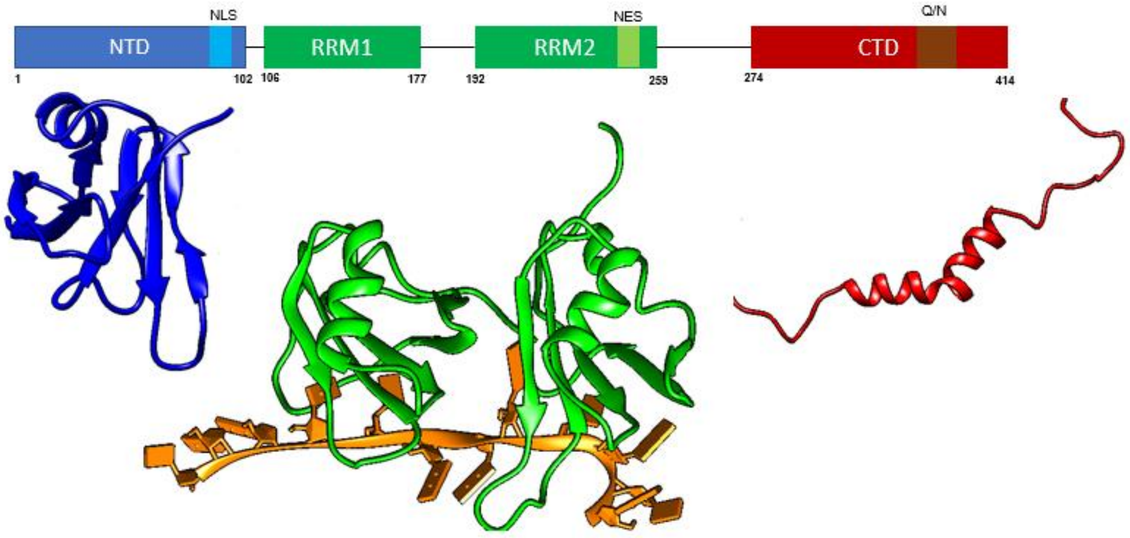


Figure 5. Crystal structure of the full-length TDP-43 bound to ssDNA. N-terminal domain (blue) comprised of the first 102 residues (PDB: 2N4P). It contains the nuclear localization signal (NLS) from residues 82 to 98. Two tandem RRM domains (green) connected through a 15-residue flexible linker is locked in conformation through binding of a ssDNA (PDB:4BS2). The nuclear exportation site (NES) spans from residue 239 to 250 in the RRM2. The unstructured domain at the C-terminal end (red) spans residues 274 to 414 (PDB:2N3X). It contains a hydrophilic Q/N sequence (residues 345-366) that facilitates LLPS and a glycine rich tail.

1.4 Oligomerization of TDP-43

1.4.1 Dimerization

In a healthy neuron, TDP-43 was shown to be in a dimer state bound to RNA in the nucleus and cytoplasm for physiological function by Shiina et al. 2010 through immunoprecipitation followed by western-blot characterization of TDP-43 in SK-N-SH cells. Using anti-TDP-43 antibodies the western blot showed a band with a molecular weight of 86-kDa, which is the expected weight of the dimer (Shiina 2016). Previously, there have been several dimerization interfaces proposed through analysis of solution NMR and crystallography; several in the NTD and one in the RRM2 domain (Jiang 2017; Wang 2018; Mompean 2017; Mompean 2016; Toro 2018). In our study two dimer interfaces were explored; one in the NTD and one in the RRM2 domain. The first of these involved serine at residue 48 (S48) in one NTD which forms hydrogen bonds with the tyrosine at residue 4 (Y4) and glutamate at residue 17 (E17) of a second NTD (Fig. 6) (Wang 2018; Toro 2018). When a phosphomimetic mutation, serine to glutamate (S48E), is made this interaction is disrupted and has been proposed to form a monomer (Wang 2018; Toro 2018). The second interface occurs between two antiparallel β4 sheets in RRM2 with the nitrogen on the aspartate at residue 247 (D247) on one TDP-43 hydrogen bonding with the oxygen on D247 on a second copy of TDP-43 (Fig. 7) (Kuo et al. 2009). This dimer interface is further stabilized by a hydrogen bond interaction between the glutamate at residue 245 (E245) and the isoleucine at residue 249 (I249) (Kuo 2009). There have been other interfaces proposed that are not further discussed in this work; namely a head-to-head dimerization between valine at residue 72 (V72) and leucine at residue 71 (L71) through hydrophobic interactions around the β7 strand (Fig. 7) (Jiang 2017). This dimerization exposes two cysteine residues that can form disulfide bonds with a second homodimer of TDP-43 to form a stable tetramer complex (Fig.8) (Jiang 2017).

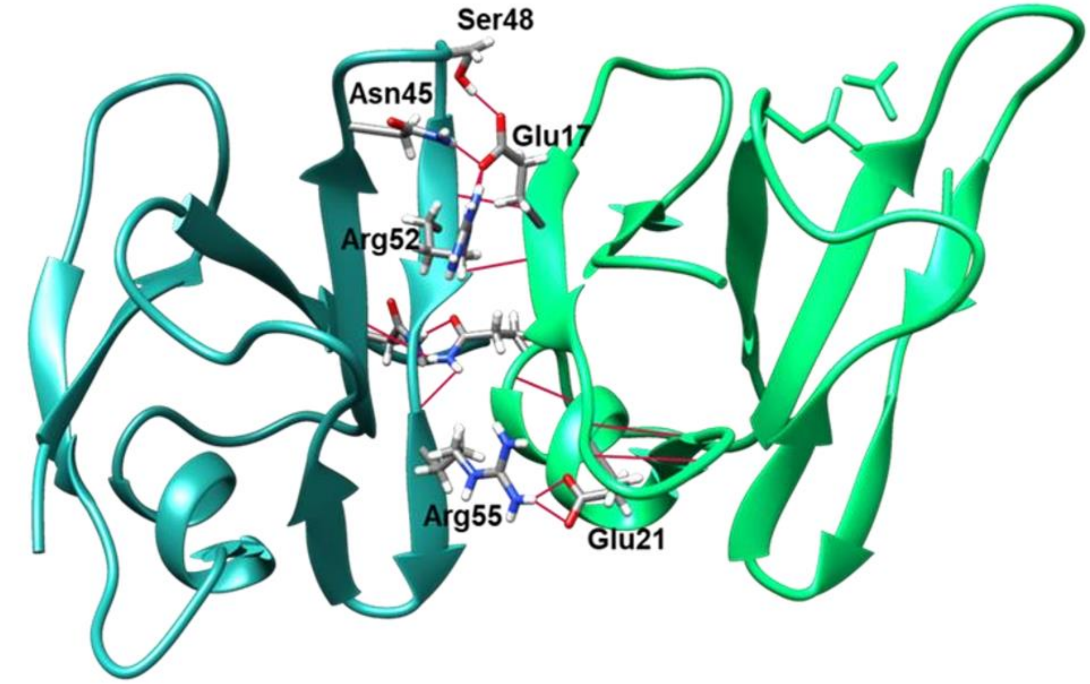


Figure 6. The head-to-tail dimerization of two NTDs with hydrogen bonds (in red) between Ser48 and Glu17 (PDB: 5MDI). Image from Toro 2018.

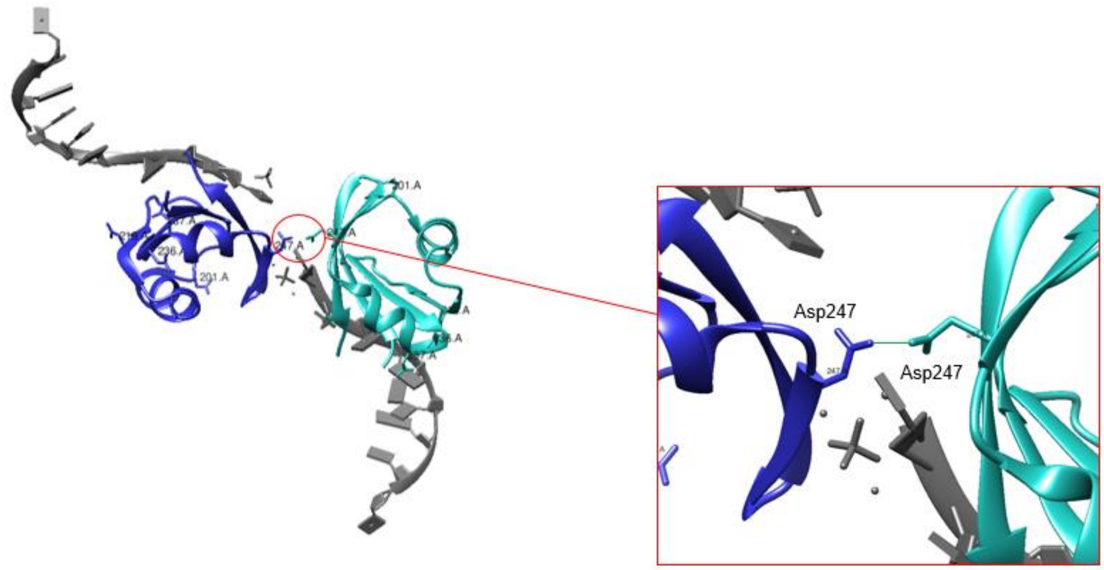


Figure 7. The head-to-head dimerization of two RRM2 domains complexed with DNA (PDB:3D2W) through hydrogen bonding (in green) between the two Asp247 residues.

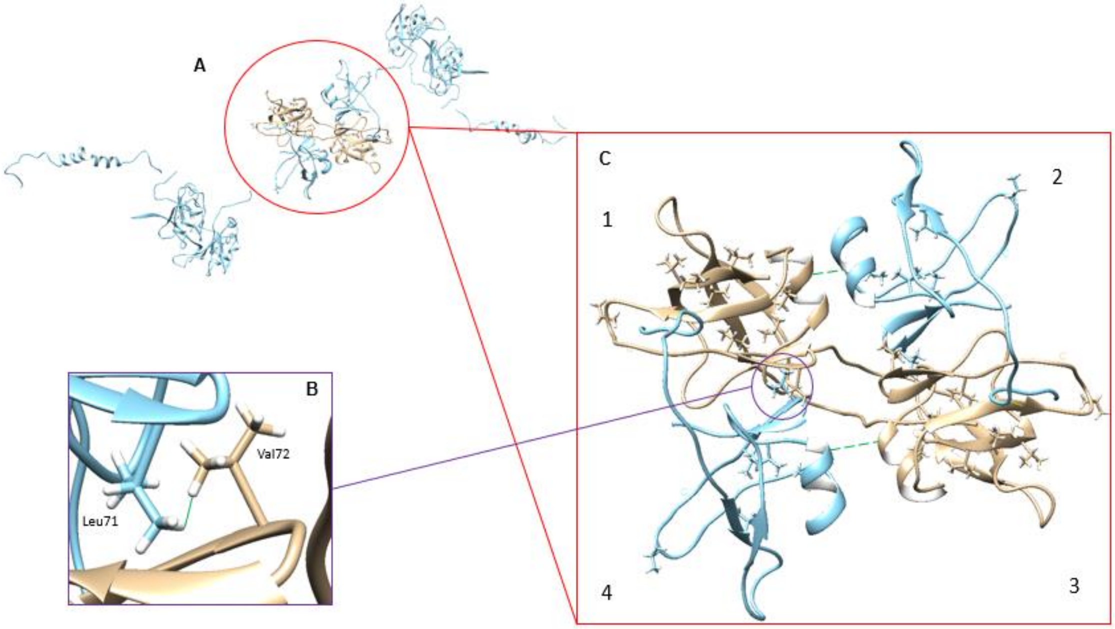


Figure 8. (A) The stable tetrameric conformation of two homodimers of TDP-43; full length depicted in blue and NTD only in brown. (B) The dimerization is happening through a head-to-head interaction of two NTDs (PDB: 2N4P) with hydrophobic interaction between Leu71 and Val72. (C) The tetrameric complex is stabilized through disulfide bonds between Cys39 and Cys50.

1.4.2 Oligomerization

The TDP-43 dimer can further self-assemble into higher order oligomers (Chang et al. 2016; Chang 2013). At a small oligomer size, the protein is still in a normal physiological state and the oligomerization is reversible, such as the formation of SG during cellular stress then dissociation once the stress has diminished (Fig. 3) (Chang 2016; Prasad 2019). This oligomerization is promoted by the NTD in a concentration dependent manner (Chang 2012; Chang 2013). The biophysical mechanism by which the protein goes from dimer to oligomer is not well known, however there have been several factors found to influence this process. Under oxidative stress, the protein begins to form into larger oligomers through LLPS, from this state they can form SGs in the cytoplasm (Chang 2016; Pillali et al. 2018; Colombrita et al. 2009). This transition is still reversible, and the oligomer can transition back to the monomer or dimer state

(Fig. 3) (Pillali 2018). This process is one way that the CTD mediates protein and RNA interactions, which regulates the RNA processes in the cells (Mann 2019; Sun 2014).

1.4.3 Toxic Aggregation

Mutations and/or prolonged stress can cause SGs to continue to assemble through liquid-solid phase separation (LSPS) into toxic irreversible aggregates that appear to be amorphous, ubiquitinated and phosphorylated (Fig. 3) (Prasad 2019; Sun 2017; Pillali 2018; Chang 2013; Mann 2019). These irreversible aggregates can also form through an independent pathway not dependent on SG formation (Sun 2017). Cleavage of TDP-43 can cause oligomerization of the NTD along with aggregation of the CTD, promoting the formation of the irreversible aggregates (Fig.3) (Sun 2017; Wei 2017; Heyburn 2019). In both cases, the aggregates that are formed can then seed the aggregation of other TDP-43 constructs in the cytoplasm (Shimonaka 2016; Sun 2017; French 2019). TDP-43's CTD's sequence resembles that of the yeast prion protein Sup35; enriched in polar and uncharged amino acids (McAlary et al. 2019). This supports that TDP-43 aggregates can seed aggregation of full-length TDP-43 through prion-like propagation as seen in many neurodegenerative disorders (McAlary 2019; Shimonaka 2016). The conformational change from α-helicies to β-sheets in the beginning of the pathway triggers the formation of amyloid like fibrils characteristic of those found in patients with ALS and FTLD (McAlary 2019).

1.5 Electrolytes and Protein Aggregation

1.5.1 Hofmeister Series

In the late 1800s Franz Hofmeister characterized a series of cations and anions by their ability to precipitate out proteins (Fig. 9) (Okur et al. 2017). The initial rationalization, that has held true for anions but not cations, was that it was just a phenomenon driven by the difference between their ability to drive order within the water molecules (Okur 2017). It was later discovered that the reasoning for the anion order is not just due to water molecule stability, but also dependent on the interaction with either the backbone or side chain of the protein (Okur 2017). When interacting with the negatively charged backbone the effects of anions follow the order of the Hofmeister series, however since the side chain is positively charged the interaction order is reversed (Okur 2017). The major driving force behind the cation sequence is the hydration of the cations. Strongly hydrated cations (e.g. calcium) interact with the water more than the weakly hydrated cations (e.g. sodium), so strongly hydrated cations have a better ability to salt out proteins (Okur 2017).

This correlation between Hofmeister series order and ability to salt out proteins was analyzed in different environmental conditions. First, it was determined that the effect of the Hofmeister series can be overridden by different steric arrangements of binding sites and ion channels (Okur 2017). In addition to the identity of the ions present in solution, the pH of the environment can play a role (Bostrom et. al. 2005). When the pH of the solution is less than the isoelectric point (pI) the direct Hofmeister series works for anions, but when the pH is greater than the pI the reverse Hofmeister series is observed (Bostrom 2005). Thus, under different conditions ions can play mixed roles in their ability to induced aggregation of a protein.

Sun et al. 2018 isolated a metastable intermediate form of TDP-43 that they used to study the magnitude that each cation in the Hofmeister series had on the induction of the aggregation of TDP-43. They found that the effect of cations on the ability to make a meta-stable intermediate of the protein aggregate followed the reverse Hofmeister series (Sun 2018). This was done at a pH 7.5, which is greater than the pI of TDP-43 (pI = 5.8) showing a similar trend to the reverse interaction that anions and cations in the Hofmeister series can have (Sun 2018; Okur 2017).

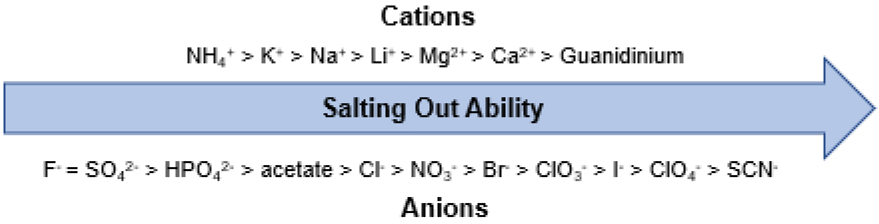
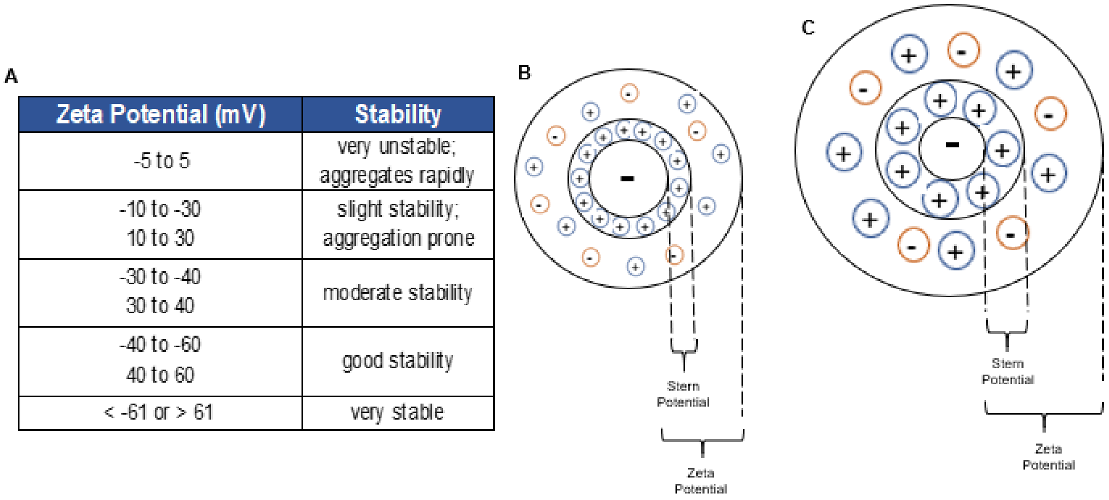


Figure 9. Hofmeister series.

1.5.2 Zeta Potential

Another relationship found between ions and their ability to induce aggregation of proteins is based on the zeta potential of the protein; the larger the zeta potential the more stable the protein (Fig. 10A) (Salgin et al. 2017). Zeta potential is the magnitude of electrostatic interactions between charged surfaces of adjacent proteins (Salgin 2017). The strength of the zeta potential (mV) is dependent on the charge and size of the ion in the solution (Salgin 2017). Conceptually, there are two layers important to consider when determining the effect of the zeta potential (Fig. 10B). The first layer is the Stern potential, which is dependent on the size of ions included in the first hydration shell of the protein (Fig. 10B). The larger the atomic radii of these ions the larger the Stern potential (Fig. 10C) (Liu et al. 2016). The next layer of the hydration sphere is the zeta potential, which is affected by each ions' charge (Liu 2016; Salgin 2017). For both cations and anions, divalent ions cause greater reduction of the zeta potential than monovalent ions. Cations induce this change by compressing the double layer bringing two charged particles closer together, increasing the likelihood for aggregation (Salgin 2017). Anions provide charge screening that causes a similar reduction in zeta potential as cations (Salgin 2017). Similar to the Hofmeister series, there is a reverse relationship between the ions ability to promote aggregation and the pH of the environment for anions and cations. Anions have a greater effect on the zeta potential at a lower pH and cations at a higher pH (Salgin 2017). Importantly, the studies included in this paper were performed at pH 7.5, which is higher than the pI of TDP-43. This means the protein is expected to have an overall negative charge and our results portray the influence of the cations on GFP-TDP-43 aggregation. Following this hypothesis, smaller cations with a higher valency have a larger effect on the aggregation of TDP-43 than larger monovalent cations.



Zeta Potential (mV)	Stability
-5 to 5	very unstable; aggregates rapidly
-10 to -30 10 to 30	slight stability; aggregation prone
-30 to -40 30 to 40	moderate stability
-40 to -60 40 to 60	good stability
< -61 or > 61	very stable

Figure 10. (A) Table of the relationship between a proteins zeta potential and stability. (B) Schematic of the double-layer surrounding the negatively charged protein background. The zeta potential encompasses the Stern potential and the diffuse ion layer surrounding it. (C) Change in zeta potential with larger cations.

1.6 Nucleic Acid Binding

1.6.1 Physiological Function

In its native state, TDP-43 is dimeric in structure and bound to mRNA in the nucleus (Sun 2017). Cytoplasmic shuttling of this complex transports the mRNA into the cytoplasm for transcription and translation, and the unbound TDP-43 will relocate back into the nucleus (Sun 2017; Colombrita 2009; Flores 2009; Prasad 2019). The formation of the homodimer is independent of RNA binding but necessary for normal biological function (Zhang et al. 2013). Since TDP-43’s main biological function is facilitating RNA metabolism, then it can be hypothesized that the dimerization may either be necessary for RNA binding or provide more efficient binding capability (Zhang 2013). It has also been shown that TDP-43 binds UG-rich RNA sequences with a higher affinity (K_d = 6.3 nm for UG-rich RNA_{12}; K_d = 1.9 µM for $NegRNA_{12}$) through biolayer interferometry (Zacco 2019; Coyne 2017; Sun 2017; Bhardwaj 2013; Chang

2012). This binding occurs through pi-stacking with aromatic rings in the five phenylalanine's in the RRM domains (Fig. 11) (Huang 2013). RRM1 contains two phenylalanine's at residue 147 and 149 (F147/F149). F147 and F149 have been demonstrated to have the highest nucleic acid binding affinity (Buratti 2001; Bhardwaj 2013; Huang 2013; Flores 2019; Zacco 2018). RRM2 contains the other three phenylalanine's at residue 194, 229 and 231 (F194/F229/F231) that only show a two-fold reduction in binding affinity when mutated (Furukawa et al. 2016). The hypothesized role of RRM2 in RNA-binding is that once the RNA is bound to RRM1 it binds RRM2 to lock it into the appropriate conformation (Furukawa 2016; Bhardwaj 2013; Flores 2019). The other proposed binding site is in the NTD in the Lys/Arg rich region and in the ubiquitin-like fold (Wang 2018; Chang 2016). This occurs through weak aromatic ring-cation interactions at a sub-millimolar affinity and lacks sequence specificity (Wang 2018; Chang 2016).

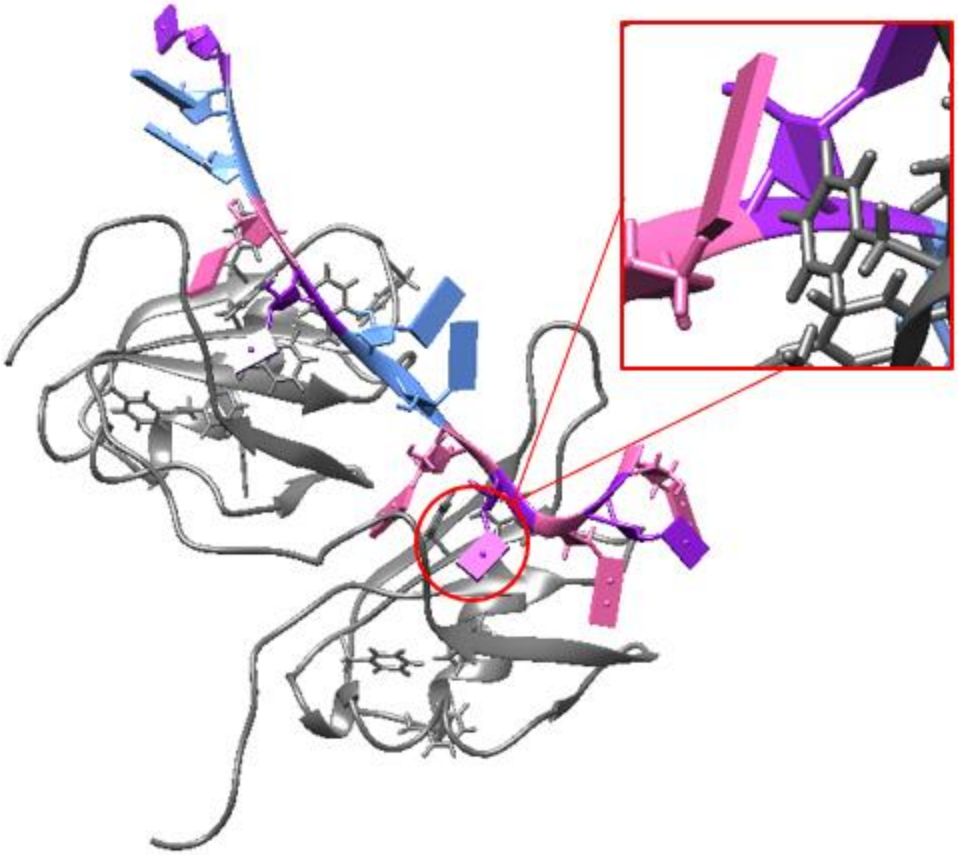


Figure 11. Tandem RRM domains (gray) bind RNA with sequence specificity. Solution NMR structure shows the phenylalanine's (stick figures) interact with ssRNA through pi-stacking with uracil and guanine nucleic acids (purple/pink respectively), locking the conformation in place (PDB:4BS2).

1.6.2 Induction/Enhancement of LLPS in the NTD and CTD

LLPS of the CTD is present in normal TDP-43 processes (Wang 2018; Li 2019). ssDNA has been shown to induce LLPS formation by the NTD and enhance LLPS in the CTD through weak aromatic ring interactions (Wang 2018). In the CTD the three tryptophan residues provide the strongest affinity for interaction with ssDNA. The residue at 334 (W334) being the most important (Wang 2018). Although LLPS to form SGs is a physiological function of TDP-43, it can also be argued that RNA may influence the cytotoxic aggregation of TDP-43. The formation of SGs is necessary to help reduce cellular stress, however it can also promote the formation of toxic aggregates if it does not dissociate once the stress has been relieved. RNA's function in this formation is apparent by the presence of RNA in these SGs (McGurk 2018). There are contradicting studies that state whether RNA binding is indeed beneficial to the cell or if certain strands of RNA can cause the toxic aggregates characteristic of those found in patients with ALS (Zacco 2019).

1.6.3 Inhibition of the Formation of Larger Order Oligomers

Binding of polyanionic RNA to TDP-43 has been shown to inhibit or mitigate the aggregation of TDP-43 from forming the cytoplasmic inclusions found in the neurons of patients with ALS *in vitro* (Sun 2014; Huang 2013). When the DNA oligomer $(TG)_n$ is present in a 2:1 DNA to RNA ratio max inhibition in observed, with n=12 being the optimal length (Sun 2014; Huang 2013). The specificity is seen when observing the binding of UG-rich RNA sequences versus lower-affinity RNA oligomers. The strongly bound UG RNA showed an inhibition of aggregation, whereas the lower-affinity RNA oligomers showed no effect or a slight increase in aggregation propensity (Zacco 2019). The potential of RNA-binding to inhibit aggregation of TDP-43 and the prevalence of RNA therapeutics emerging in modern day medicine makes the aggregation mechanism and inhibition a targeted field of study for ALS research.

1.7 Objective

Building on what has already been determined about the structure of TDP-43, our first aim was to characterize the proposed dimers in the NTD and RRM2 domain. Mutant constructs of GFP-TDP-43 that disrupt the dimer interface in either the NTD, RRM2 or both interfaces will be characterized using size exclusion chromatography and dynamic light scattering to see if disruption of either interface could prevent the dimerization of GFP-TDP-43 as indicated by the presence a smaller species than the dimer of the WT-GFP-TDP-43.

The second aim of our study was to determine the salt-dependent aggregation kinetic trends of GFP-TDP-43. A metastable intermediate form of the protein will be obtained by using a denaturing purification previously reported by Sun et al. 2017. Using monovalent and divalent cations, aggregation will be induced and monitored using turbidity assays. Our hypothesis is that the rate of aggregation will be dependent on the valency and atomic radius of the cation, rather than solely on the Hofmeister series as previously reported. Using the turbidity aggregation assay, we will also look to show that TG12 can reduce the aggregation of GFP-TDP-43 and that AC12 has no effect. This would confirm the sequence specificity of the binding interaction between nucleic acids and TDP-43 as well as indicating that nucleic acids can have a reducing effect on toxic aggregation.

CHAPTER 2: METHODS & MATERIALS

2.1 Site-Directed Mutagenesis (SDM)

Q5 Site-Directed Mutageneis Kit was used to mutate the aspartate to leucine at residue 247 in the pET-GFP-TEV-TDP43 wild-type and S48E variant (LIC cloned by previous graduate students). Forward and reverse primers were designed using NEB Base Change designer (Table S1). The plasmid was PCR amplified with Q5 phusion and run on a 1% agarose gel for 1.5 hours at 130 eV to verify linearization of the plasmid. The PCR product was treated with 2X KLD (Kinase, Ligase, Ddpn1) reaction buffer and 10X KLD enzyme mix and incubated overnight at room temperature. The KLD mix was transformed into super-competent DH5α cells. Cells were plated on a kanamycin resistant plates and incubated overnight at 37^{o}C. Mutations were verified through DNA sequencing by QuintaroBio.

2.2 Expression/Purification in Native Conditions of GFP-TDP-43

eGFP-TEV-TDP43 (wild-type and mutant variants) in a 1GFP vector was transformed into BL21 cells for expression (Fig. S1). Cultures were grown in 50 mL LB media with kanamycin resistance overnight. The 50 mL culture seeded 2 L of Miller LB media with kanamycin resistance and grown to OD600 of 0.4 at 37^{o}C. They were induced with IPTG, final concentration of 1 mM, and incubated at 19^{o}C for 6 hours. After induction, the cultures were spun down at 4 ^{o}C, 3000 rpm for 30 minutes to collect the cells. Each pellet was resuspended in 30 mL lysis buffer [300 mM NaCl, 1 mM $MgCl_{2,}$ 40 mM HEPES (pH 7.5), 50 mM Imidazole, 5 mM β-Me, 10% Glycerol] and lysed on ice with 1% CHAPS, 0.25 mg/mL lysozyme, and 20 µg/mL Dnase for 1 hour with gentle stirring. Lysate was centrifuged at 4 ^{o}C and 13000 rpm for 1 hour 15 min to pellet the cellular debris. The supernatant was filtered through 0.22 µM bottle top filters.

Purification was done by Ni-NTA affinity chromatography using the NGC^{TM} chromatography system at 4 ^{o}C. The sample was eluted through an isocratic elution with the lysis buffer and elution buffer [300 mM NaCl, 1 mM $MgCl_{2,}$ 40 mM HEPES (pH 7.5), 500 mM Imidazole, 5 mM β-Me, 10% Glycerol]. Peak absorbance elution was filtered through 0.22 μM syringe tip filters and further purified with size exclusion chromatography (SEC) eluting with SEC buffer [300 mM NaCl, 5 mM DTT, 40 mM HEPES (pH 7.5)]. Protein was collected off the column, filtered through 0.22 μM syringe tip filters and stored at -80 ^{o}C.

2.3 Expression/Purification of GFP-TDP43 Under Denaturing Conditions

2.3.1 GFP-TDP-43

eGFP-TEV-TDP43 (wild-type and mutant variants) in a 1GFP vector was transformed into BL21 cells for expression. Cultures were grown in 50 mL LB media with kanamycin resistance for 5 hours. The 50 mL culture seeded 1 L LB cultures with 1 mL 1000x kanamycin and grown to OD600 of 0.4-0.6 at 37^{o}C. They were then induced with 1 mL 1M IPTG for a final concentration of 1 mM IPTG and incubated at 19^{o}C for 12 hours. After induction, the cultures were spun down at 4 ^{o}C, 2700 rpm for 30 minutes to collect the cells. Each pellet was resuspended in 25 mL lysis buffer [8M urea, 500 mM KCl, 20 mM HEPES (pH 7.4), 20 mM imidazole, 2 mM β-Me, 10% glycerol] and lysed by sonication for 5 minutes (5 secs on/off). Lysate was centrifuged at 4 ^{o}C, 13600 rpm for 1 hour to pellet the cellular debris. The supernatant was filtered through 0.22 μM bottle top filters. Purification was done by Ni-NTA affinity chromatography using the AKTA chromatography system at room temperature. The sample was eluted through an isocratic elution with the lysis buffer and elution buffer [8M urea, 500 mM KCl, 20 mM HEPES (pH 7.4), 250 mM imidazole, 2 mM β-Me, 10% glycerol]. Highest absorbance peak was collected and stored at -80 ^{o}C. When needed for experiments, 3 mL aliquots were thawed on ice and placed in dialysis buffer [8M urea, 10 mM HEPES (pH 7.4)] overnight to remove the salts, and then switched to a another dialysis buffer [10 mM HEPES (pH 7.4)] overnight to remove the urea. Remaining protein is either stored at -80 ^{o}C or used for aggregation studies.

2.3.2 pDest17-TDP-43

pDest17-TDP43 was transformed into BL21 cells for expression (Fig. S2). Cultures were grown in 50 mL LB media with ampicillin resistance overnight. The 50 mL culture seeded 1L of LB media with ampicillin resistance and grown to OD600 of 0.4-0.6 at 37°C. They were then induced with 1 mL 1M IPTG for a final concentration of 1 mM IPTG and incubated at 19°C for 12 hours. After induction, the cultures were spun down at 4 °C, 2700 rpm for 30 minutes to collect the cells. Each pellet was resuspended in 25 mL lysis buffer [8M urea, 500 mM KCl, 20 mM HEPES (pH 7.4), 20 mM imidazole, 2 mM β-Me, 10% glycerol] and lysed by sonication for 5 minutes (5 secs on/off). Lysate was centrifuged at 4 °C, 13600 rpm for 1 hour to pellet the cellular debris. The supernatant was filtered through 0.22 μM bottle top filters. Purification was done by Ni-NTA affinity chromatography using the AKTA chromatography system at room temperature. The sample was eluted through an isocratic elution with the lysis buffer and elution buffer [8M urea, 500 mM KCl, 20 mM HEPES (pH 7.4), 250 mM imidazole, 2 mM β-Me, 10% glycerol]. Highest absorbance peak was collected and stored at -80 °C. When needed for experiments, 3 mL aliquots were thawed on ice and placed in dialysis buffer [8M urea, 10 mM HEPES (pH 7.4)] overnight to remove the salts, and then switched to a another dialysis buffer [10 mM HEPES (pH 7.4)] overnight to remove the urea. Remaining protein is either stored at -80 °C or used for aggregation studies.

2.3.3 Green Fluorescent Protein (GFP)

GFP in a pBADk vector was transformed into BL21 cells for expression. Cultures were grown in 10 mL LB media with ampicillin resistance overnight. The 10 mL culture seeded 1 L of LB media with ampicillin resistance and grown to OD600 of 0.65 at 37°C. They were induced with 0.2% arabinose and incubated at 20°C for 16 hours. After induction, the cultures were spun down at 4 °C, 4000 rpm for 20 minutes to collect the cells. Each pellet was resuspended in 30 mL lysis buffer [8M Urea, 400 mM NaCl, 20 mM Tris buffer (pH 8.0), 20 mM Imidazole, 10 mM β-Me] and lysed by sonication for 15 min (10 seconds on/ 1 min off). Lysate was centrifuged at 4 °C and 14000 rpm for 1 hour 15 min to pellet the cellular debris. The supernatant was filtered through

0.22 μM bottle top filters. Purification was done by Ni-NTA affinity chromatography using the AKTA chromatography system at room temperature. The sample was eluted through an isocratic elution with the lysis buffer and elution buffer [8M Urea, 400 mM NaCl, 20 mM Tris buffer (pH 8.0), 500 mM Imidazole, 10 mM β-Me]. Peak absorbance eluent was filtered through 0.22 μM syringe tip filters and was stored at -80 °C. For aggregation assays, GFP samples were thawed on ice and placed in dialysis buffer [8M urea, 10 mM HEPES (pH 7.4)] overnight to remove the salts, and then switched to a another dialysis buffer [10 mM HEPES (pH 7.4)] overnight to remove the urea. Remaining protein is either stored at -80 °C or used for aggregation studies.

2.4 Aggregation Assays

2.4.1 Dynamic Light Scattering (DLS)

Dynamic light scattering (DLS) measurements were taken to obtain the hydrodynamic radius of each protein sample after size exclusion chromatography and aggregation assays on a DynaPro® Plate Reader II. Protein samples (5-25 μM) were filtered through 0.22 μM PES syringe tip filters and aliquoted into 96-well black half-area plates to a volume of 100 μL in either SEC buffer [300 mM NaCl, 40 mM HEPEs (pH 7.5), 5 mM DTT] or 10 mM HEPES (pH 7.5). 5 measurements of each well were taken in succession over 5 seconds for each measurement and the light intensity is analyzed using DYNAMICS software. The analysis derives the diffusion coefficient from the autocorrelation curves and then uses Stokes-Einstein equation to obtain the hydrodynamic radius based on a spherical model (Eqn. 1). The output gives a histogram showing the polydispersity of each sample plotting the radius against the percent mass. Each sample was done in triplicate.

$$D = \frac{Tk_B}{6\pi n R_H}$$

$$D = Translational\ diffusion\ coefficent\ \left[{}^{m^2}/_{s}\right]$$

$$T = Temperature\ [K]$$

$$k_B = Boltzman\ constant\ \left[{}^{m^2kg}/_{Ks^2}\right]$$

$$n = viscosity\ [Pa * s]$$

$$R_H = Hydrodynamic\ radius\ [m]$$

Equation 1. Stokes-Einstein Equation

2.4.2 Turbidity Measurements

Protein samples were diluted to ~20 µM in 10 mM HEPES (pH7.5) in a 96-well black half-area plates in 100 µL aliquots. DLS was used to obtain the starting hydrodynamic radius of each protein construct. Turbidity was analyzed over 30 minutes with absorbance measured at 600 nm every 10 minutes using Tecan Infinite® M1000 plate reader. Samples were then spiked with 10 µL of varying electrolytes to have a final concentration of 10, 50, 150 or 300 mM. Turbidity was analyzed again over 3600 min with absorbance measured at 600 nm every 10 minutes. After 3600 min, the protein sample size was characterized again using DLS to obtain the hydrodynamic radius of the aggregated species.

2.4.3 ZetaView® Analysis

1 mL proteinaceous fractions containing ~25 µM GFP-TDP-43 in 10 mM HEPES (pH 7.5) were incubated with either $CaCl_2$ (final concentration = 10 mM) or NaCl (final concentration = 150 mM) overnight. Through serial dilution fractions were brought down to a concentration of 100000x less than initial. The eGFP fluorescent channel (488 nm) on ZetaView® was used for sample analysis to minimize buffer noise. Video images were obtained from 11 positions in each sample and nanoparticle tracking analysis (NTA) was used to obtain the average diameter (nm) for the particles in solution and the zeta potential (mV) of the sample.

2.4.4 DNA Addition

DNA oligomers TG12 and AC12 were purchased from Integrated DNA Technologies and resuspended to a final volume of 1 mM using molecular biology grade water (Table S2). GFP-TDP43 protein samples were diluted to ~20 µM in 10 mM HEPES (pH7.5) in a 96-well black bottom plates in 100 µL aliquots. Samples containing TG12 or AC12 were added to the wells prior to dilution in a 1:1 1:5 or 1:10 ratio of protein to DNA respectively and then 10 mM HEPES (pH7.5) was added to reach a final volume of 100 µL. DLS was used to obtain the starting hydrodynamic radius of each protein construct. Turbidity was analyzed over 30 minutes with absorbance measured at 600 nm every 10 minutes. Samples were then spiked with 10 µL of

NaCl (final concentration = 150 mM) or buffer. Turbidity was then analyzed again over 3600 min with absorbance measured at 600 nm every 10 minutes. After 3600 min, the protein sample size was characterized again using DLS to obtain the final hydrodynamic radius of the aggregated species.

2.5 SDS-PAGE Protein Gel

20 µL proteinaceous samples are mixed with 10 µL 2x Laemmli Sample Buffer containing β-Me. Fractions were incubated at 95^{o}C for 7 minutes. Samples were run on a stain-free gel (Mini-PROTEAN® TGX Stain-Free^{TM} Precast Gels) in a 1x Tris/Glycine/SDS buffer [2.5 mM Tris, 19.2 mM glycine, 0.01% SDS (pH 8.3)] at 300eV for 20 minutes along with 10-250 kDa protein ladder (Precision Plus Protein^{TM} Unstained Protein Standards). The protein gel was incubated with Coomassie blue stain for 20 min and then washed several times with destain [10% acetic acid, 20% methanol, 70% MilliQ H_2O] with gentle rocking for optimal imaging.

CHAPTER 3: RESULTS

3.1 Characterization of the Dimer Interface of TDP-43

Wild-type TDP-43 with an N-terminal green fluorescent protein (GFP) tag, used to increase solubility, was obtained through ligation independent cloning (LIC-cloned) by a previous graduate student (Toro 2018). Using the wild-type (WT) TDP-43 in the eGFP vector three mutant variants were obtained through site-directed mutagenesis; S48E, D247L and S48E/D247L. These mutations are predicted to disrupt the N-terminal head-to-tail dimerization, RRM2-RRM2 dimerization or both dimer interfaces respectively by breaking the hydrogen bond interactions (Kuo 2009; Wang 2018).

To examine the effect of these mutations on GFP-TDP-43 dimerization, each construct was expressed and purified using a previously optimized protocol that included preparatory scale size-exclusion chromatography (SEC) and was verified by SDS=PAGE (Fig. 12) The predicted molecular weight of GFP-TDP-43 is 71.3 kDa and in these experiments full length GFP-TDP-43 appears as a band just below a 75 kDa marker (Fig. 12). Also present to varying degrees in our experiment are smaller fragments of purified protein. Excision of these products followed by peptidic digest and mass spectrometry is being pursued in order to fully determine their identity. If this result shows a cleaved TDP-43 fragment with a molecular weight indicative of a 63 kDa species, then it could be indicative of the cleavage that has been shown to seed aggregation of TDP-43 (Sun 2014). This would be an interesting result since Fig 12 shows that the D247L mutant primarily forms this smaller species, which would indicate that it may be a more toxic mutant than the other constructs studied.

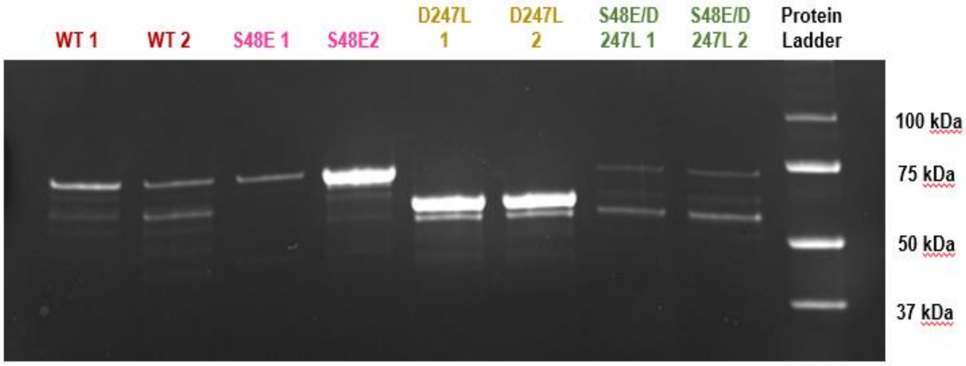


Figure 12. SDS-PAGE gel of purified GFP-TDP-43 collected off the size-exclusion column for the wild-type (WT), single-point mutants (S48E and D247L) and double-point mutant (S48E/D247L) of GFP-TDP-43 with 1 kb unstained protein ladder. Colors correspond to SEC chromatogram in Figure 13. 1 = Aggregate peak, 2 = Dimer/Monomer peak. Predicted molecular weight is 71.6 kDa for GFP-TDP-43. Second band present in lanes 1, 2, 7 and 8 could indicate presence of cleaved protein.

During each purification, the GFP-TDP-43 is first purified using a Nickel affinity column, which recovers several different size species of GFP-TDP43 constructs present in E. coli lysates. This mixture is then further purified using SEC to analyze and separate different size species. In these experiments large oligomers elute out first and smaller oligomers or monomeric species come off later. The WT, S48E and D247L constructs each yielded two species off the size exclusion column that, using sizing standards, appeared to be a larger aggregated species and a dimeric species (Fig. 13 and Fig S4). Interestingly, the construct with the double mutant that would disrupt both dimer interfaces, yielded a slightly smaller species for the second peak off the sizing column (Fig. 13). Each constructs chromatogram was plotted against sizing standards run with the same elution protocol to give an estimate of the molecular weight of the species (Fig. S3). The breadth of each peak shows the room for error for these estimates, so the molecular weight is only a rough estimate and needs additional analysis to confirm the size of the species.

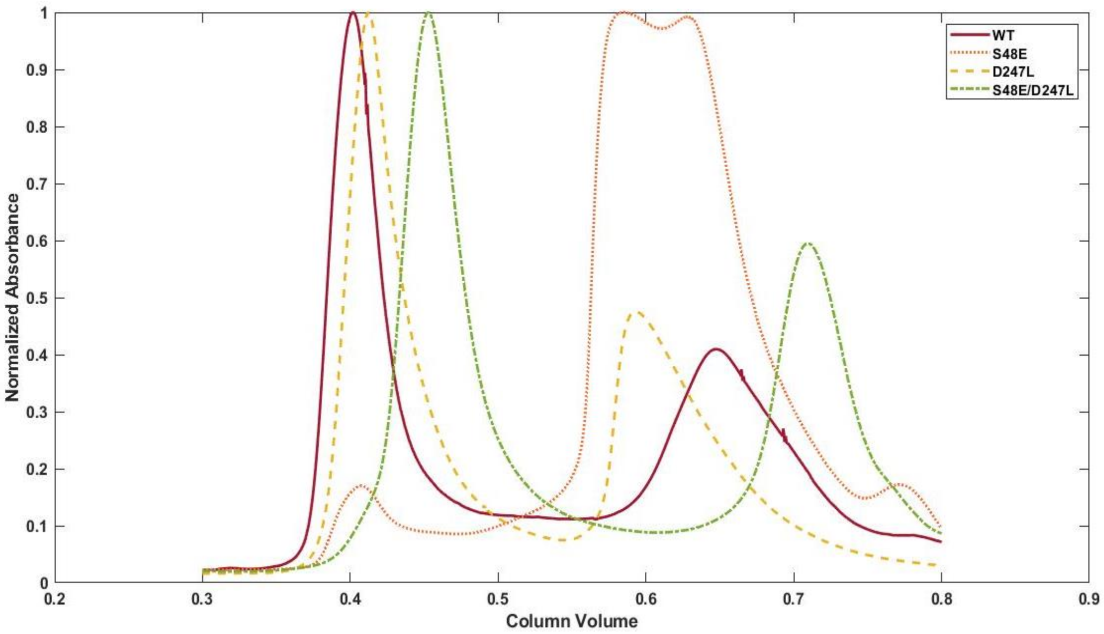


Figure 13. Chromatogram from size exclusion chromatography showing the two peaks from the purification each construct: Wild-type (WT) (solid red), single-point mutant S48E (pink dots), single-point mutant D247L (yellow dash) and double-point mutant S48E/D247L (green dot/dash). Left most peak is the aggregated species approximately 640 kDa based on sizing standards. The right most peaks of WT, S48E and D247L constructs run closer to 158 kDA, whereas the right most peak of S48E/D247L lands around 44 kDa.

In order to further characterize the different TDP-43 species DLS confirmed that the double mutant yielded a smaller species than the WT or either single point mutant (Fig. 14). DLS assumes spherical particles in solution move randomly in accordance with the principle of Brownian motion, which allows the system to use Stokes-Einstein equation to calculate the hydrodynamic radius (Eqn. 1). By measuring the translational diffusion coefficient from the intensity of light scattered in solution, it obtains the hydrodynamic radius of each particle in the sample. The molar mass is then approximated using DYNAMICS software which derives the hydrodynamic radius based on the Brownian motion of the particles in solution, however, since the mass is determined based on the assumption that the particles are spherical the estimate is not accurate for the actual species weight but is useful for a relative comparison of the different size oligomers of each species. Using this method, the first peak off the SEC for each construct had a hydrodynamic radius of approximately 104.2 nm (n=5, SD=12.8) which is indicative of an aggregated species. The second peak for WT, S48E and D247L had hydrodynamic radii of approximately 5.7 nm (n=3, SD=0.24), 5.2 nm (n=3, SD=0.14) and 5.7 nm (n=3, SD=0.33), respectively (Fig. 14). However, the S48E/D247L construct was approximately 4.5 nm (n=3, SD=0.29), indicating the possible presence of a monomeric species (Fig. 14). The difference in the elution profile and hydrodynamic radius size of the double mutant, S48E/D247L, versus the single point mutants and wild-type GFP-TDP-43 supports the hypothesis that the double mutant is disrupting the formation of a dimer through the disruption of two separate interfaces, whereas the single point mutants still have the ability to form a dimer through the undisrupted interface.

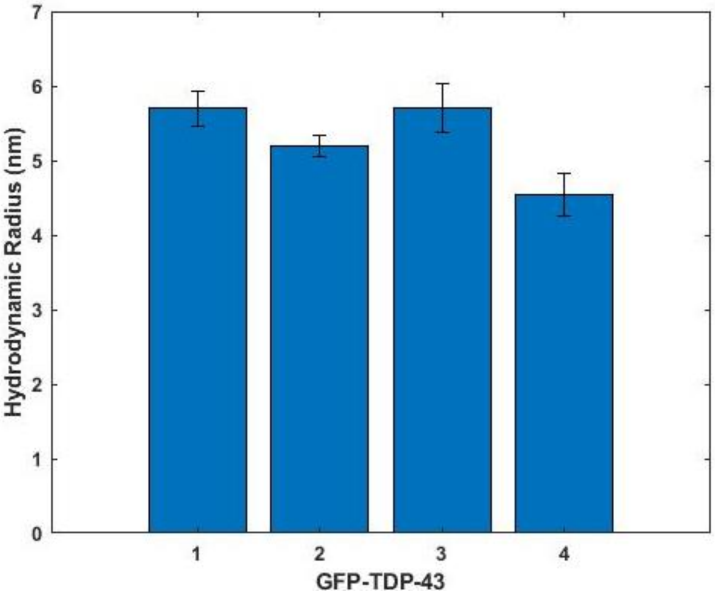


Figure 14. Average hydrodynamic radius of the double mutant, both single mutants and wild-type GFP-TDP-43 determined through dynamic light scattering of second peak off the size-exclusion column. 1=WT, 2=S48E, 3=D247L, 4=S48E/D247L.

To analyze aggregation of the GFP-TDP-43 dimer into higher order oligomers, time-course analysis of DLS measurements of the S48E dimer were taken over the course of 12 hours. These data showed an increase in aggregation over time, indicating slow aggregation at room temperature (rate = 0.021 nm/min) (Fig. 15). This trend shows GFP-TDP-43's propensity to aggregate at room temperature, which could pose a problem for aggregation assays.

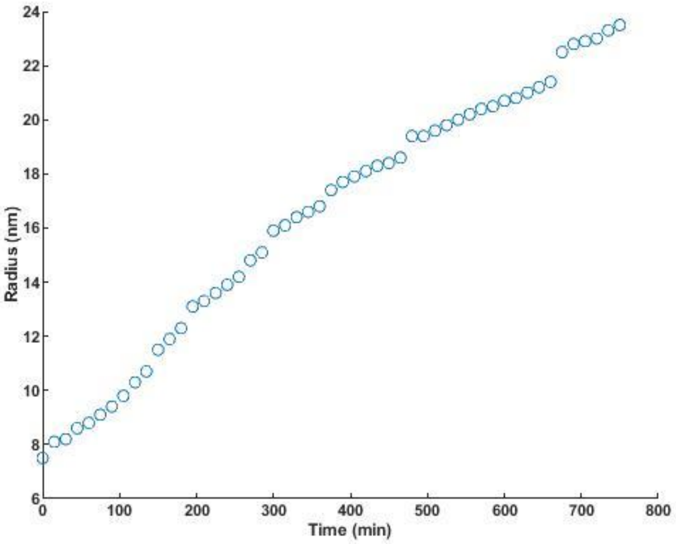


Figure 15. Change in radius of the dimer of GFP-TDP43 S48E [10 μM] at room temperature over 800 minutes shows slow aggregation over time. Hydrodynamic radius derived from dynamic light scattering analysis using DYNAMICS software. Starting radius is 7.5 nm and the final radius is 23.5 nm, which has not reached full aggregation (~100 nm).

3.2 Purification and Characterization of a Meta-Stable Intermediate Species of TDP-43

Since the GFP-TDP-43 dimer obtained from the purification as described above will slowly aggregate over time (Fig. 15), we needed to be able to isolate a metastable form of the protein in order to study the aggregation kinetics. This was achieved through a denaturing purification and then buffer exchanging the protein into 10 mM HEPES (pH 7.5), as reported by Sun *et al.* 2019, which is stable at room temperature. The final species is characterized using DLS and SDS-PAGE. The SDS-PAGE gel showed a band with high intensity landing at approximately the same spot as the 75 kDa band of the protein ladder indicating that we are working with GFP-TDP43 (Fig. 16A). DLS data showed that this metastable intermediate had approximately 7.1 nm hydrodynamic radius (Fig. 16B) indicating a low order oligomeric state.

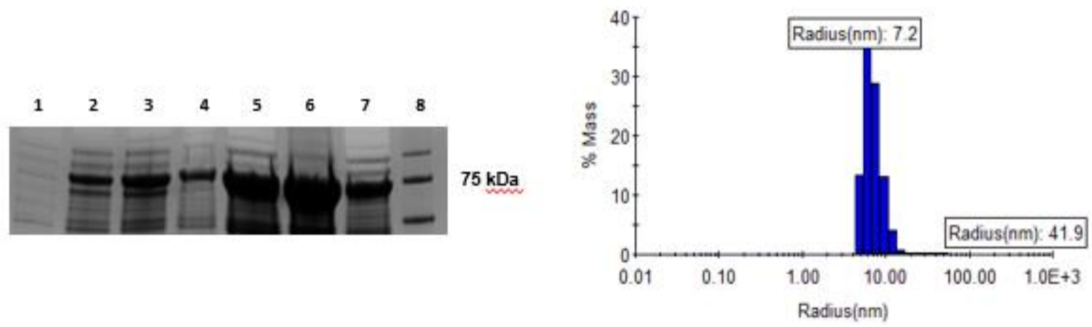


Figure 16. (A) SDS-PAGE gel of GFP-TDP43 WT of the denaturing purification. Lanes are: 1 – uninduced; 2 – post-induction; 3 – lysate; 4 – pellet post-lysing; 5 – Ni column elution; 6 – post-first dialysis; 7 – post-second dialysis (metastable intermediate species); 8 – protein ladder. (B) DLS measurement of the metastable intermediate (~20 µM) using DYNAMICS software shows the presence of a species with a hydrodynamic radius of 7.2 nm.

3.3 Electrolytic Ions can Induce Aggregation of the Meta-Stable Intermediate Form of GFP-TDP-43

Sun *et al.* 2019 reported that the meta-stable intermediate in 10 mM HEPES (pH 7.5) at room temperature will begin to aggregate when induced with electrolytes. To test this hypothesis, GFP-TDP-43 and each of its mutant variants were diluted to ~20-25 µM using 10 mM HEPES (pH 7.5) and then aliquoted into a 96 well plate with an initial volume 100 µL. DLS measurements were taken of the initial species (Fig. 16B) and then subjected to turbidity measurements using a plate reader measuring the absorbance at 600 nm every 10 min to establish an initial baseline, which remained stable indicating the lack of aggregation (Fig. 17). To each well, 10 µL of either salt solution (final concentration equaling 10-300 mM) or buffer (10 mM HEPES, pH 7.5) was added and the turbidity was measured again over 6 hours taking absorbance measurements at 600 nm every 10 min. Over time, if the absorbance measurement increases it is indicating that the solution is becoming more turbid, likely due to the aggregation of the protein in solution. This correlation allows us to establish the relative aggregation that each electrolytic cation can induce on the meta-stable GFP-TDP43. The results showed that at low concentrations the divalent cations (Mg^{2+}, Ca^{2+}, Zn^{2+}) were able to induce aggregation (Fig. 18A), whereas the monovalent cations (K^+, Na^+, NH_4^+, Li^+) showed little to no change in absorbance. All measurements were taken using the same anion, Cl^-, to ensure the effect was due to the cation. The same results were replicated with the S48E mutant (Fig. 18B) This difference between the divalent and monovalent cations shows the effect that the charge has on the salt-induced aggregation kinetics of GFP-TDP43.

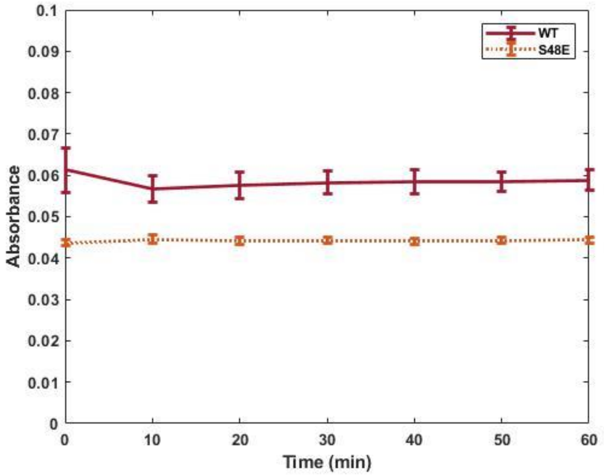


Figure 17. Average absorbance measurements of the meta-stable intermediate form of GFP-TDP-43 WT [19.66 μM] (red) and S48E [22.43 μM] (orange) shows no significant change in turbidity at room temperature over 60 minutes. Samples done in triplicate and error bars represent the standard deviation.

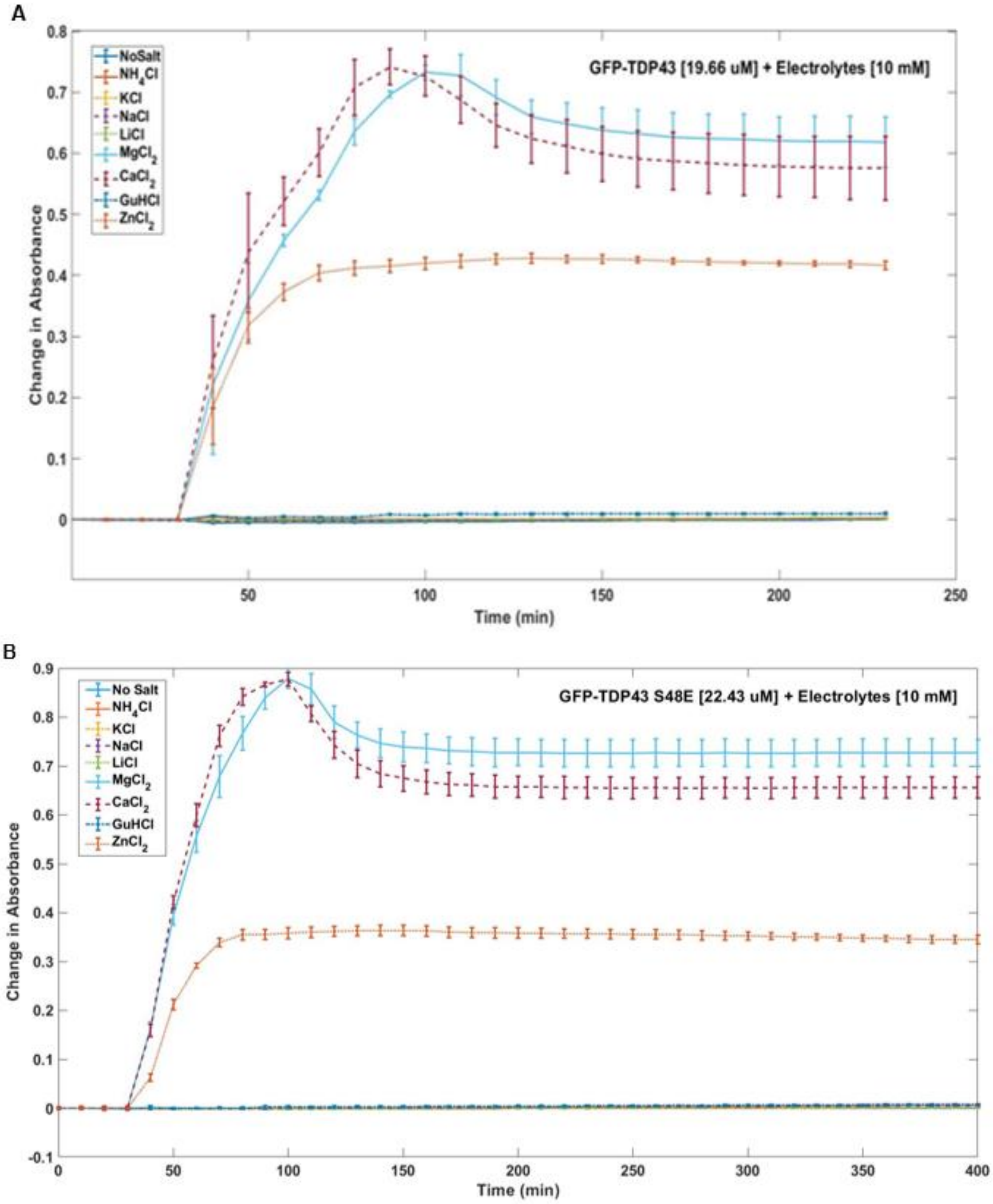


Figure 18. Absorbance measurements of 100 µL samples of (A) GFP-TDP43 [21.63 µM] and (B) GFP-TDP43 S48E mutant [24.67 µM] at 600 nm were taken every 10 minutes over the course of 240 minutes. After the first 30 minutes each sample was spiked with 10 µL of specified cations with a final concentration of 10 mM in the 110 µL each sample and measurements continued to assess the aggregation induced based on the increase in turbidity measurements. Each sample was done in triplicate and the averages are depicted in the plot with error bars indicating the standard deviation from the mean.

As the concentration increases for each cation there is an increase in absorbance readings indicating that the aggregation kinetics is concentration dependent (Fig. 19A). This concentration dependence was also seen in time lapse experiments using DLS, measuring the hydrodynamic radius approximately every 15 minutes for 300 minutes (Fig. 19B). The increase in concentration is shown to increase the aggregate size in both experiments with a similar rate of aggregation (Fig. 19C/D).

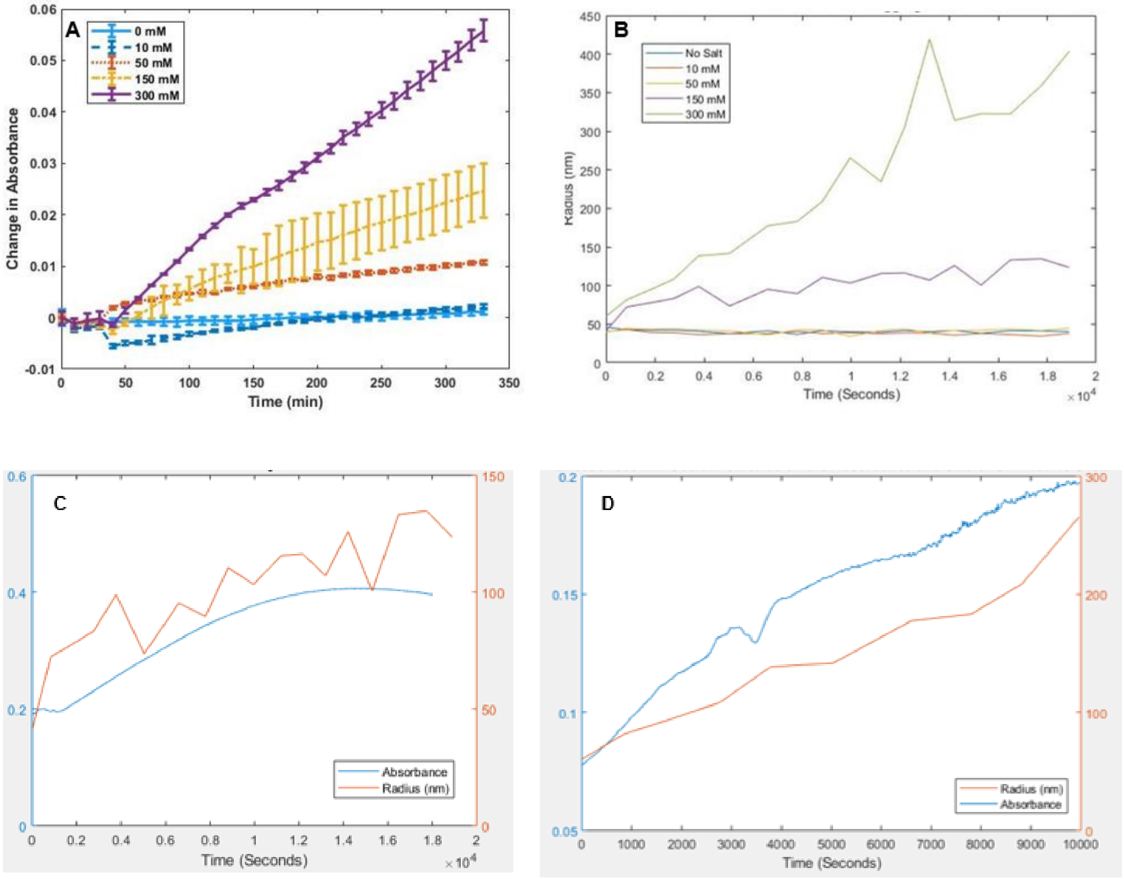


Figure 19. (A) Absorbance measurements of 100 µL samples of GFP-TDP43 [21.63 µM] at 600 nm were taken every 10 minutes over the course of 240 minutes. After the first 30 minutes each sample was spiked with 10 µL of NaCl at varying concentrations [10mM, 50mM, 150mM, 300mM]. Each concentration was done in triplicate and averaged with the standard deviation indicated by the error bars. (B) DLS measurements of the average hydrodynamic radius for each concentration after induced with 10 µL of NaCl at varying concentrations [10mM, 50mM, 150mM, 300mM]. Measurements were taken every 15 minutes for 300 minutes. When the turbidity measurements (blue) and DLS measurements (orange) for the samples containing (C) 150 mM NaCl and (D) 300 mM NaCl were plotted on the same graph the similarity between the perceived rate of aggregation is seen.

At 150 mM salt concentrations the difference in each cations' propensity to cause aggregation is apparent and we were able to order the cations from potassium causing the least amount of aggregation and calcium causing the most aggregation: *Gu*$^+$ *(150 mM)* > Ca^{2+} > Mg^{2+} > Zn^{2+} > *Gu*$^+$ *(10 mM)* > NH_4^+ > Li$^+$ > Na$^+$ > K$^+$ (Fig. 20). The cations represented here are all part of the Hofmeister series, except for zinc, but there is a slight discrepancy from the series. The ammonium ion is out of place from the series. If the Hofmeister series explained this trend in aggregation then ammonium chloride should cause slower aggregation than potassium chloride (Fig. 9). It has been seen in some previous studies that the ammonium ion can deviate from the expected order of the Hofmeister series due to its unique interaction with proteins (Nucci et al. 2009). Through vibrational coupling, the nitrogen has been shown to favor specific hydrogen bond interactions with proteins rather than disrupting the bulk water network through hydrogen bonds (Nucci 2009). Also, interestingly, the guanidinium ion shows the appropriate placement at 150 mM in the Hofmeister series trend, but at 10 mM concentration it shows slightly more aggregation than the other monovalent cations, but significantly less aggregation induction than the divalent cations (Fig. 18).

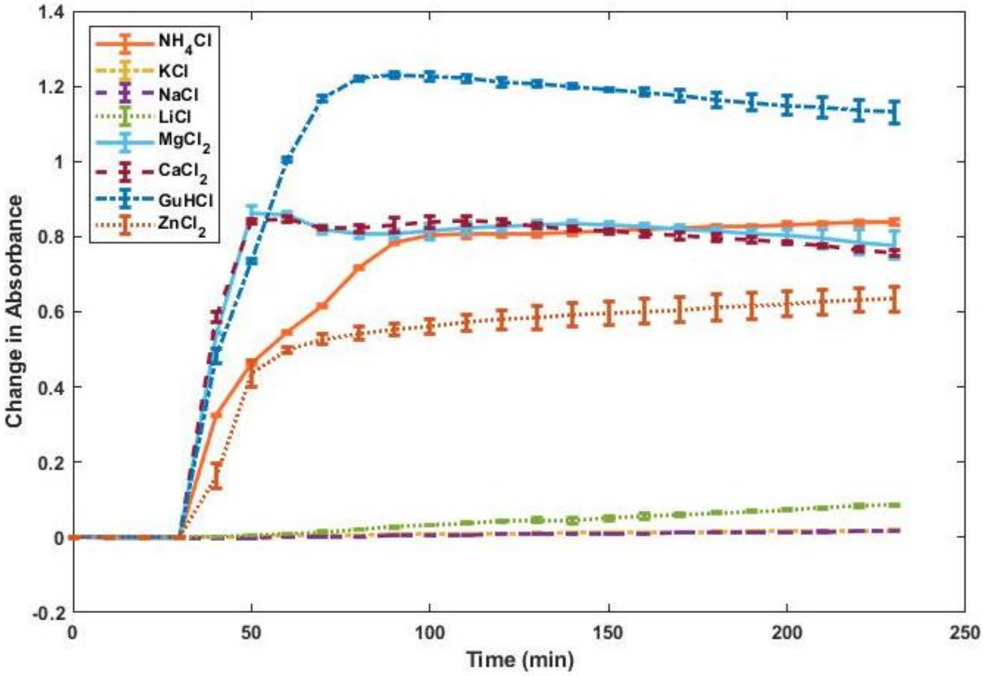


Figure 20. Absorbance measurements of 100 µL samples of GFP-TDP43 [21.63 µM] 600 nm were taken every 10 minutes over the course of 240 minutes. After the first 30 minutes each sample was spiked with 10 µL of specified cations with a final concentration of 150 mM in the 110 µL each sample and measurements continued to assess the aggregation induced based on the increase in turbidity measurements. Each sample was done in triplicate and the averages are depicted in the plot with error bars indicating the standard deviation from the mean.

GFP was purified under the same denaturing conditions as the GFP-TDP-43 and dialyzed into 10 mM HEPES (pH 7.5) buffer. GFP was then subjected to the same experimental procedures to ensure that the cations were indeed causing the TDP-43 to refold and not the GFP on the N-terminal of the protein, as seen in Fig. 21, the salts had no effect on the GFP alone.

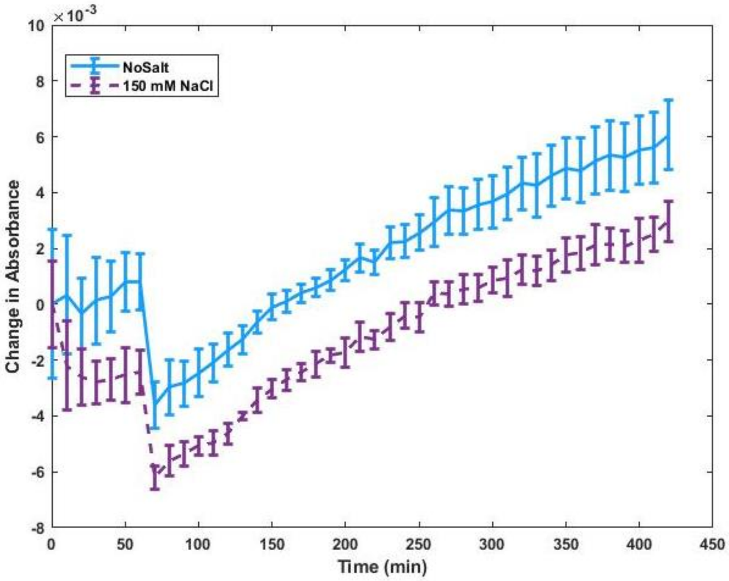


Figure 21. Absorbance measurements of 100 µ L samples of GFP [~20 µM] at 600 nm were taken every 10 minutes over the course of 420 minutes. After the first 30 minutes each sample was spiked with 10 µL of NaCl with a final concentration of 150 mM in the 110 µL each sample or 10 µL of buffer (10 mM HEPEs pH 7.5) and measurements were continued to assess the aggregation induced based on the increase in turbidity measurements. Each sample was done in triplicate and the averages are depicted in the plot with error bars indicating the standard deviation from the mean.

$$v = \frac{V_{max}[S]}{K_M + [S]}$$

$v = rate\ of\ aggregation$
$V_{max} = \max rate\ achieved$
$[S] = concentraiton\ of\ cation$
$K_M = concentraiton\ of\ caiton\ at\ half \max rate$

Equation 2. Michaels-Menten Equation

Since our data showed some discrepancies from the Hofmeister series hypothesis, we wanted to see if the zeta potential could explain the aggregation trends. As mentioned previously, the zeta potential of a protein indicates the stability and its readiness to aggregate. Buffers containing salt ions have been shown to affect the zeta potential, which we hypothesized would cause the same aggregation trends we had observed. The Stern potential of the electrostatic double layer surrounding the proteins negative core is determined by the radius of the cation in solution. The positive charge will be drawn to the negative charge of the protein and surround the core (Fig. 10B). The larger the atomic radius of the cation the larger the Stern potential, therefore the larger the zeta potential (Fig. 10C). A larger zeta potential indicates a more stable and less aggregation prone protein (Fig. 10A). To calculate the effect of the atomic radius of each cation on the aggregation of the protein we wanted to obtain a kinetic graph correlating the radius to the initial rate of aggregation. Looking at the divalent cations, the aggregation rate appears to decrease with time when induced with a higher concentration of 150 mM or 300 mM of the salt. However, when looking at the slope of the initial increase in absorbance, we can relate it to Michaelis-Menten (MM) kinetics model (Eqn. 2) which shows a trend in concentration dependence when there is a maximum rate possible for the substrate to reach. It is apparent that the rate of aggregation still increases with concentration and that it has reached a maximum rate of aggregation at 150 mM (Fig. 22), which follows the trend for MM. Since the concentration of GFP-TDP-43 is held constant in this experiment, the plateau seen for the maximum initial rate of aggregation in the divalent cations is showing complete aggregation of the protein in solution within the first 30 minutes after induction with either divalent cation (Fig. 20). For the monovalent cations, the initial rate of aggregation does not appear to hit a max even at a concentration of 300 mM as seen by the linear trend when comparing the concentration versus the rate of change in absorbance (Fig. 23A). Using this slope for each monovalent cation, we were able to compare the concentration dependent rate of aggregation against the atomic radii of the cation. The results yielded an exponential graph that fits a trend line of $y = (2*10^{-8})e^{-0.04x}$ with an correlation (R^2) value of 0.9994 (Fig. 23B). This correlation indicates that the atomic radii of the cations may affect the rate of aggregation, which supports the hypothesis that the zeta potential is a reason behind the aggregation trends we are seeing with salt addition.

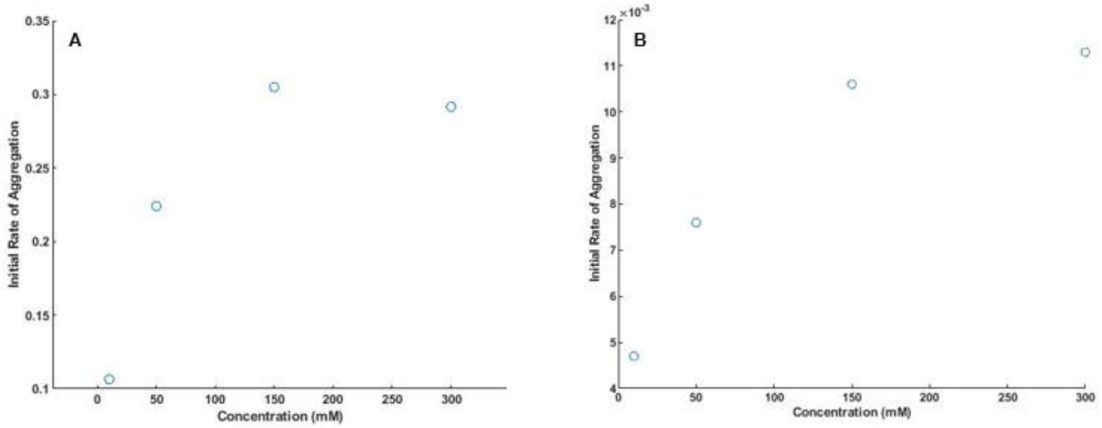


Figure 22. Initial rate of aggregation from the first 30 minutes after induction of (A) $CaCl_2$ and (B) $MgCl_2$ plotted against the concentration of the salt. Both plots show the rate plateauing between 150 mM and 300 mM salt concentration, showing the minimum amount of salt needed to reach the maximum rate of aggregation under each condition.

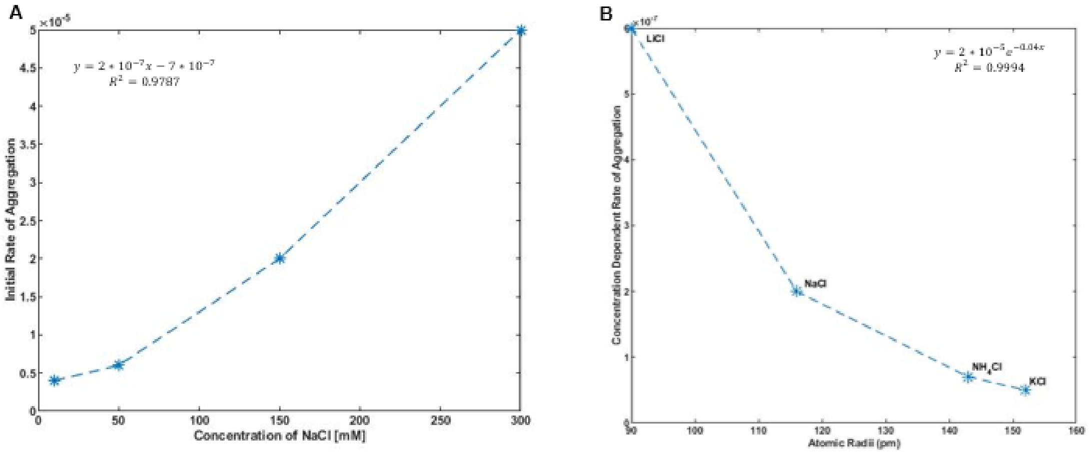


Figure 23. (A) Rate of aggregation during the first 30 minutes after NaCl induction is measured by obtaining the slope from time versus turbidity plot for each concentration (10 mM, 50 mM, 150 mM, 300 mM). The slot is plotted as the initial rate of aggregation versus the concentration. Line of best fit yielded a linear equation with a confidence interval (R^2) of 0.9787. (B) The slope for each monovalent cation extrapolated from concentration of cation versus initial rate of aggregation was obtained to show the concentration dependent rate of aggregation. To determine the correlation between atomic radii and aggregation the atomic radii (pm) was plotted against the concentration dependent rate of aggregation. Line of best fit showed an exponential decay relationship between atomic radii and concentration dependent rate of aggregation with an R^2 of 0.9994.

3.4 ZetaView® Data Analysis Shows Different Size Aggregates Present for Each Cation

To better visualize the aggregated species present 12 hours post-induction with salt, we ran two different samples through a ZetaView® to obtain images of the particles in solution and the hydrodynamic diameter distribution. We used one sample that had been induced with 150 mM NaCl and a second one that had been induced with 10 mM $CaCl_2$. After 12 hours the 150 mM NaCl solution appeared turbid, whereas the 10 mM $CaCl_2$ protein sample had phase separated. When the sample was loaded to the ZetaViewer®, the phase separated species was resuspended in solution. Imaging was done through an eGFP channel to remove most buffer noise from the analysis. The resulting images showed small particles present in the sample with 150 mM NaCl (Fig. 24B) and a range of small to large particles present in the sample with 10 mM $CaCl_2$ (Fig. 24C). The software computed the hydrodynamic diameter of the particles from the Brownian motion and electrophoretic mobility of each particle. It uses scanning nanoparticle tracking analysis (NTA) to obtain measurements from 11 different places in the sample. The resulting hydrodynamic diameter distribution was polydisperse for the 10 mM $CaCl_2$ solution with particles ranging from 134.3 nm to 640.8 nm and monodisperse for the 150 mM NaCl solution with a hydrodynamic radius of 78.0 nm (Fig. 24A). This corresponds with the DLS data we obtained from the DynaPro® Plate Reader II in that the 150 mM NaCl produced a monodisperse peak (Fig. 24D) and the 10 mM $CaCl_2$ produced multiple peaks (Fig, 24E).

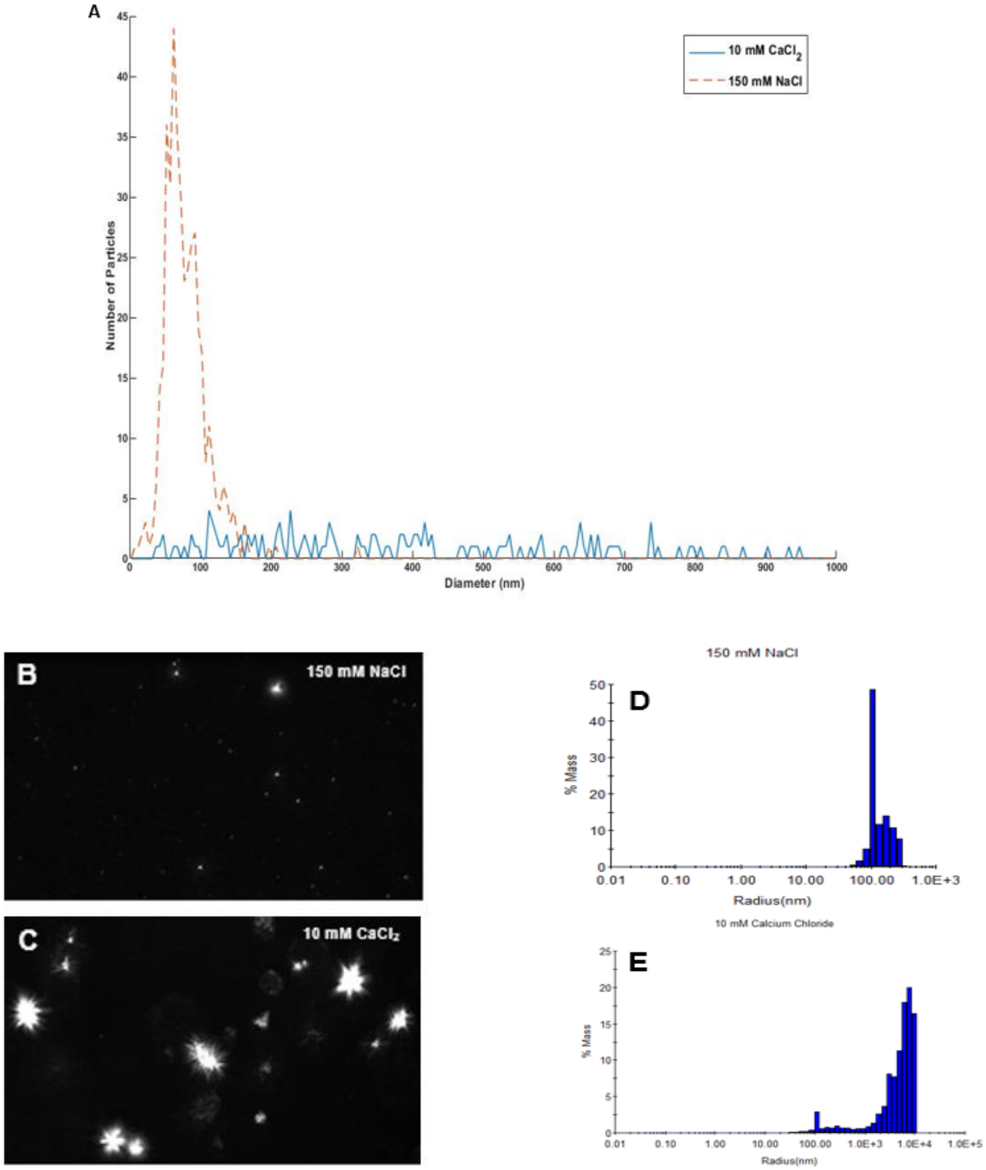


Figure 24. (A) Size distribution measurement of GFP-TDP43 S48E aggregates in 10 mM HEPES (pH 7.5) incubated with either 10 mM CaCl$_2$ (244 particles detected) or 150 mM NaCl (215 particles detected) by ZetaViewer. In this experiment, initial protein concentration was 25 uM with a diameter of approximately 2 nm. ZetaViewer was run after salt was induced into the metastable samples and allowed to aggregate overnight. Measurements taken using eGFP fluorescence to isolate the protein from the light scattering noise. Images of samples were taken during analysis of the sample containing 150 mM NaCl (B) and 10 mM CaCl$_2$ (C). DLS measurements of each ending species containing either 10 mM CaCl$_2$ (D) and 150 mM NaCl (E) shows the same polydisperse and monodisperse distribution of particle size, respectively, as the ZetaViewer. Hydrodynamic radius (nm) derived from DYNAMICS software using Stokes-Einstein equation based on a spherical model to give the predicted diameter of the particles in solution.

3.5 GFP-Tag is Necessary for TDP-43 to Remain Soluble

A previous graduate student had determined that the eGFP vector was optimal for protein expression of TDP-43 because the N-terminal GFP tag helped keep the protein soluble. Since our aggregation studies required a denaturing purification of the protein before allowing it to refold into a meta-stable intermediate, we wanted to see if it was possible to obtain this protein without the GFP tag. A denaturing expression and purification (See Methods) was done for TDP-43 in a pDest17 vector (Fig. S3), which removed the solubility tag but still had a His6 tag allowing for Ni affinity column purification. After the first dialysis at room temperature the solution containing the protein appeared more turbid than initially and continued to become more turbid during the second dialysis into 10 mM HEPES (pH 7.5). Using absorbance measurements, the same experiment was repeated using pDest17-TDP-43. The turbidity results showed an aggregated species to begin with based on the initial absorbance measurement being approximately 0.2 whereas the GFP-TDP-43’s baseline was around 0.04 (Fig. 25). This aggregated species was confirmed using DLS which showed a hydrodynamic radius of 415.3 nm. After salt addition there was little to no change in absorbance indicating that it had already reached its max aggregation for the volume allotted in the well plates (Fig. 26). This confirms that the solubility tag was necessary, even with a denaturing purification, to study the aggregation kinetics of TDP-43. Since GFP alone shows no aggregation propensity when subjected to the same salt conditions this data can be inferred to represent TDP-43 without a GFP tag.

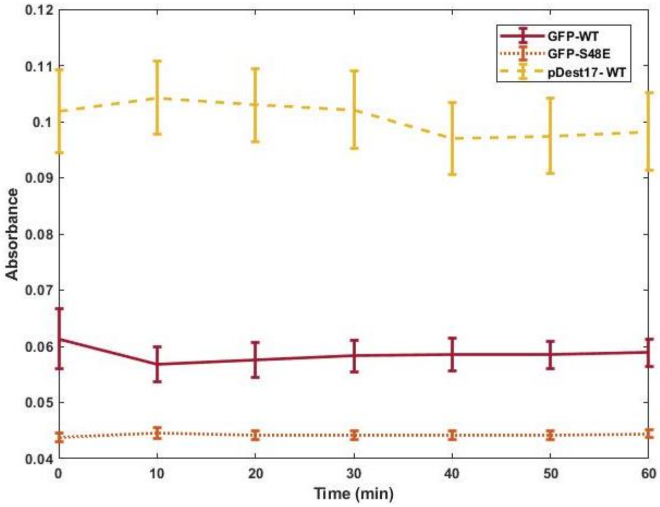


Figure 25. Average absorbance measurements of the meta-stable intermediate form of GFP-TDP-43 WT [19.66 µM] (red), GFP-TDP-43 S48E [22.43 µM] (orange) and pDest17-TDP-43 WT [~20 µM] (yellow) shows a larger species with the sample in the pDest17 vector compared to the eGFP vector. Samples done in triplicate and error bars represent the standard deviation.

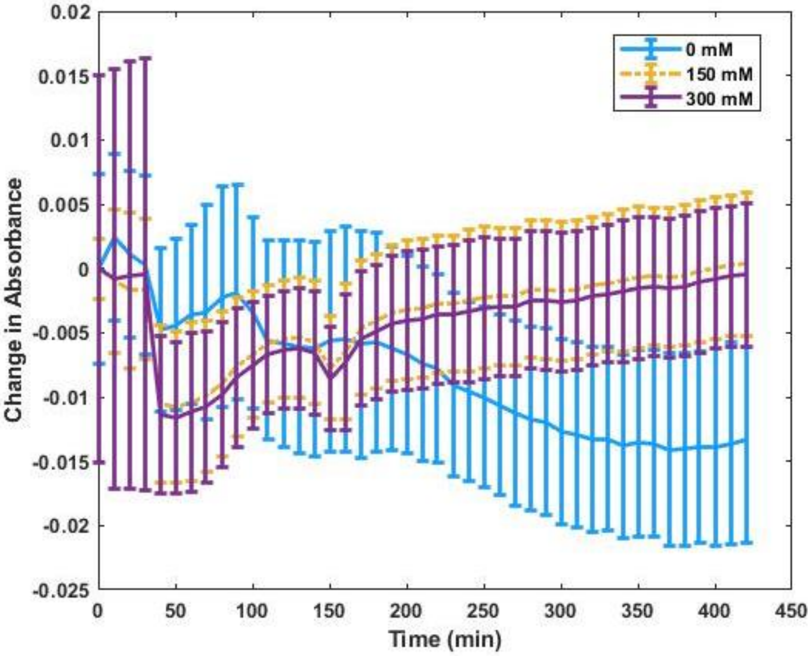


Figure 26. Absorbance measurements of 100 µ L samples of pDest17-TDP-43 WT [~20 µM] at 600 nm were taken every 10 minutes over the course of 420 minutes. After the first 30 minutes each sample was spiked with either 10 µL of NaCl (final concentration = 150 mM or 300 mM) or 10 µL of buffer (10 mM HEPEs pH 7.5) and measurements were continued to assess the aggregation induced based on the increase in turbidity measurements. Each sample was done in triplicate and the averages are depicted in the plot with error bars indicating the standard deviation from the mean.

3.6 TG12 DNA Causes a Decrease in Rate of Aggregation of Meta-Stable Intermediate

Previous studies have shown that TDP-43 binds DNA and RNA oligomers with sequence specificity for TG and UG repeats (Sun 2019). This binding has been hypothesized to reduce or inhibit the aggregation of TDP-43. To test this hypothesis, we incubated TG12 and AC12 DNA oligomers at varying concentrations and repeated the salt-induced aggregation assay. DLS measurements of the initial solution showed a polydisperse solution with multiple species present (Fig.27B). Data analysis after salt induced aggregation showed a concentration dependent inhibition of aggregation with TG12 but no significant change in aggregation with AC12, even at high concentrations of the DNA oligomer (Fig. 27A). At a 1:2 protein to TG12 DNA oligomer there is a slight decrease, and then an even more significant decrease when the concentration of TG12 is increased to 1:5 and then again 1:10 (Fig. 27A). Therefore, with excess DNA we do see that it plays a mitigating role in the aggregation kinetics of GFP-TDP-43. DLS measurements 12 hours post induction showed that the samples that contained 1:10 TG12 had a monodisperse profile, indicating that this inhibition may be coming from the interaction between the protein and DNA oligomer (Fig. 27C).

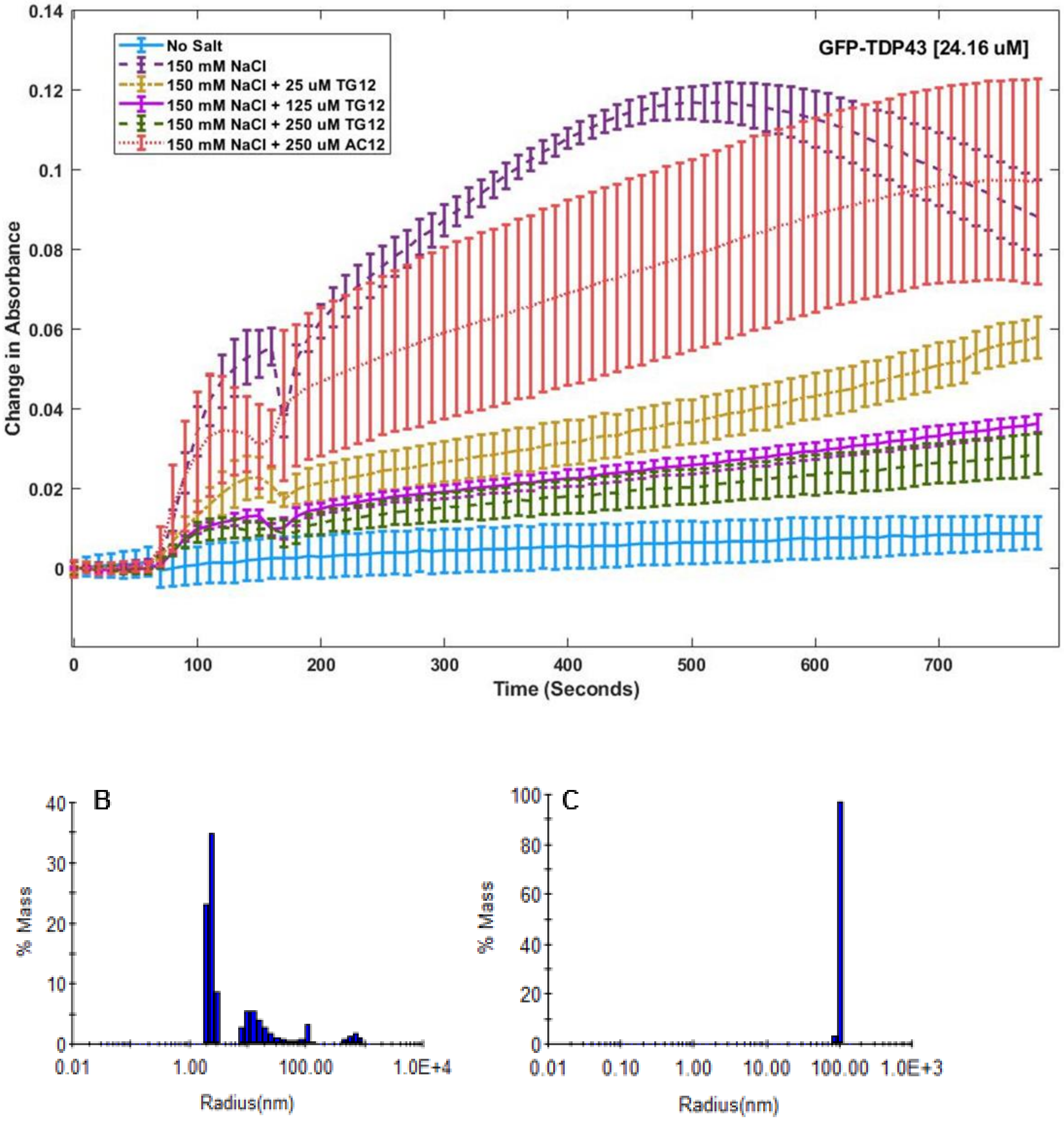


Figure 27. (A) Absorbance measurements at 600 nm of 100 µL samples of GFP-TDP43 [26.58 µM] with either TG12 DNA [25 µM, 125 µM, 250 µM], AC12 DNA [250 µM] or buffer (10 mM HEPES, pH 7.5) were taken every 10 minutes over the course of 800 minutes. After the first 30 minutes each sample was spiked with 10 µL of NaCl with a final concentration of 150 mM or buffer for a final volume of 110 µL each sample well and measurements continued to assess the aggregation induced based on the increase in turbidity measurements. Each sample was done in triplicate and the averages are depicted in the plot with error bars indicating the standard deviation form the mean. DLS measurements were taken before (B) and after (C) turbidity assay for the sample containing 250 µM TG12. Hydrodynamic radius (nm) derived from DYNAMICS software using Stokes-Einstein equation based on a spherical model.

CHAPTER 4: DISCUSSION AND SUMMARY

TDP-43 is a nucleic acid binding protein linked to the pathology of approximately 97% of the cases in ALS and approximately 45% in FTLD (Prasad 2019). It consists of 414 amino acid residues making up 4 main domains: the structured N-terminal domain, 2 RNA-recognition motifs and an unstructured C-terminal domain (Prasad 2016; Shimonaka 2016; Huang 2013; Sasaguri 2016). In its native form, TDP-43 is found in the dimeric state bound to RNA in the nucleus (Sun 2017; Shiina 2010). Its physiological function in the cell is to facilitate RNA processing mechanisms (e.g. RNA metabolism, mRNA transport, translation, splicing regulation) (Sasaguri 2016; Weskamp 2018; Zacco 2019; Coyne 2017; French 2019; Prasad 2019; Sun 2014). To promote these biological functions TDP-43 binds mRNA to transport it into the cytoplasm from the nucleus, and once it dissociates it moves back into the nucleus with support from its nuclear localization signal (NLS) region at the end of the NTD (Prasad 2019; McGurk 2018). The pathological hallmark of this protein in ALS is the presence of cytoplasmic inclusions of TDP-43 in the neurons of the patients. The mechanism of aggregation of this protein from the native dimeric state into a toxic aggregated species is still not well known. There are several different pathways the protein can take once in the cytoplasm that will disrupt its normal function. Some mutations can increase the aggregation propensity by disrupting RNA binding or disrupt its ability to re-localize to the nucleus (Prasad 2019; Zacco 2019; Mann 2019). Formation of stress granules can also occur when the cell is under oxidative stress to help facilitate the transcription of the RNA important in relieving the stress and then, under normal biological scenarios, it will dissociate back into the monomeric or dimeric state and transport back to the nucleus. However, when these stress granules do not dissociate then the protein can further oligomerize until it reaches an irreversible toxic state (Prasad 2019; Huang 2013; Colombrita 2009; Coyne 2017; Mann 2019).

Finally, there is also evidence that the cleavage of TDP-43 to TDP-35 can seed toxic aggregation by the 35 kDa fragments' propensity to oligomerize with full-length TDP-43 (Che *et al.* 2015). Before being able to test any potential therapeutics for ALS this mechanism of aggregation and the point at which TDP-43 goes from reversible oligomer to irreversible aggregate needs to be further defined.

To be able to study the mechanism of aggregation of TDP-43, the starting species needs to be classified. It has been shown that the native species of TDP-43 is found in the dimeric state bound to RNA in neurons (Shiina 2010). As described in the introduction to this thesis, there have been several dimer interfaces proposed in literature, so our first goal was to further characterize where this dimerization is occurring. Starting with the WT TDP-43 we purified and expressed the protein in an eGFP vector, with a N-terminal GFP tag, to promote solubility in solution per the protocol developed by a previous graduate student (Toro 2018). The purification yielded a large amount of the aggregated species and a small amount of what appears to be a dimeric species based on sizing standards and DLS measurements. Using this protein, we wanted to see if we were able to obtain a monomeric species through disruption of the dimer interfaces. The first dimer interface explored in this study was the head-to-tail dimerization of the NTDs occurring in an interface formed by a serine residue (S48) on the first NTD and a tyrosine (Y4) and glutamate (E17) of a second NTD. Through site-directed mutagenesis of WT GFP-TDP-43 we obtained an S48E mutant form of TDP-43 that breaks the proposed hydrogen bond between S48 and E17 (Kuo 2009). The second dimer interface explored was a head-to-head dimerization of the RRM2 domain proposed to involve an asparagine (D247) on two RRM2s' stabilized by an interaction with glutamate (E245) and isoleucine (I249). Using the same methodology as for the S48E mutant we obtained a D247L mutation that disrupts these interactions. Expression and purification of these mutants yielded two separate species off the size-exclusion column, which were verified by SDS-PAGE to be the same size as GFP-TDP-43, ~75 kDa. Through SEC analysis and DLS it is apparent that the smaller species from the purification was still in the dimer state with a hydrodynamic radius of 5.2 nm for the S48E mutant and 5.7 nm for the D247L mutant. We then used site-directed mutagenesis of the GFP-TDP-43

S48E to make a double mutant with the mutation in the NTD and RRM2. The purification and expression of this double mutant also yielded two separate species off the size-exclusion column that again was verified by SDS-PAGE to be the same size as expected for GFP-TDP43. However, this double mutant yielded a slightly right shifted species off the column in comparison to the WT and single point mutants second elution peak. This indicates the double mutant is present as a smaller species than the dimer species found for the WT, S48E and D247L. DLS measurements further indicated that the S48E/D247L species is likely a smaller species, potentially monomeric, due to the presence of a hydrodynamic radius of approximately 4.5 nm. Based on this result, we are predicting that TDP-43 can form a dimer in both the NTD and the RRM2, so when we disrupted one of the dimer interfaces it was still able to dimerize through the other interface. However, when both interfaces are disrupted the dimerization is lost. Dimerization is necessary for RNA binding and normal biological function, so understanding the areas that can form this dimerization interface and how it is disrupted allows for some insight into the beginning mechanics of the aggregation of TDP-43. To further confirm the characterization of the smaller species of the double mutant we can pursue electrospray ionization mass spectroscopy of each sample and directly compare the indicated molecular weights of the proposed dimeric single point mutants and wild-type versus monomeric double mutant.

The next aim of our study focused on the ability to induce aggregation of a meta-stable intermediate of WT GFP-TDP-43 and the S48E mutant. Being able to induce aggregation gives an opportunity to study the unknown kinetics and what influences the aggregation propensity. Based on previous studies done by Sun *et al.* 2018, aggregation of this metastable form of TDP-43 was induced by cations with increasing ability based on the reverse Hofmeister series. In our study, we similarly confirmed that a variety of cations induce aggregation of metastable TDP-43. However, several cations did not follow the Hofmeister relationship at all concentrations. Based on the series, ammonium ion should show less ability to induce aggregation than the potassium ion. At 10 mM concentration of each salt, we see that they have approximately the same aggregation propensity, however at 150 mM ammonium induces faster aggregation and a larger aggregate upon completion. In addition, if the trend followed the Hofmeister relationship then guanidinium should cause faster aggregation than calcium and magnesium, but at 10 mM it is significantly lower on the turbidity graph indicating a slower and smaller aggregate.

Since there were some deviance from the Hofmeister series, we started to look at other theories that show the effects that salts can have on proteins. There are claims that salts in buffers can change the zeta potential of a protein (Salgin 2012). The zeta potential can determine the proteins likelihood to aggregate; the larger the potential the more stable the protein. Cations can affect the zeta potential by compressing the electrostatic double layer of the hydration shell surrounding the protein (Salgin 2012). When the layers are compressed, proteins become closer together and have a higher likelihood of self-assembly. The layer determines the Stern potential and is determined by the size of the cation (Liu 2016; Salgin 2012). Based on this hypothesis, if the cations are inducing aggregation then there should be a correlation between atomic radius and rate of aggregation. The monovalent cations' have a slower rate of aggregation, which made them the optimal species to analyze for this correlation. We determined the cation concentration dependent rate of aggregation of the protein for each cation and plotted it against the atomic radii of the cations. This graph had an exponential relationship between the atomic radius and rate of aggregation with an R^2 value of 0.99, showing that there is likely a correlation between size and aggregation rate. The next part of the zeta potential would then be determined by the valency of the cations. A higher charge should suppress the hydration shell's double layer more, meaning the divalent cations should cause more aggregation than the monovalent cations. As seen in the turbidity assays, this indeed held true by the magnesium and calcium cations having significantly higher aggregation propensity than the potassium, sodium, lithium and ammonium at all concentrations. To further validate this claim we induced GFP-TDP-43 with $ZnCl_2$, a divalent cation that is not in the Hofmeister series but has shown an effect on protein aggregation in previous literature. The zinc cation also induced a faster aggregation and larger oligomer than the monovalent cations, supporting there is also a likely correlation between valency and aggregation kinetics. These aggregation assays indicate that the salt-induced aggregation of GFP-TDP-43 is more dependent on the effect that the cations have on the zeta potential of the protein rather than the Hofmeister series.

Lastly, we were able to begin to investigate factors that may either inhibit or mitigate the aggregation. TDP-43 is a DNA binding protein, so we first wanted to test the sequence specificity and if the interaction can inhibit or mitigate salt-induced aggregation. To test the sequence specificity, we used DNA oligomers of 12 TG repeats or 12 AC repeats, since it has been previously reported that TDP-43 has a higher affinity for TG/UG repeats with a length greater than 6 repeats for optimal binding (Buratti 2001; Sun 2014; Zacco 2019; Kitamura 2015). With an excess amount of DNA to GFP-TDP-43 (10:1) there was a significant difference between the effect of TG12 and AC12. Each sample was induced with 150 mM NaCl after being incubated with the DNA oligomer for 30 minutes. After 6 hours, the sample with no DNA and the sample with AC12 DNA had similar aggregation profiles indicating that AC12 had no effect on the aggregation of GFP-TDP-43. However, TG12 showed a significant decrease in the turbidity measurements, indicating that it had the ability to slow down the aggregation and decrease the final oligomer size. To test the concentration dependence of this interaction we ran the same experiment with increasing ratios of TG12 to GFP-TDP-43; 1:1, 5:1 and 10:1. The results showed that the ability of TG12 to mitigate the aggregation of GFP-TDP-43 increased as concentration increased. With the same concentration of TG12 to GFP-TDP-43, there was slight reduction of aggregation and then that significantly reduced when increasing the DNA to five times the amount of protein. From five to ten times the amount of DNA to protein there was a slight decrease in aggregation, but it appears to be reaching its limit showing that excess DNA was needed to decrease the aggregation of GFP-TDP-43 with a limit of how many DNA oligomers can interact with a single TDP-43.

Overall, this study was able to better classify the native state of TDP-43 and some of the aggregation kinetics of the protein. TDP-43 in an eGFP vector was used to study the dimer interface by comparing the species present after purification of the WT, S48E, D247L and S48E/D247L mutant variants. It was apparent that the WT had the largest aggregate to dimer ratio present and that the S48E and D247L single point mutants had a relatively mixed ratio of aggregate and dimer species. However, the double mutant yielded a species that is characteristic of a monomeric structure of GFP-TDP-43. These results indicate that TDP-43 can form a dimer through head-to-tail interactions in the NTD or head-to-head interactions in the RRM2. When only one of the interfaces is disrupted then it will still form a dimer through the second, but disruption of both interfaces causes a loss of dimerization. There are other proposed interfaces in the NTD, specifically a head-to-head dimer that forms a stable tetramer with a second homodimer, which should be tested to further characterize the dimerization (Jiang 2017).

Next, we used a meta-stable intermediate species of GFP-TDP-43 that is stable at room temperature in 10 mM HEPES (pH 7.5) buffer to determine the salt induced aggregation kinetics. We found that the ability for the cation present in the salt to induce aggregation correlated with its effect on the zeta potential of the protein. This is shown by divalent cations having a greater affect than monovalent cations due to their ability to allow for more compression of the electrostatic double layer surrounding the protein, and that the cations with a larger atomic radii had less of an effect on the induction of aggregation due to having a larger Stern potential. To further validate the effect of the atomic radii lower concentrations of divalent cations need to be tested to in order to get the concentration dependent rate of aggregation to plot against their atomic radii. In addition, the pH dependence of this aggregation should be determined. GuHCl causes the buffer to become more acidic therefore this could be the reason for the discrepancies in its aggregation propensity at low versus high concentrations (Salgin 2012; Mompean 2017). This experiment also needs to be repeated using anions from the Hofmeister series to see if they follow the zeta potential hypothesis we are proposing for the cations (Okur 2017). In addition, zinc is part of the Irving-Williams Series that orders transition metals based on their ability to stabilize complexs. Given that trend, we should test the other transition metals in the Irving-Williams Series (Ni^{2+}, Cu^{2+}, Co^{2+}, Mn^{2+}, Fe^{2+}) to see if the effect of Zn^{2+} is due to its valency and size according to the zeta potential or its complex stability potential (Garnier *et al.* 2017, Irving *et al.* 1953). If it follows the zeta potential theory then zinc would cause the fastest initial rate of aggregation, whereas if it follows the Irving-Williams series then the sample induced with copper would have the fastest initial rate of aggregation (Salgin 2012; Irving 1953). It would also be useful to see the effect of the cations on the D247L mutant. Aspartate is a negatively charged amino acid and when mutated to leucine the charge is lost. This could cause the salt induced aggregation propensity to increase since there would be less repulsive forces present between the proteins.

Finally, we began to investigate the ability to inhibit or mitigate the aggregation using DNA oligomers. We found that excess TG12 DNA did not inhibit, but it did significantly mitigate the aggregation, whereas excess AC12 DNA had no significant effect. This shows that TDP-43 can interact with DNA oligomers, with sequence specificity, in a way that could prevent aggregates from forming. These experiments have been done in both the WT and S48E point mutant constructs and yielded similar results. The next steps would be to test the S48E/D247L double mutant to see if TG12 has the same effect. Previous literature has proposed that dimerization is needed for DNA binding, so if this is the case then TG12 should not influence the aggregation of GFP-TDP43 S48E/D247L. Similarly, this should be repeated with an RRM mutant that has the five phenylalanine’s necessary for nucleic acid binding mutated to leucine’s to disrupt the binding capacity (Mann 2019; Buratti 2001; Sun 2014; Wang 2018; Bhardwaj 2013; Huang 2013; Flores 2019; Kitamura 2015; French 2019). Both the double mutant and the RRM mutant should yield similar profiles when incubated with excess TG12 and AC12, both having no effect on the aggregation. This result would indicate that it is the binding specificity that is causing the alleviated aggregation seen with the WT and S48E when incubated with excess TG12.

REFERENCES

Arrasate, M., & Finkbeiner, S. (2012). Protein aggregates in Huntington's disease. *Experimental Neurology*, *238*(1), 1–11. https://doi.org/10.1016/j.expneurol.2011.12.013

Asua, J. M. (2008). Polymer Reaction Engineering. *Polymer Reaction Engineering*, 1–367. https://doi.org/10.1002/9780470692134

Babinchak, W. M., Haider, R., Dumm, B. K., Sarkar, P., Surewicz, K., Choi, J. K., & Surewicz, W. K. (2019). The role of liquid-liquid phase separation in aggregation of the TDP-43 low-complexity domain. *Journal of Biological Chemistry*, *294*(16), 6306–6317. https://doi.org/10.1074/jbc.RA118.007222

Bostro, M. (2005). *Hofmeister Effects in Surface Tension of Aqueous Electrolyte Solution*. *3*, 2619–2623. https://doi.org/10.1021/la047437v

Chang, C. K., Chiang, M. H., Toh, E. K. W., Chang, C. F., & Huang, T. H. (2013). Molecular mechanism of oxidation-induced TDP-43 RRM1 aggregation and loss of function. *FEBS Letters*, *587*(6), 575–582. https://doi.org/10.1016/j.febslet.2013.01.038

Chang, C. K., & Huang, T. H. (2016). Untangling the structure of the TDP-43 N-terminal domain. *FEBS Journal*, *283*(7), 1239–1241. https://doi.org/10.1111/febs.13676

Chang, C., Wu, T., Wu, C., Chiang, M., Toh, E. K., Hsu, Y., Lin, K., Liao, Y., Huang, T., & Huang, J. J. (2012). Biochemical and Biophysical Research Communications The N-terminus of TDP-43 promotes its oligomerization and enhances DNA binding affinity. *Biochemical and Biophysical Research Communications*, *425*(2), 219–224. https://doi.org/10.1016/j.bbrc.2012.07.071

Chemistry, S., Liu, X., Ding, W., Tian, R., Li, R., Li, H., & Ideas, C. (2016). *How Ionic Polarization Affects Stern Potential: An Insight into Hofmeister Effects*. *July*, 1181–1189. https://doi.org/10.2136/sssaj2016.04.0095

Colombrita, C., Zennaro, E., Fallini, C., Weber, M., Sommacal, A., Buratti, E., Silani, V., & Ratti, A. (2009). TDP-43 is recruited to stress granules in conditions of oxidative insult. *Journal of Neurochemistry*, *111*(4), 1051–1061. https://doi.org/10.1111/j.1471-4159.2009.06383.x

Controls, S. (2004). *Letters to the Editor*. 1321–1322.

Cooper-Knock, J., Hewitt, C., Highley, J. R., Brockington, A., Milano, A., Man, S., Martindale, J., Hartley, J., Walsh, T., Gelsthorpe, C., Baxter, L., Forster, G., Fox, M., Bury, J., Mok, K., McDermott, C. J., Traynor, B. J., Kirby, J., Wharton, S. B., … Shaw, P. J. (2012). Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. *Brain*, *135*(3), 751–764. https://doi.org/10.1093/brain/awr365

Coyne, A. N., Zaepfel, B. L., & Zarnescu, D. C. (2017). Failure to deliver and translate—new insights into RNA dysregulation in ALS. *Frontiers in Cellular Neuroscience*, *11*(August), 1–13. https://doi.org/10.3389/fncel.2017.00243

Denver), N. T. (University of. (2018). *Characterization of a Phosphomimetic Mutant of the ALS Associated Protein TDP-43* (Issue November).

Fields, C. R., Bengoa-Vergniory, N., & Wade-Martins, R. (2019). Targeting Alpha-Synuclein as a Therapy for Parkinson’s Disease. *Frontiers in Molecular Neuroscience*, *12*(December), 1–14. https://doi.org/10.3389/fnmol.2019.00299

Flores, B. N., Li, X., Malik, A. M., Martinez, J., Beg, A. A., & Barmada, S. J. (2019). An Intramolecular Salt Bridge Linking TDP43 RNA Binding, Protein Stability, and TDP43-Dependent Neurodegeneration. *Cell Reports*, *27*(4), 1133-1150.e8. https://doi.org/10.1016/j.celrep.2019.03.093

François-Moutal, L., Perez-Miller, S., Scott, D. D., Miranda, V. G., Mollasalehi, N., & Khanna, M. (2019). Structural Insights Into TDP-43 and Effects of Post-translational Modifications. *Frontiers in Molecular Neuroscience*, *12*(December), 1–22. https://doi.org/10.3389/fnmol.2019.00301

French, R. L., Grese, Z. R., Aligireddy, H., Dhavale, X. D. D., Reeb, A. N., Kedia, N., Kotzbauer, P. T., Bieschke, J., & Ayala, Y. M. (2019). *cro Detection of TAR DNA-binding protein 43 (TDP-43) oligomers*. *43*, 6696–6709. https://doi.org/10.1074/jbc.RA118.005889

Furukawa, Y., Suzuki, Y., Fukuoka, M., Nagasawa, K., Nakagome, K., Shimizu, H., Mukaiyama, A., & Akiyama, S. (2016). A molecular mechanism realizing sequence-specific recognition of nucleic acids by TDP-43. *Scientific Reports*, *6*(January), 1–12. https://doi.org/10.1038/srep20576

Heyburn, L., Sajja, V. S. S. S., & Long, J. B. (2019). The role of TDP-43 in military-relevant TBI and chronic neurodegeneration. *Frontiers in Neurology*, *10*(JUN). https://doi.org/10.3389/fneur.2019.00680

Jarred M. Kaiser, PhD*,†, Michael F. Vignos, MS†, Richard Kijowski, MD‡, Geoffrey Baer, MD, PhD§, and Darryl G. Thelen, PhD†,§, |. (2017). 乳鼠心肌提取 HHS Public Access. *Physiology & Behavior*, *176*(1), 139–148. https://doi.org/10.1016/j.physbeh.2017.03.040

Jiang, L. L., Xue, W., Hong, J. Y., Zhang, J. T., Li, M. J., Yu, S. N., He, J. H., & Hu, H. Y. (2017). The N-terminal dimerization is required for TDP-43 splicing activity. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-06263-3

Kiang, K. M. Y., & Leung, G. K. K. (2018). A Review on Adducin from Functional to Pathological Mechanisms: Future Direction in Cancer. *BioMed Research International*, *2018*. https://doi.org/10.1155/2018/3465929

Kitamura, A., Nakayama, Y., Shibasaki, A., Taki, A., Yuno, S., Takeda, K., Yahara, M., Tanabe, N., & Kinjo, M. (2016). Interaction of RNA with a C-terminal fragment of the amyotrophic lateral sclerosis-associated TDP43 reduces cytotoxicity. *Scientific Reports*, *6*(January), 1–14. https://doi.org/10.1038/srep19230

Kuo, P. H., Doudeva, L. G., Wang, Y. T., Shen, C. K. J., & Yuan, H. S. (2009). Structural insights into TDP-43 in nucleic-acid binding and domain interactions. *Nucleic Acids Research*, *37*(6), 1799–1808. https://doi.org/10.1093/nar/gkp013

Li, H. R., Chiang, W. C., Chou, P. C., Wang, W. J., & Huang, J. rong. (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), 6090–6098. https://doi.org/10.1074/jbc.AC117.001037

Mcalary, L., Plotkin, S. S., Yerbury, J. J., & Cashman, N. R. (2019). *Prion-Like Propagation of Protein Misfolding and Aggregation in Amyotrophic Lateral Sclerosis*. *12*(November), 1–21. https://doi.org/10.3389/fnmol.2019.00262

McEwan, W. A., Falcon, B., Vaysburd, M., Clift, D., Oblak, A. L., Ghetti, B., Goedert, M., & James, L. C. (2017). Cytosolic Fc receptor TRIM21 inhibits seeded tau aggregation. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(3), 574–579. https://doi.org/10.1073/pnas.1607215114

Mcgurk, L., Gomes, E., Guo, L., Mojsilovic-petrovic, J., Tran, V., Kalb, R. G., Shorter, J., & Bonini, N. M. (2019). *contact*. *71*(5), 703–717. https://doi.org/10.1016/j.molcel.2018.07.002.Poly

Mompeán, M., Baralle, M., Buratti, E., & Laurents, D. V. (2016). An amyloid-like pathological conformation of TDP-43 is stabilized by hypercooperative hydrogen bonds. *Frontiers in Molecular Neuroscience*, *9*(NOV2016), 1–8. https://doi.org/10.3389/fnmol.2016.00125

Mompeán, M., Romano, V., Pantoja-Uceda, D., Stuani, C., Baralle, F. E., Buratti, E., & Laurents, D. V. (2017). Point mutations in the N-terminal domain of transactive response DNA-binding protein 43 kDa (TDP-43) compromise its stability, dimerization, and functions. *Journal of Biological Chemistry*, *292*(28), 11992–12006. https://doi.org/10.1074/jbc.M117.775965

Morgan, S., & Orrell, R. W. (2016). Pathogenesis of amyotrophic lateral sclerosis. *British Medical Bulletin*, *119*(1), 87–97. https://doi.org/10.1093/bmb/ldw026

Needham, P. G., Copley, K. E., & Hurtle, B. (2020). *HHS Public Access*. *102*(2), 321–338. https://doi.org/10.1016/j.neuron.2019.01.048.RNA

Neumann, M., Sampathu, D. M., Kwong, L. K., Truax, A. C., Micsenyi, M. C., Chou, T. T., Bruce, J., Schuck, T., Grossman, M., Clark, C. M., Mccluskey, L. F., Miller, B. L., Masliah, E., Mackenzie, I. R., Feldman, H., Feiden, W., Kretzschmar, H. A., Trojanowski, J. Q., & Lee, V. M. (2006). *Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis*. *314*(October), 130–134.

Nucci, N. V., & Vanderkooi, J. M. (2008). Effects of Salts of the Hofmeister Series on the Hydrogen Bond Network of Water. *Journal of molecular liquids*, *143*(2-3), 160–170. https://doi.org/10.1016/j.molliq.2008.07.010

Okur, H. I., Hlad, J., Rembert, K. B., Cho, Y., Heyda, J., Dzubiella, J., Cremer, P. S., & Jungwirth, P. (2017). *Beyond the Hofmeister Series: Ion-Speci fi c E ff ects on Proteins and Their Biological Functions*. https://doi.org/10.1021/acs.jpcb.6b10797

Papegaey, A., Eddarkaoui, S., Deramecourt, V., Fernandez-Gomez, F. J., Pantano, P., Obriot, H., Machala, C., Anquetil, V., Camuzat, A., Brice, A., Maurage, C. A., Le Ber, I., Duyckaerts, C., Buée, L., Sergeant, N., & Buée-Scherrer, V. (2016). Reduced Tau protein expression is associated with frontotemporal degeneration with progranulin mutation. *Acta Neuropathologica Communications*, *4*(1), 1–14. https://doi.org/10.1186/s40478-016-0345-0

Pillai, M., & Jha, S. K. (2019). *The Folding and Aggregation Energy Landscapes of Tethered RRM Domains of Human TDP-43 Are Coupled via a Metastable Molten Globule-like Oligomer*. https://doi.org/10.1021/acs.biochem.8b01013

Prasad, A., Bharathi, V., Sivalingam, V., Girdhar, A., & Patel, B. K. (2019). Molecular mechanisms of TDP-43 misfolding and pathology in amyotrophic lateral sclerosis. *Frontiers in Molecular Neuroscience*, *12*(February), 1–36. https://doi.org/10.3389/fnmol.2019.00025

Salgin, S., Salgin, U., & Bahadir, S. (2012). Zeta potentials and isoelectric points of biomolecules: The effects of ion types and ionic strengths. *International Journal of Electrochemical Science*, *7*(12), 12404–12414.

Shiina, Y., Arima, K., Tabunoki, H., & Satoh, J. I. (2010). TDP-43 dimerizes in human cells in culture. *Cellular and Molecular Neurobiology*, *30*(4), 641–652. https://doi.org/10.1007/s10571-009-9489-9

Shimonaka, S., Nonaka, T., Suzuki, G., Hisanaga, S. I., & 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), 8896–8907. https://doi.org/10.1074/jbc.M115.713552

Sun, Y., Arslan, P. E., Won, A., Yip, C. M., & Chakrabartty, A. (2014). Binding of TDP-43 to the 3′UTR of Its Cognate mRNA Enhances Its Solubility. *Biochemistry*, *53*(37), 5885–5894. https://doi.org/10.1021/bi500617x

Sun, Y., & Chakrabartty, A. (2017). Phase to Phase with TDP-43. *Biochemistry*, *56*(6), 809–823. https://doi.org/10.1021/acs.biochem.6b01088

Sun, Y., Medina-Cruz, A., Hadley, K. C., Galant, N. J., Law, R., Vernon, R. M., Morris, V. K., Robertson, J., & Chakrabartty, A. (2018). Physiologically Important Electrolytes as Regulators of TDP-43 Aggregation and Droplet-Phase Behavior [Research-article]. *Biochemistry*, *58*, 590–607. https://doi.org/10.1021/acs.biochem.8b00842

Toro, N. (University of. Denver) (2018). *Characterization of a Phosphomimetic Mutant of the ALS Associated Protein TDP-43*.

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., & Fawzi, N. L. (2018). A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. *The EMBO Journal*, *37*(5), 1–18. https://doi.org/10.15252/embj.201797452

Wei, Y., Lim, L., Wang, L., & Song, J. (2017). Biochemical and Biophysical Research Communications ALS-causing cleavages of TDP-43 abolish its RRM2 structure and unlock CTD for enhanced aggregation and toxicity. *Biochemical and Biophysical Research Communications*, *485*(4), 826–831. https://doi.org/10.1016/j.bbrc.2017.02.139

Weskamp, K., Barmada, S. J., & Arbor, A. (2019). *HHS Public Access*. *1693*, 67–74. https://doi.org/10.1016/j.brainres.2018.01.015.TDP43

Zacco, E., Graña-Montes, R., Martin, S. R., de Groot, N. S., Alfano, C., Tartaglia, G. G., & Pastore, A. (2019). RNA as a key factor in driving or preventing self-assembly of the TAR DNA-binding protein 43. *Journal of Molecular Biology*, *431*(8), 1671–1688. https://doi.org/10.1016/j.jmb.2019.01.028

Zhang, Y. J., Caulfield, T., Xu, Y. F., Gendron, T. F., Hubbard, J., Stetler, C., Sasaguri, H., Whitelaw, E. C., Cai, S., Lee, W. C., & Petrucelli, L. (2013). The dual functions of the extreme N-terminus of TDP-43 in regulating its biological activity and inclusion formation. *Human Molecular Genetics*, *22*(15), 3112–3122. https://doi.org/10.1093/hmg/ddt166

APPENDIX

	Primer Sequence (5' → 3')
Forward	TTGTGGAGAGaagTTGATCATTAAAGGAATCAGCG
Reverse	AGAGACTGCGCAATCTGA

Table S1. Primers for Site-Directed Mutagenesis of WT GFP-TDP-43 and S48E to D247L and S248E/D247L, respectively.

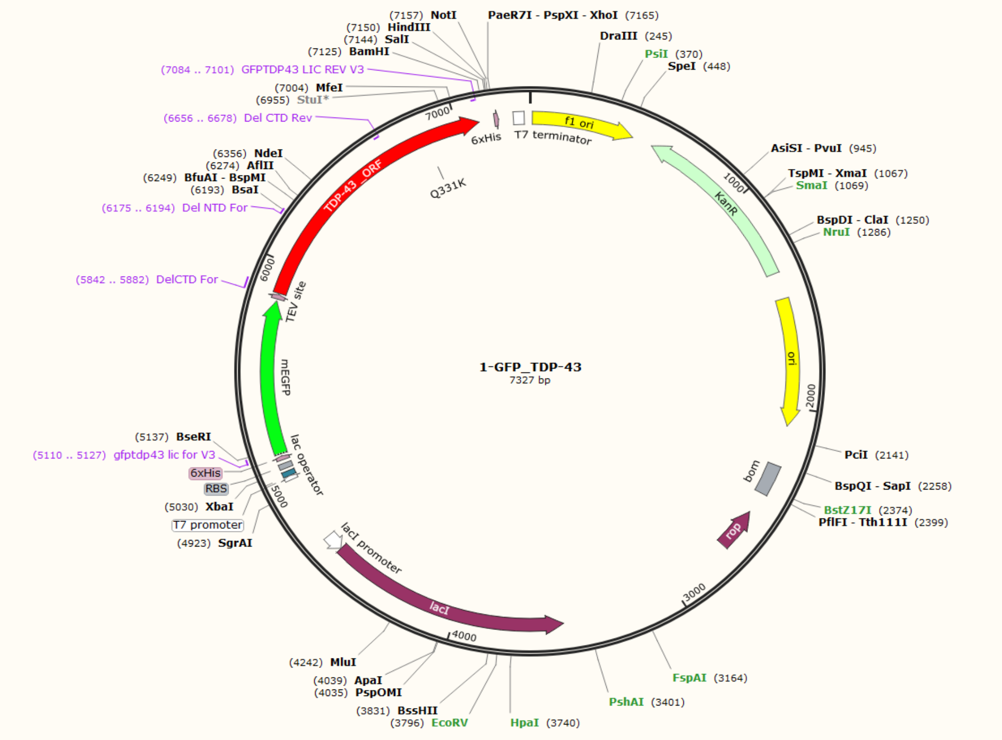


Figure S1. Expression vector containing full-length TDP-43. It shows a 6x-histamine (6xHis) tag and enhanced green fluorescent protein (eGFP) on the N-terminal with a Tobacco Etch Virus (TEV) cleavage site and kanamycin (KanR) resistance.

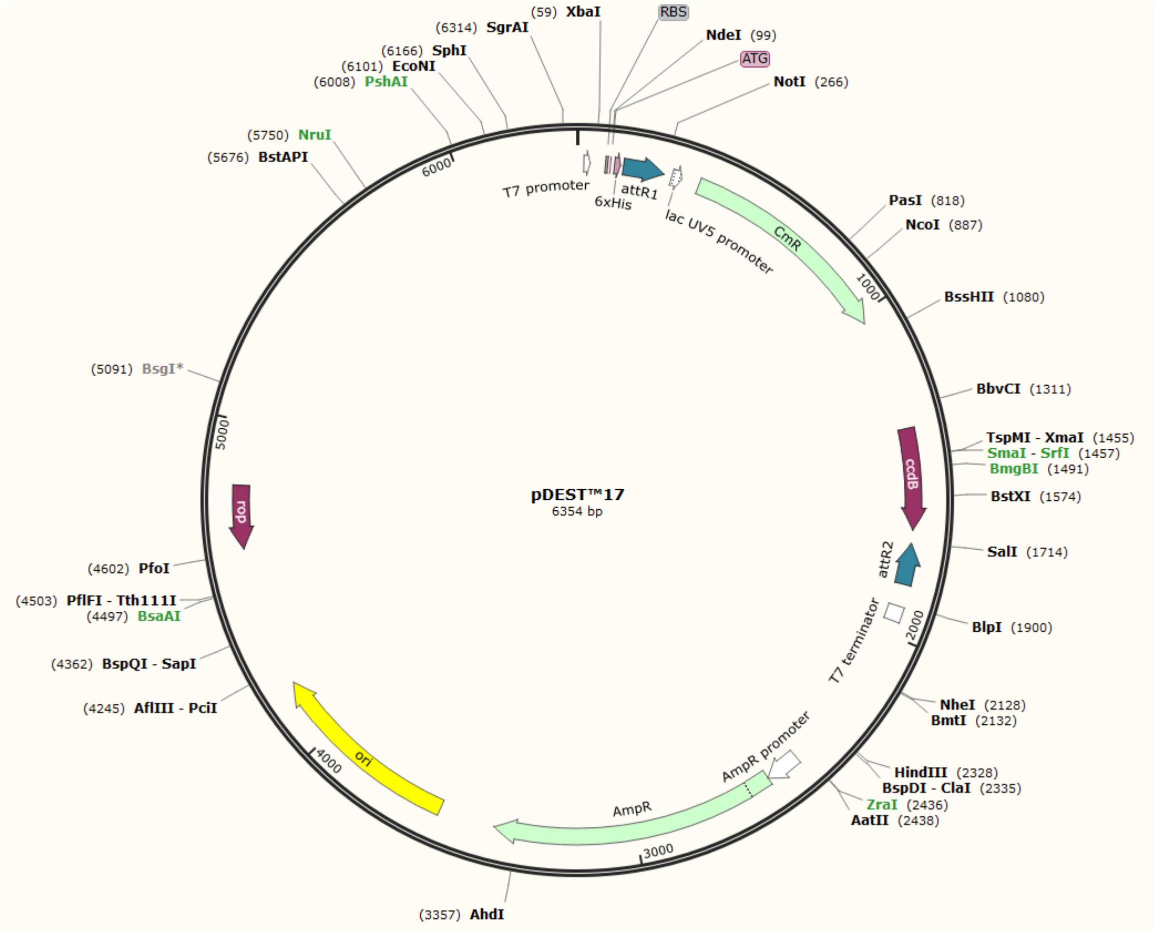


Figure S2. Gateway^{TM} pDest^{TM} 17 expression vector. It has a 6x-histadine (6xHis) tag, ampicillin (AmpR) resistance and *attR* site for recombination.

	Nucleic Acid Sequence
TG12	TGTGTGTGTGTGTGTGTGTGTGTG
AC12	ACACACACACACACACACACACAC

Table S2. Sequence of DNA oligomers used in aggregation assays.

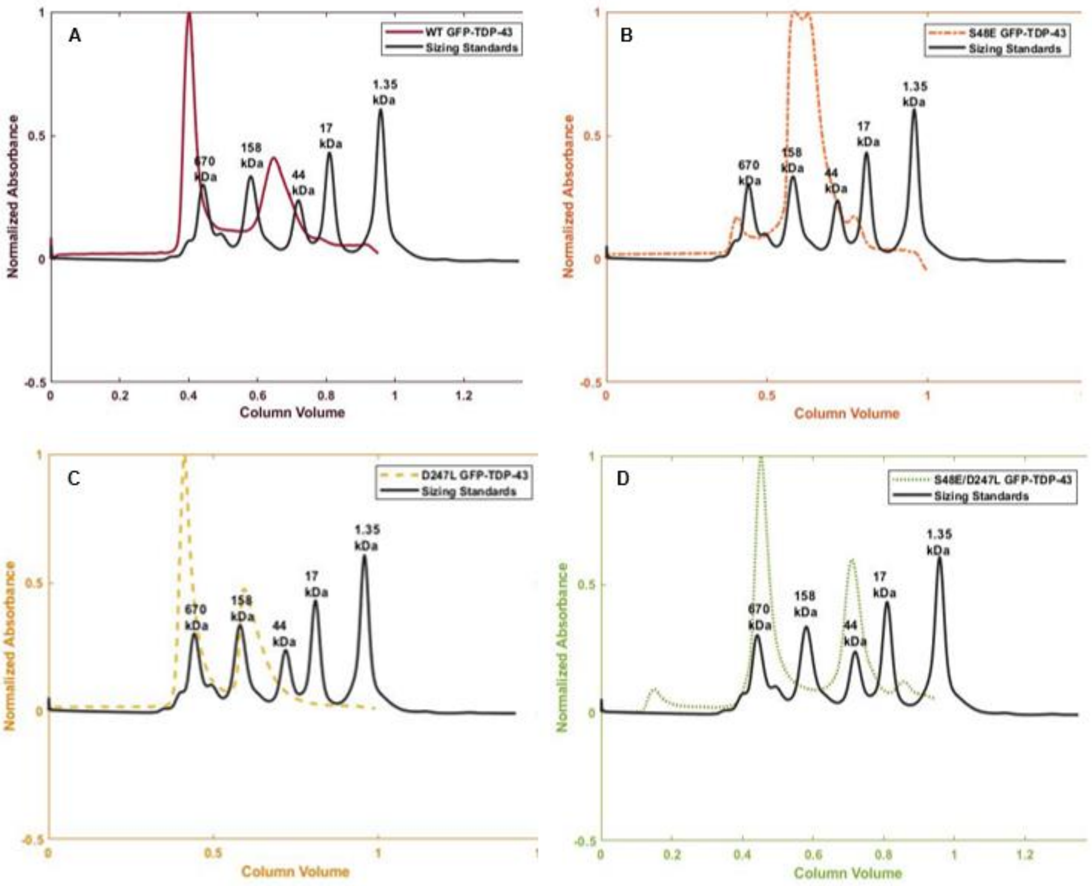


Figure S3. SEC chromatograms of GFP-TDP-43 (A) WT (solid dark red), (B) S48E mutant (dash/dot orange), (C) D247L mutant (dashed yellow) and (D) S48E/D247L double mutant (dotted green) plotted against Bio-Rad's Gel Filtration Standards (solid black). Sizing standards were run with same elution protocol as the GFP-TDP-43 constructs (see methods).