

University of Denver

Digital Commons @ DU

---

Electronic Theses and Dissertations

Graduate Studies

---

1-1-2019

## Characterization of a Phosphomimetic Mutant of the ALS Associated Protein TDP-43

Nicole Toro  
*University of Denver*

Follow this and additional works at: <https://digitalcommons.du.edu/etd>



Part of the [Neuroscience and Neurobiology Commons](#)

---

### Recommended Citation

Toro, Nicole, "Characterization of a Phosphomimetic Mutant of the ALS Associated Protein TDP-43" (2019). *Electronic Theses and Dissertations*. 1628.  
<https://digitalcommons.du.edu/etd/1628>

This Thesis is brought to you for free and open access by the Graduate Studies at Digital Commons @ DU. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ DU. For more information, please contact [jennifer.cox@du.edu](mailto:jennifer.cox@du.edu), [dig-commons@du.edu](mailto:dig-commons@du.edu).

---

# Characterization of a Phosphomimetic Mutant of the ALS Associated Protein TDP-43

## Abstract

Trans-activation response (TAR) DNA-binding protein 43 (TDP-43) is a natively dimeric 414-residue protein that is encoded by the human *TARDBP* gene that has important implications in the pathogenesis of the neurodegenerative disorders ALS, FTD, and CTE. TDP-43 has been found hyperphosphorylated and ubiquitinated in the aggregates of the affected neurons of these diseases. The discovery of the presence of TDP-43 positive inclusions in brain matter of patients with CTE has made repetitive brain injury a possible environmental stimulus for aggregation in TDP-43 proteinopathies. We expand upon the hypothesis that TDP-43 readily aggregates under agitation conditions and that the addition of poly-TG repeats to TDP-43 in aggregation conditions attenuates its aggregation propensity. We expressed a recombinant S48E phosphomimetic mutation of TDP-43 with an N-terminal GFP fluorescent tag in *Escherichia coli* and induced aggregation by agitation. We examined the extent to which different DNA/RNAs and differing stoichiometric concentrations of these nucleic acids affected the aggregation *in vitro*. We show that the addition of RNA/DNA to the S48E mutant does not have profound effects on aggregation attenuation under *in vitro* aggregation conditions.

## Document Type

Thesis

## Degree Name

M.S.

## Department

Chemistry and Biochemistry

## First Advisor

Erich G. Chapman, Ph.D.

## Keywords

Aggregation, Amyotrophic lateral sclerosis, Neurodegenerative disorders, Phosphomimetic, S48E, TDP-43, TAR DNA-binding protein 43

## Subject Categories

Life Sciences | Neuroscience and Neurobiology

## Publication Statement

Copyright is held by the author. User is responsible for all copyright compliance.

Characterization of a Phosphomimetic Mutant of the ALS Associated Protein TDP-43

---

A Thesis

Presented to

the Faculty of Natural Sciences and Mathematics

University of Denver

---

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

---

by

Nicole Toro

November 2018

Advisor: Erich G. Chapman

©Copyright by Nicole Toro 2018

All Rights Reserved

Author: Nicole Toro  
Title: Characterization of a Phosphomimetic Mutant of the ALS Associated Protein TDP-43  
Advisor: Erich G. Chapman  
Degree Date: November 2018

## ABSTRACT

Trans-activation response (TAR) DNA-binding protein 43 (TDP-43) is a natively dimeric 414-residue protein that is encoded by the human *TARDBP* gene that has important implications in the pathogenesis of the neurodegenerative disorders ALS, FTD, and CTE. TDP-43 has been found hyperphosphorylated and ubiquitinated in the aggregates of the affected neurons of these diseases. The discovery of the presence of TDP-43 positive inclusions in brain matter of patients with CTE has made repetitive brain injury a possible environmental stimulus for aggregation in TDP-43 proteinopathies. We expand upon the hypothesis that TDP-43 readily aggregates under agitation conditions and that the addition of poly-TG repeats to TDP-43 in aggregation conditions attenuates its aggregation propensity. We expressed a recombinant S48E phosphomimetic mutation of TDP-43 with an N-terminal GFP fluorescent tag in *Escherichia coli* and induced aggregation by agitation. We examined the extent to which different DNA/RNAs and differing stoichiometric concentrations of these nucleic acids affected the aggregation *in vitro*. We show that the addition of RNA/DNA to the S48E mutant does not have profound effects on aggregation attenuation under *in vitro* aggregation conditions.

## TABLE OF CONTENTS

<b>CHAPTER 1: INTRODUCTION</b> .....	1
1. 1 Protein Aggregation and Human Disease .....	1
1.1.1 Intermediate Unfolded States.....	3
1.1.2 Self-association .....	5
1.1.3 Chemical degradation .....	5
1.2 TDP43 Biology and its implications in neurodegenerative disorders .....	6
1.2.1 ALS .....	8
1.2.2 FTD .....	8
1.2.3 CTE.....	9
1.3 Structure of TDP43 .....	11
1.3.1 N-Terminal Domain.....	13
1.3.2 RNA Recognition Motifs.....	16
1.3.3 C-Terminal Domain.....	19
1.3.4 Domain Contributions to Aggregation.....	21
1.4 Nucleic Acid Binding .....	22
1.4.1 Uncertainty in RNA Binding Models .....	24
1.5 Ribonucleoprotein Complexes.....	24
<b>CHAPTER 2: MATERIALS AND METHODS</b> .....	26
2.1 LIC Cloning .....	26
2.2 Site Directed Mutagenesis .....	27
2.3 Expression and Purification .....	27
2.4 Denaturing Expression and Purification .....	28
2.5 <i>In vitro</i> Aggregation Studies.....	29
2.6 Fluorescence Assays .....	30
2.7 Sedimentation Analysis of Aggregation Experiments.....	30
2.8 Dynamic Light Scattering .....	30
2.9 Determining Concentration of TDP-43 .....	31
2.10 AUG12 RNA .....	32
<b>CHAPTER 3: RESULTS</b> .....	33
3.1 WT TDP-43 Readily Forms Aggregates.....	33
3.2 Urea Denaturing Purification yields an aggregated TDP-43 .....	36
3.3 A S48E Phosphomimic Mutant Maintains Dimer Conformation after Purification.....	38
3.5 Cleaved protein visible in aggregates .....	46
<b>CHAPTER 4: DISCUSSION AND SUMMARY</b> .....	48
<b>REFERENCES</b> .....	56
<b>APPENDIX</b> .....	64

## LIST OF FIGURES

### CHAPTER 1

Figure 1. Schematic illustrations of Mechanisms of Protein Aggregation .....	4
Figure 2. TDP-43's roles in RNA metabolism. ....	7
Figure 3. TDP-43 forms cytoplasmic aggregates in neurodegenerative diseases.....	10
Figure 4. Structure of TDP-43 .....	12
Figure 5. NTD of TDP-43.....	13
Figure 6. NTD of TDP-43 adopts a ubiquitin-like fold. ....	14
Figure 7. Domain interaction of TDP-43 NTD Dimer. ....	15
Figure 8. Individual RRMs .....	17
Figure 9. TDP-43 RRMs bind UG nucleotides.....	18
Figure 10. C-Terminal Domain.....	19
Figure 11. Illustration of proposed CTD aggregation by $\beta$ -strand expansion modeled by MD Simulations .....	20
Figure 12. Phe $\rightarrow$ Leu mutation in RRM1 disrupts RNA binding. ....	23

### CHAPTER 3

Figure 13. Characterization of WT GFP-TDP-43.....	34
Figure 14.Characterization of GFP-TDP-43 in a native purification followed by Urea. .	37
Figure 15. Characterization of GFP-S48E-TDP-43.....	39
Figure 16. Characterization of GFP-S48E-TDP-43 in denaturing conditions.....	41
Figure 17. Aggregation of GFP-S48E-TDP-43 <i>in vitro</i> .....	43
Figure 18. Individual aggregation data of TDP-43 aggregation <i>in vitro</i> .....	44
Figure 19. Aggregation of S48E assessed by sedimentation analysis .....	45
Figure 20. Supernatant/ Pellet analysis of S48E aggregates after <i>in vitro</i> aggregation studies .....	47

### CHAPTER 4

Figure 21. S48E Mutation.....	49
-------------------------------	----

### APPENDIX

Figure S1. Nickel affinity chromatogram of a denaturing urea purification with WT TDP-43.....	64
Figure S2. Nickel affinity chromatogram of modified denaturing urea purification.....	64
Figure S3. Visual Fluorescence Assay.....	65

## CHAPTER 1: INTRODUCTION

### *1. 1 Protein Aggregation and Human Disease*

Cellular proteins are mortal and have cell cycles that are defined by their half-life. Two processes, synthesis and degradation, work in synchrony to maintain protein concentration in cells. Turnover of cells is physiologically important because intracellular levels of proteins should change in response to cellular stimuli during developmental and cell cycle processes. A dysfunction in the clearance system leads to accumulation of protein that can be toxic to the cell (Rubinsztein, 2006). Such structural aberrations ultimately aggregate either within the cell or extracellularly. Protein aggregates are a feature of several human diseases such as neurodegenerative diseases (Ross & Poirier, 2004), prion diseases (Aguzzi & Calella, 2009), cancer (Yang-hartwich, Bingham, Garofalo, Alvero, & Mor, 2015), metabolic diseases (De Baets, Van Doorn, Rousseau, & Schymkowitz, 2015), immune disorders (Kumar et al., 2016), and cystic fibrosis (Du et al., 2015).

Misfolded proteins originate from multiple pathways. Genetic mutations are often a tragic event that predispose a protein to unsuccessful folding. A defect in translational modification also contributes towards misfolding (Scheper, Van Der Knaap, & Proud, 2007). Failure of protein folding by chaperones is one of the other important factors that debilitates its function. In addition to compromising an individual proteins function, some of these misfolded proteins play a role in aberrant

RNA processing and formation of stress granules that subsequently direct towards aggregate formation. There are mechanisms in place that keep the protein properly folded, however, the mechanisms by which the protein is kept in homeostasis tends to devolve with time in a normal lifecycle of a cell. The ability of cells to control these mechanisms is a significant issue for research in aging (Ross & Poirier, 2004). A dysfunction in any of the mechanisms that maintain proper protein folding leads to aggregation, making protein aggregation an important prognosis in disease pathology (Chiti & Dobson, 2006; De Baets et al., 2015).

Protein aggregation occurs when a protein monomer or collection of proteins associates into a complex larger than its native state (Miyazaki et al., 2016). Misfolded proteins can become toxic to the cell when they aggregate because of exposed sequences that are usually buried within the hydrophobic core of the protein to the solvent solution surrounding it (Chiti & Dobson, 2006; De Baets et al., 2015). The hydrophobic cores of amyloidogenic proteins have aggregation prone regions (APRs) that, when exposed, have the potential to form intra- and intermolecular  $\beta$ -strand interactions that are observed in aggregates (De Baets et al., 2015; Westerlund, von Heijne, & Emanuelsson, 2009). The consequences of protein aggregation as loss or gain of function is still debated. Nevertheless, this process which leads to cellular death is considered a symptom and in some cases a cause of disease (De Baets et al., 2015).

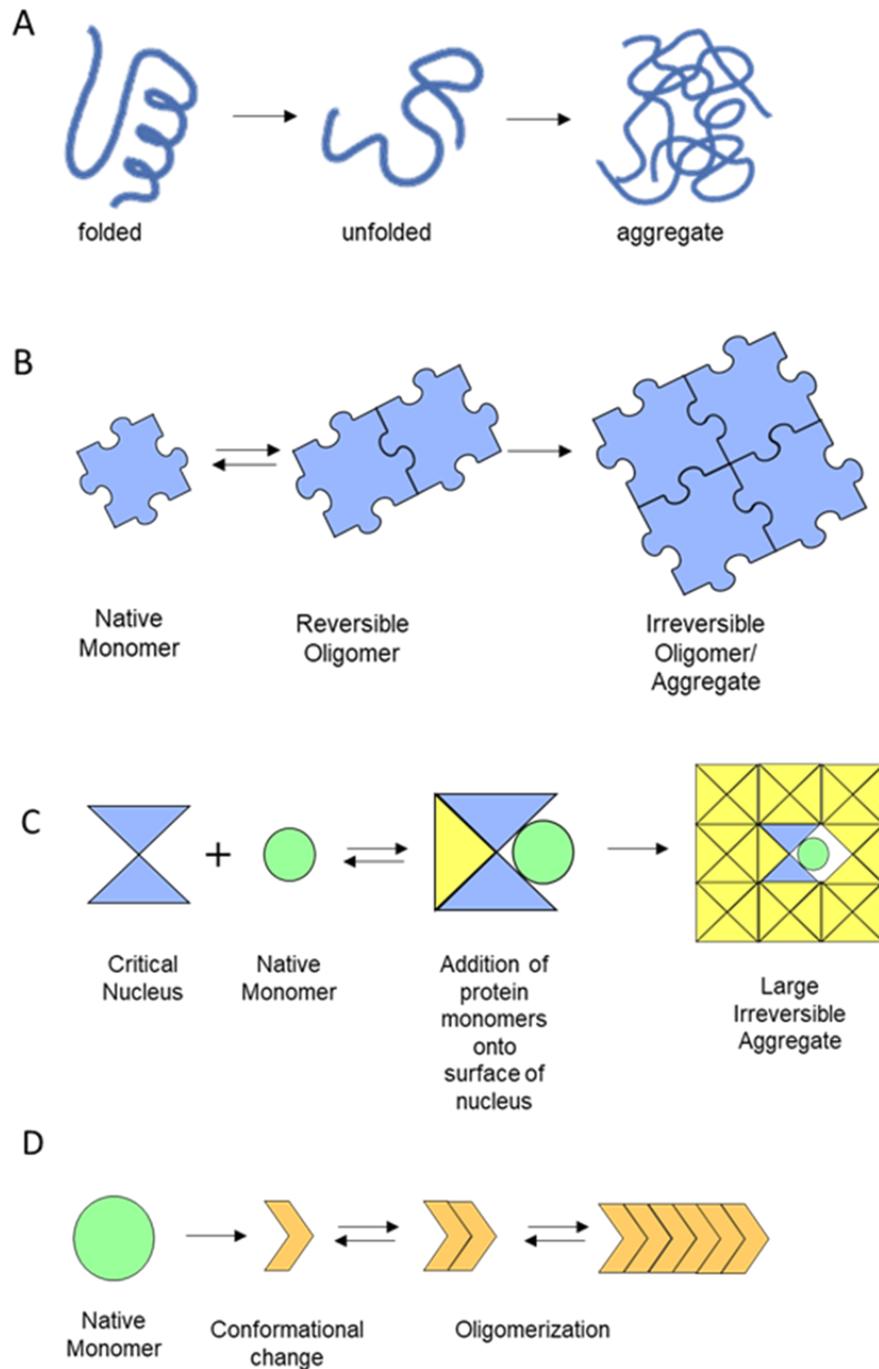
There are many different characteristics to aggregates and aggregation pathways. For example, aggregation can be reversible as it is for yeast pyruvate kinase, Cdc19, where its low complexity domain (LCD) is dephosphorylated to reverse aggregation

(Höhfeld, 2017). On the other hand, aggregation can be irreversible as is seen with TDP-43 where large SDS insoluble inclusions are formed (Lim, Wei, Lu, & Song, 2016). In addition, proteins have differing behaviors and localize to different cellular compartments. This may result in various, multistep aggregation pathways both between multiple proteins and within a single population of protein (E. Chi, Krishnan, & Kendrick, 2003; Van Buren, Rehder, Gadgil, Matsumura, & Jacob, 2009). Aggregation is a complex process that can be a product of protein instability (Treuheit, Kosky, & Brems, 2002; W. Wang, 2005). The multiple mechanisms by which proteins aggregate can be more or less divided into three categories: 1) intermediate unfolded states, 2) self-association or chemical linkages and 3) chemical degradation (W. Wang, Nema, & Teagarden, 2010).

### *1.1.1 Intermediate Unfolded States*

Proteins in solution exist in an equilibrium of folded, intermediate, and unfolded/denatured states. Contrary to natively folded proteins which are less likely to aggregate because they do not have as many solvent exposed hydrophobic side chains, the intermediate states expose more hydrophobic patches. (W. Wang et al., 2010).

Normal cell processes do one of three things with these intermediates: a protein will follow a traditional post- or co-translational pathway and fold to its native state, have a chaperone bind the intermediate and assist with proper folding to the native state, or the protein can be degraded by proteolysis mechanisms. In the absence of these processes



**Figure 1. Schematic illustrations of Mechanisms of Protein Aggregation** A) the folding pathway of a native monomer exposes the hydrophobic core in the intermediate state which can serve as a catalyst for aggregation B) Complementary patches on the surface of a native protein can lead to the formation of oligomers C) Formation of a thermodynamically disfavored critical nucleus drives the formation of large irreversible aggregates D) Chemical degradation of the protein causes a conformational change that can drive the protein to aggregate

### 1.1.2 Self-association

Some protein monomers have the propensity to self-associate due to the self-complementarity of their surfaces (Fig. 1B). This favors the formation of smaller reversible oligomers that encourage oligomer growth through processes such as nucleation and seeding, which is an aggregation process that produces large insoluble aggregates (E. Y. Chi, Krishnan, Randolph, & Carpenter, 2003). In this mechanism, a small oligomer is formed, termed the critical nucleus, that grows in size to form a large oligomeric species (Fig. 1C)(Philo & Arakawa, 2009).

### 1.1.3 Chemical degradation

Chemical degradation changes the physical properties of the protein bringing about a conformational change that may drive the protein towards aggregation. Conformational changes of the native fold to a nonnative state may spur aggregation by exposing the hydrophobic residues (Fig. 1D) (Chaturvedi, Siddiqi, Alam, & Khan, 2016; W. Wang et al., 2010).

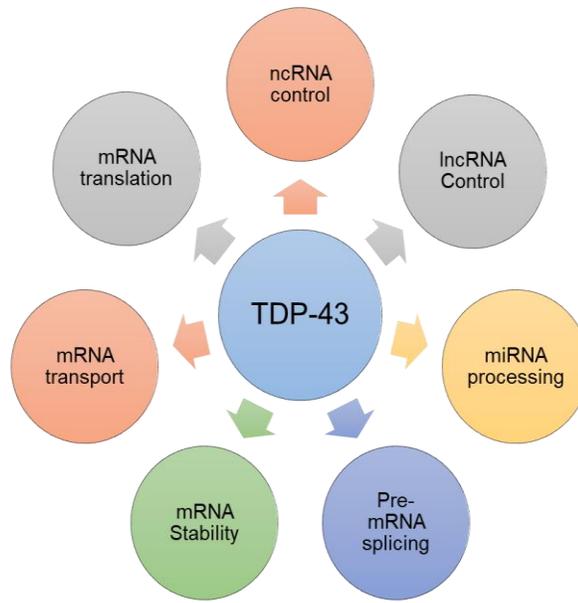
Aggregation is prompted by varying environmental stress conditions *in vivo* that can be mimicked experimentally *in vitro* (Kiese, Pappenberg, Friess, & Mahler, 2008). Stimuli of protein aggregation include: extreme temperature conditions, freeze-thaw cycles, mechanical stressors, and chemical modifications. Increasing temperatures disrupt intramolecular interactions in the protein and accelerate unfolding (Day, Bennion, Ham, & Daggett, 2002). At higher temperatures, proteins can be oxidized and deaminated leading to oligomerization (Brange, Havelund, & Hougaard, 1992). Oxidation of proteins is a major chemical degradation pathway that leads to aggregation (Shihong Li, ’

Christian Schoneich, 1995). In addition, freeze-thaw introduces interactions with ice crystals which can be a damaging stressor for proteins (Hillgren, Lindgren, & Aldén, 2002; Kuelto, Wang, Randolph, & Carpenter, 2008). The native structure of the protein fold is not stable due to the protein-ice interactions that increase the solvent accessible surface area of the native fold (Strambini & Gonnelli, 2007).

Aggregated proteins can be either soluble or insoluble in nature (Krielgaard et al., 1998). Aggregation may occur following the chemical modification of the protein or its surroundings. Solubility can be dependent on buffer conditions which may stabilize the native state of the protein (Katayama et al., 2006) or lead to aggregation (Kameoka, Masuzaki, Ueda, & Imoto, 2007; Krishnan et al., 2002). Because major changes in pH can greatly affect protein stability, aggregation is correlated to changes in pH values (Kuelto et al., 2008). Mechanical stressors disturb the processes that mediate the control mechanisms which target protein misfolding in a cell (Hachiya, Kozuka, & Kaneko, 2008). Stirring and shaking are both mechanical stressors which induce aggregation in proteins (Mahler, Müller, Frieß, Delille, & Matheus, 2005). Each of these driving factors make protein aggregation a complex process with important pathological implications in humans.

### *1.2 TDP43 Biology and its implications in neurodegenerative disorders*

Gaining knowledge about the mechanisms by which neurological diseases work has become more accessible with the identification of specific proteins in the aggregates of affected cells. Here, we focus on the presence of trans-activation response (TAR) DNA-binding protein 43 (TDP-43) (Fig. 4) and its role in neurodegenerative diseases. TDP-43



**Figure 2. TDP-43's roles in RNA metabolism.** TDP-43 plays pleiotropic effects in RNA metabolism pathways

was first discovered to bind TAR element in human immunodeficiency virus 1 (HIV-1) (Ou et al., 1995). TDP-43's discovery in human disease occurred when it was found to induce exon skipping on a human cystic fibrosis transmembrane conductance regulator (CFTR) exon 9 by binding to the TG(m) repeat near the 3' splice site (Morimoto & Cuervo, 2009). Since this discovery TDP-43 has been widely studied and has been shown to play major roles in RNA metabolism such as: pre-mRNA splicing, miRNA processing, lncRNA and ncRNA control, mRNA stability, mRNA transport, and mRNA translation (Ratti & Buratti, 2016)(Fig 2). TDP-43 has also been found in inclusions of postmortem brain tissue of patients with Amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD) making it a pathological marker for both of these disorders (Woollacott & Rohrer, 2016).

### *1.2.1 ALS*

ALS is a fatal motor neuron disease characterized by motor neuron degradation that leads to cell death in the central nervous system (CNS)(Bonafede & Mariotti, 2017; Brown & Al-Chalabi, 2017; Johnson et al., 2009; Morgan & Orrell, 2016). The effects of this disease are paralytic as the most common first symptom is loss of limb function with symptoms progressing to death resulting from respiratory failure (Brown & Al-Chalabi, 2017). ALS is the most common motor neuron disease of the middle age community with symptoms appearing around the age of 50 (Cleveland & Rothstein, 2001). Recent data show that there are approximately 3 per 100,000 cases of ALS worldwide where 90% of cases are sporadic and 10% of cases are familial (Morgan & Orrell, 2016). First symptom to diagnosis is 12 months with a 3-5 year life expectancy after diagnosis (Kiernan et al., 2011). Characteristics of mutant TDP-43 in ALS are loss of axonal transport function and RNA processing. These factors are suggested preface other factors of the disease pathology, however, the mechanism of pathology for ALS is unknown leaving no cure or preventative measures to take against this disease (Morgan & Orrell, 2016; Scotter, Chen, & Shaw, 2015).

### *1.2.2 FTD*

Clinically, FTD presents with changes in cognitive abilities, personality, and behavior (Borroni et al., 2017; Haass & Neumann, 2016; Mackenzie & Neumann, 2016; Woollacott & Rohrer, 2016). It is part of a spectrum of neurodegenerative disorders being the second most common form of young onset dementia to be shown in patients under the age of 65 (Woollacott & Rohrer, 2016). FTD's pathology is diverse with the

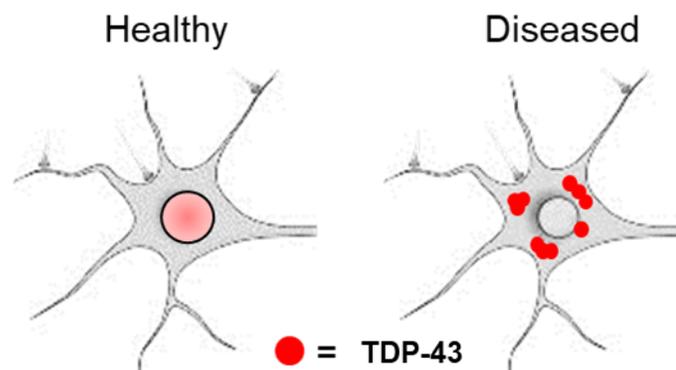
accumulation of tau and TDP-43 in cells of the frontal and temporal lobes of the brain being a consistent features in the majority of cases (Haass & Neumann, 2016). FTD is often misdiagnosed due to overlapping symptoms that present as the disease progresses (Woollacott & Rohrer, 2016). Three mutations in MAPT, GRN, and C9orf72 have been identified to provoke the formation of the inclusions that lead to cell degeneration (Mackenzie & Neumann, 2016). Even with the knowledge of the mutations that cause proteins to form inclusions, there is no known cure for FTD as the mechanisms by which these inclusions form is still unknown and there is no assay that is able to pre-screen for the mutations (Borroni et al., 2017).

### *1.2.3 CTE*

Chronic Traumatic Encephalopathy (CTE), described as “punch drunk” in the early 1900s owing to its discovery and association with boxing (Martland, 1928), is a neurodegenerative disorder with clinical symptoms similar to that of FTD and motor neuron disease: with symptoms including mental confusion, slower muscular movements, unsteadiness, and tremors (Stein, Alvarez, Mckee, & Traumatic, 2015). In the brain, CTE presents as neurofibrillary tangles and dot like inclusions (McKee & Gavett, 2010). CTE has been identified in athletes (Mckee et al., 2009), military veterans (Lee E. Goldstein, et al., 2012), victims of physical abuse (Stein et al., 2015), epileptics (Puvenna et al., 2016), and in one case a dwarf who participated in dwarf throwing events in the circus (McKee & Gavett, 2010); all occupations where repetitive head trauma is routine. CTE can be broken down into 4 stages with stage 1 being the mildest case and stage 4 the most

advanced. The extent of this disorder is still unknown and diagnoses is not thoroughly assessed until after death (Stein et al., 2015).

TDP-43 has been found hyperphosphorylated and ubiquitinated in the aggregates of affected neurons of ALS and FTD (Brown & Al-Chalabi, 2017; McKee & Gavett, 2010; Morgan & Orrell, 2016; Neumann, Sampathu, & Kwong, 2006). A known distinction of neurodegenerative diseases is the ubiquitination of misfolded proteins in the CNS; what distinguishes ALS and FTD from other similar neurodegenerative diseases is that the proteinaceous inclusion bodies are tau- and  $\alpha$ -synuclein negative (Neumann et al., 2006). A known method of pathogenesis is the relocation of TDP-43 from the nucleus to the cytoplasm of neurons (Johnson et al., 2009) (Fig.3). A proposed mechanism for why this occurs is the overexpression of TDP-43. This occurs when there are mutations in its 3' untranslated region (UTR) which may be a by-product of the decline of proteostasis (Bonafede & Mariotti, 2017). Aggregation puts stress on the cell proposedly by impairing RNA metabolism and isolating proteins important to routine cell function (Morgan & Orrell, 2016). TDP-43, originally thought to only be associated with ALS and FTD, is



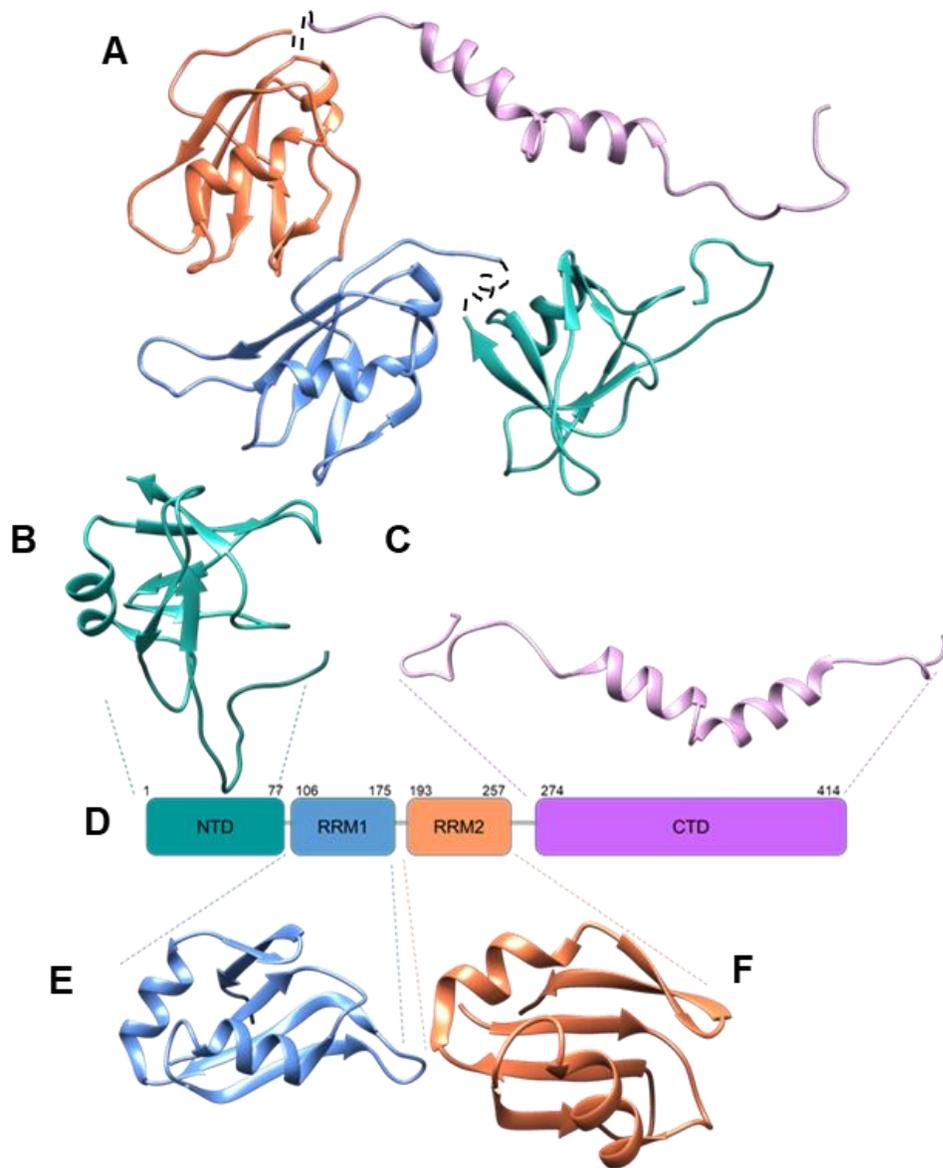
**Figure 3. TDP-43 forms cytoplasmic aggregates in neurodegenerative diseases.** Aggregated TDP-43 is mis-localized from the nucleus to the cytoplasm in diseased neurons of patients with ALS, FTD, and CTE.

quickly being recognized to be associated with other neurological disorders as a subsidiary pathology such as is seen with CTE. TDP-43 mediates the response in axonal injury, such as axonal stretching due to traumatic brain injury, drawing further association between TDP-43 proteinopathies and head injury (Mckee et al., 2009).

The discovery of the presence of TDP-43 positive inclusions in brain matter of patients with CTE has made repetitive brain injury a possible environmental stimulus for aggregation in TDP-43 proteinopathies.

### *1.3 Structure of TDP43*

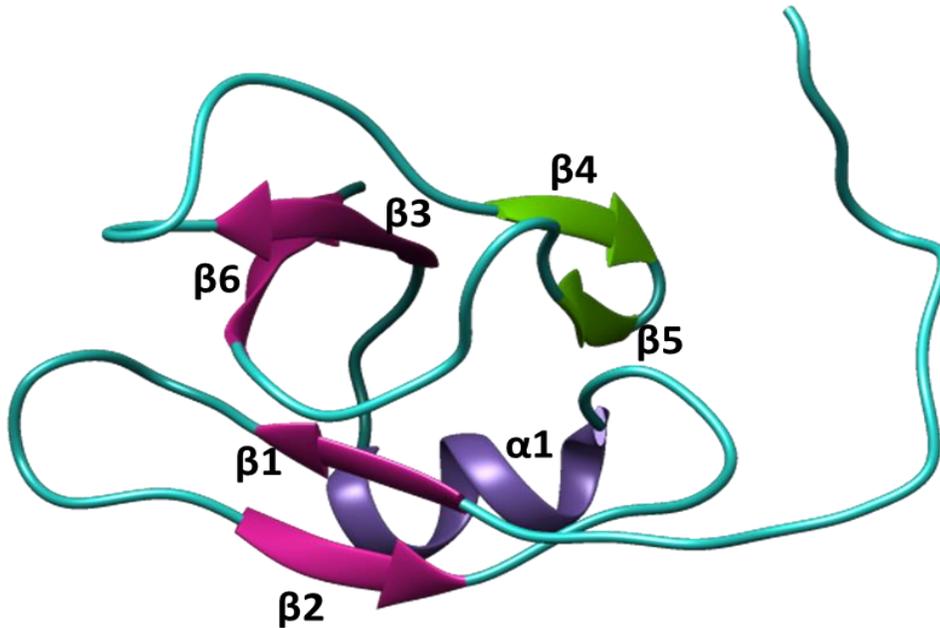
TDP-43 is a natively dimeric 414-residue protein that is encoded by the human *TARDBP* gene and present in inclusions of a wide variety of neurodegenerative disorders. The 43-kDa protein has three domains: a structured N-terminal domain, two RNA recognition motifs that bind RNA, and an unstructured C-terminal domain that mediates protein-protein interactions (Fig. 4A). All three domains are necessary for normal function of TDP-43 (Budini, Romano, Quadri, Buratti, & Baralle, 2015; Chiang et al., 2016; Kuo, Doudeva, Wang, Shen, & Yuan, 2009; Zhang et al., 2013).



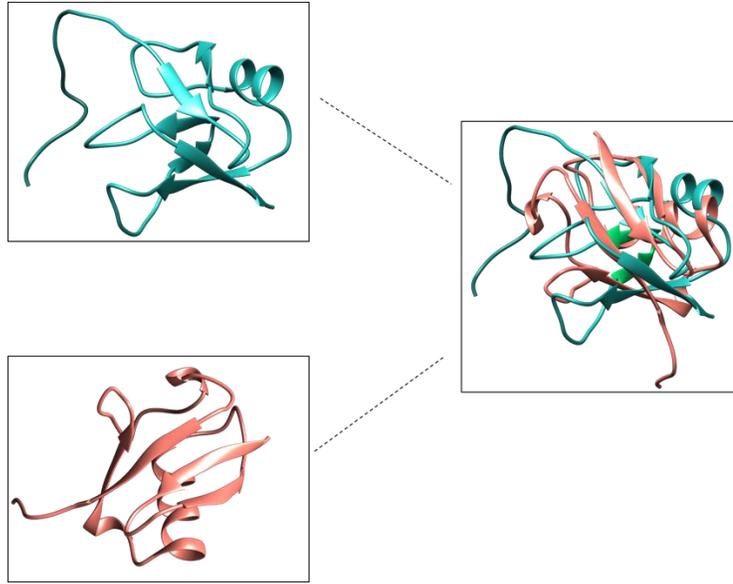
**Figure 4 Structure of TDP-43.** A) Combined structure of the TDP-43 protein. PDB files of each of the domains of TDP-43 were combined to create a proposed model for the full structure of TDP-43. Dotted lines represent the unknown residues that connect the structures. The positioning of the NTD relative to the RRMs was determined by data from small-angle X-ray scattering (Kuo et al., 2014). B) Secondary structure of the N-terminal is captured from NMR structure of the PDB file 2N4P and consists of residues 1-77. C) The flexible C-terminal tail is derived from NMR analysis PDB entry 2N3X. CTD spans residues 24-414 D) An illustrated domain analysis of the 414 amino acid protein with three domains: an N-Terminal, tandem RRMs, and a C-terminal. E) Secondary structure of RRM1, residues 106-175 bound to RNA derived from the PDB entry 4Y0F. F) Structure of RRM2 generated from the crystal structure of PDB entry 3D2W; consists of residues 193-257.

### 1.3.1 N-Terminal Domain

The highly conserved, 77 residue N-terminal domain (NTD) of TDP-43 plays an important role in TDP-43's aggregation pathology and is suggested to play an important role in TDP-43's physiology (Mompeán et al., 2016; Mompeán et al., 2017). The NTD has a unique structure with the monomer depicted as 2 sheets being comprised of 6  $\beta$ -strands and 1  $\alpha$ -helix. Sheet one is arranged with  $\beta 2$  strand antiparallel to  $\beta 1$ , which in turn is parallel to  $\beta 6$ , which is parallel to  $\beta 3$  (Fig. 5) having structural similarity to a ubiquitin like fold (Fig. 6). Sheet 2 is organized with  $\beta 4$  antiparallel to  $\beta 5$  (Fig. 5) having structural similarity to the C-terminal DIX domain of the protein axin-1 (Mompeán et al., 2016). The NTD of TDP-43 readily oligomerizes at high concentrations and is composed as a dimer interface in wild type TDP-43 (Chang et al., 2012). Hydrogen bonding and electrostatic interactions make up the intermolecular interface of the NTDs in

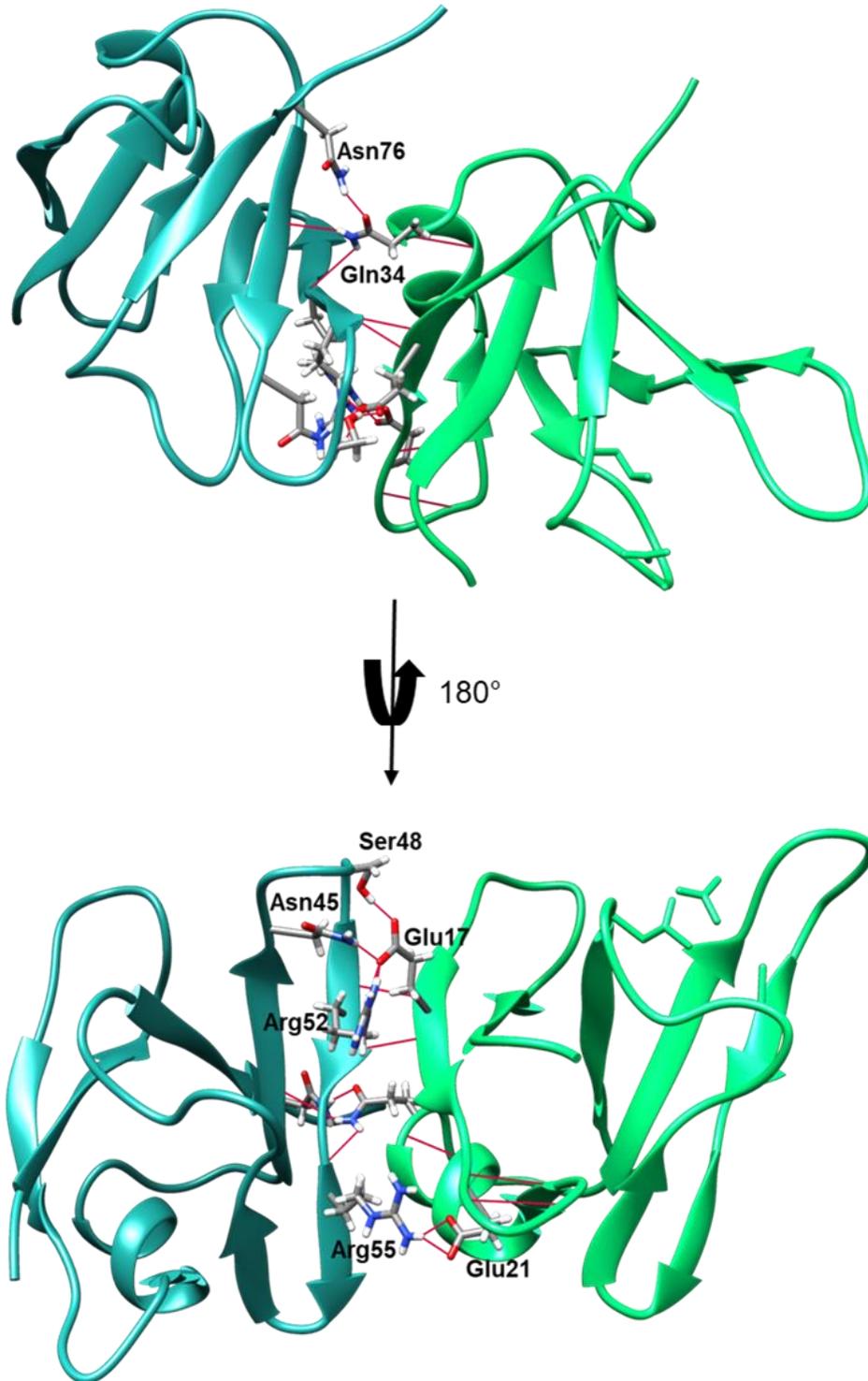


**Figure 5. NTD of TDP-43.** The NTD is composed of 6  $\beta$ -strands and 1  $\alpha$ -helix. The  $\beta$ -strands are separated into 2  $\beta$ -sheets. Sheet 1, colored pink, consists of  $\beta 1$ ,  $\beta 2$ ,  $\beta 3$ , and  $\beta 6$ . Sheet 2, green, contains  $\beta 4$  and  $\beta 5$ . The  $\alpha$ -helix is colored purple.



**Figure 6 NTD of TDP-43 adopts a ubiquitin-like fold.** Overlay of structures TDP-43 (teal) PDB: 2N4P and Ubiquitin (coral) PDB: 1TBE show that sheet 1 of TDP-43 shares structural similarities with the ubiquitin fold. Sheet 2 of TDP-43 does not share similarities with Ubiquitin.

TDP-43. The intermolecular interface of the TDP-43 NTD dimer is made up of an asymmetric unit arranged in a “head-to-tail” morphology mediated by hydrogen bonding and electrostatic interactions. The positively charged head interacts with the negatively charged tail of the second monomer. The positively charged head interacts with the negatively charged tail of the second monomer to Glu17 on  $\beta 2$  (tail); Arg55 on  $\beta 4$  (head) to Glu21 on  $\beta 1$ - $\alpha 1$  loop (tail); and Asn76 on  $\beta 5$  (head) to Gln34 on  $\alpha 1$  (tail). The Glu17 on the  $\beta 2$  strand of one monomer hydrogen bonds in a parallel manner to Met51 and Gly53 on the backbone of the  $\beta 4$  strand on the second monomer to create an intermolecular  $\beta$ -bridge. This interface is strengthened by the salt bridge created when the carboxyl of Glu17 hydrogen bonds with the hydroxyl of Ser48 and the carboxamide group of Asn45 in a site specific manner (Fig. 7)

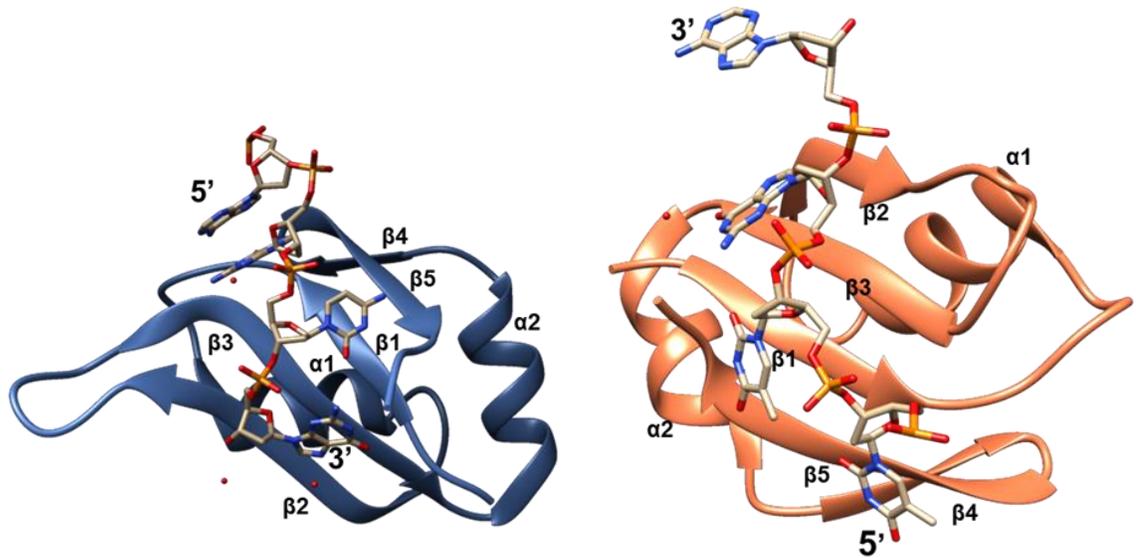


**Figure 7. Domain interaction of TDP-43 NTD Dimer.** Representation of two NTD monomers, NTD1 (Teal) and NTD2 (Green), in a dimer formation. Crystal structure of the NTD dimer (PDB:5MDI) reveal hydrogen bonds (Red dotted lines) making inter-molecular contacts at specific residues depicted as stick figures on the structure.

(Afroz et al., 2017). This thesis gives evidence of how a phosphomimetic mutation at Ser48 delays oligomerization of full length TDP-43 by disturbing NTD dimer interactions which has been previously described to interrupt self-association (A. Wang et al., 2018). Residues 77-108 of the NTD belongs to the nuclear localization signal (NLS) of TDP-43. This region serves as a tag for transport into the nucleus (Lange et al., 2007). In addition, this region is highly disordered and due to its cationic nature has the ability to bind nucleic acids (Mompeán et al., 2016). Unlike the CTD of TDP-43, the NTD fold is highly structured and very stable (Chang et al., 2012). Studies show that the NTD is necessary for physiological function and for TDP-43 aggregation; however, the method by which this domain directly contributes to the pathologic mechanism of TDP-43 aggregation is still being investigated (Mompean et al., 2016).

### *1.3.2 RNA Recognition Motifs*

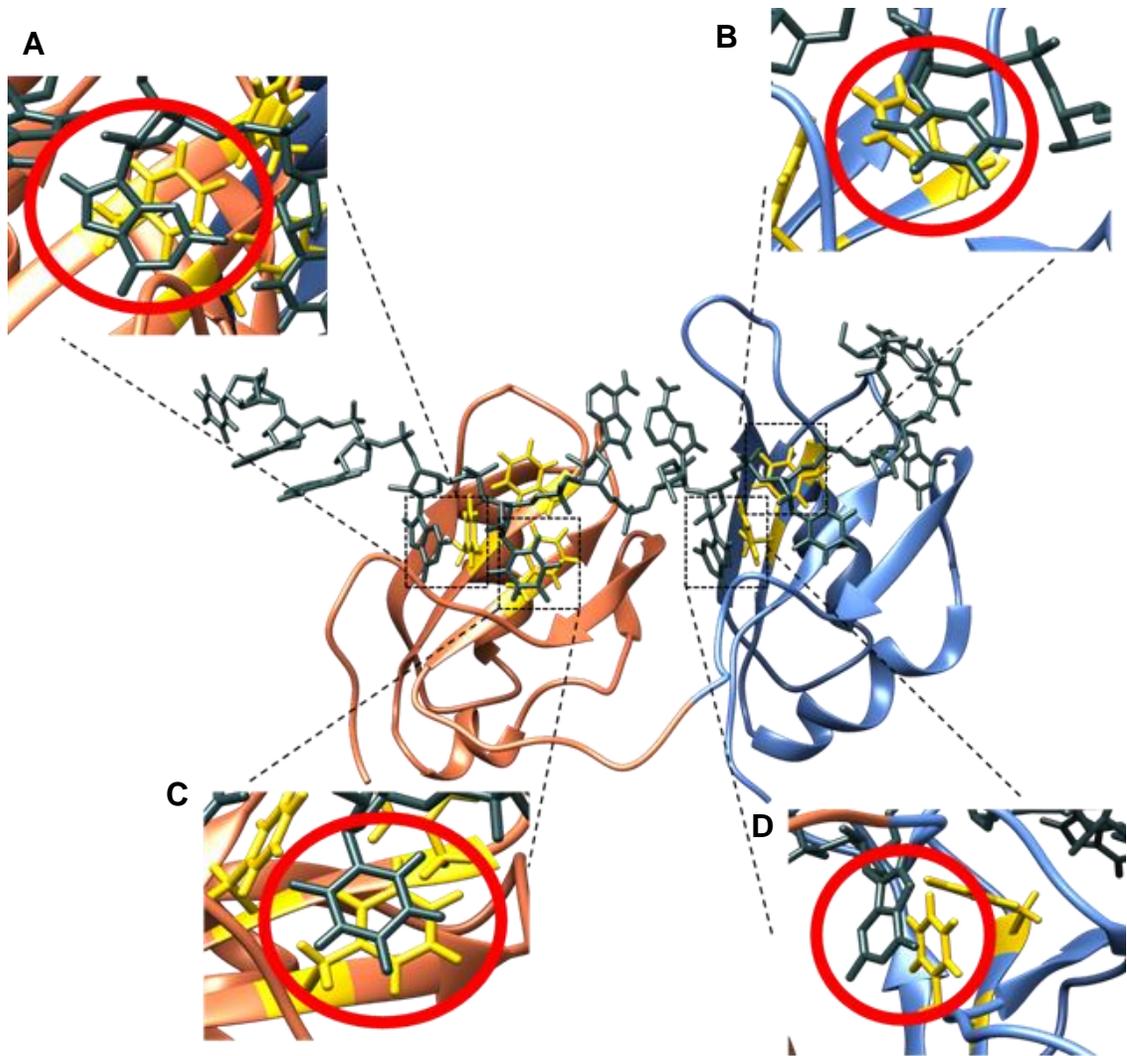
One of the most abundant eukaryotic protein domains is the RNA recognitions motif (RRM) (Cléry, Blatter, & Allain, 2008). The main functions of this protein domain are to mediate RNA binding and protein-protein interactions (Maris, Dominguez, & Allain, 2005). Typical RRM structures contain folds of a  $\beta$ -sheet packed against 2  $\alpha$ -helices. There are two highly conserved domains of 8 and 6 amino acid residues called ribonucleoprotein domain (RNP) 1 and RNP 2 respectively, that associate with single stranded nucleic acid through  $\pi$  stacking interactions to the flat side of a  $\beta$ -sheet. Each RRM has the ability to bind a minimum of 2 and a maximum of 8 nucleotides (Kuo, Chiang, Wang, Doudeva, & Yuan, 2014; Landsman, 1992).



**Figure 8. Individual RRMs.** Ribbon structures generated from the crystal structures of the individual RRMs with RNA, RRM1 (Blue, PDB: 4Y0F) and RRM2 (Orange, PDB: 3D2W), have topologies similar to that of typical RRM domains

TDP-43 forms a homodimer with 4 RRMs with each monomer containing 2 RRMs deemed RRM1, residues 106-175, and RRM2, residues 193-257 (Emanuele Buratti & Baralle, 2001). Crystal structures of RRM1 (PDB:4Y0F) (Chiang et al., 2016) and RRM2 (PDB:3D2W) (Kuo et al., 2009) reveal structures similar to that of typical RRMs with a 5 stranded  $\beta$ -sheet packed against 2  $\alpha$ -helices (Fig. 8) (A. Wang et al., 2018). Both RRMs have topologies of  $\beta 2$ - $\beta 3$ - $\beta 1$ - $\beta 5$ - $\beta 4$  with the conserved RNP1 and RNP2 located in  $\beta 1$  and  $\beta 3$  respectively. RRM1 has a longer loop between  $\beta 1$  and  $\beta 3$  making it slightly larger than RRM2 (Kuo et al., 2014).

TDP-43 has a specificity for poly-UG/TG repeats and will bind a minimum of 6 TG/UG repeats with RRM1 dominating the binding process by binding the majority of



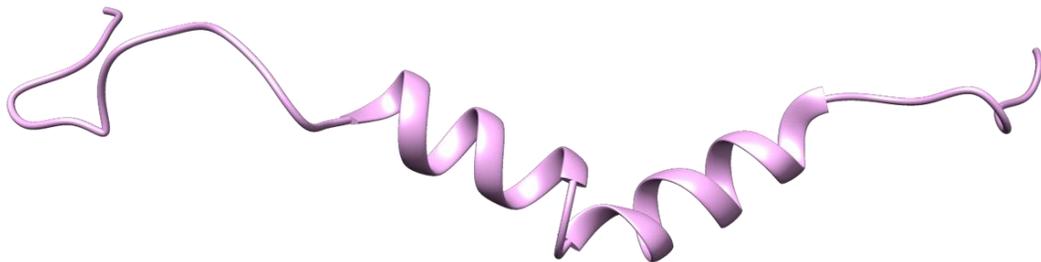
**Figure 9. TDP-43 RRM1s bind UG nucleotides.** Secondary structure of TDP-43 tandem RRM1s (RRM1: Blue; RRM2: Orange) with and AUG<sub>12</sub> RNA (5'-GUGUGAAUGAAU-3') bound (Dark Grey) (PDB: 4BS2) reveal  $\pi$  stacking interactions with Phenylalanine residues. A) Phe231  $\pi$ -stacks with G. B) Phe147  $\pi$ -stacks with U. C) Phe194  $\pi$ -stacks with U. D) Phe149  $\pi$ -stacks with G.

the nucleotides and RRM2 acting in more of a supporting role for nucleic acid binding (Furukawa et al., 2016). Phenylalanine (Phe) plays an important role in nucleic acid binding which is made apparent when mutations that change 5 Phe to Leucine (Leu) in the RRM1s inhibit nucleic acid binding (Gopal, Nirschl, Klinman, & Holzbaaur, 2017)

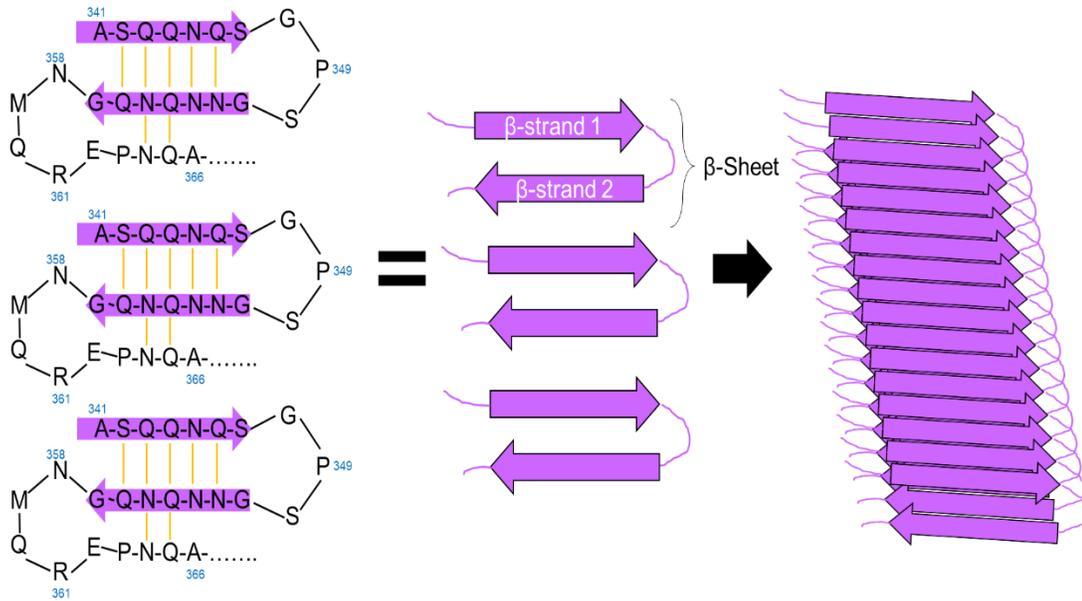
### 1.3.3 C-Terminal Domain

The C-terminal Domain (CTD) of TDP-43 is the most studied of the domains due to it containing the majority of disease related mutations in ALS (Conicella et al., 2017). Nuclear magnetic resonance spectroscopy (NMR) and molecular dynamic (MD) simulations have been used to determine the conformation of this domain because of its intrinsically disordered nature (Mompeán et al., 2015). The CTD has a low-complexity sequence, which is not uncommon for eukaryotic RNA binding proteins (Conicella et al., 2017). Only one section of the CTD, residues 319-341, have sequence conservation with vertebrates resulting in a fragmented conservation of this domain, however, the CTD has a 24% sequence identity with the N-terminus of the Sup35 yeast prion domain resulting in the CTD being “prion-like” (Lim et al., 2016).

Secondary structure predictions by NMR and MD simulations determine the CTD, spanning residues 274-414, to adopt a helix-turn-helix conformation with three defined domains: a glycine rich domain (GRR) which functions to recruit cellular factors that modulate cytoplasmic inclusion of TDP-43 (D’Ambrogio et al., 2009), an Asn/Gln rich motif which is capable of aggregation and has been seen to recruit normal TDP-43



**Figure 10. C-Terminal Domain.** The C-terminal domain of TDP-43 is highly disordered. Secondary structure analysis by NMR (PDB: 2N3X) reveals a prominent  $\alpha$ -helical domain.



**Figure 11. Illustration of proposed CTD aggregation by  $\beta$ -strand expansion modeled by MD Simulations.** The glycine-rich region of the CTD has been described by molecular dynamic simulations to form anti-parallel  $\beta$ -hairpins. 2  $\beta$ -strands form a single  $\beta$ -sheet. Aggregates are formed by interactions between  $\beta$ -sheets.

into aggregates (Mompeán et al., 2014), and a hydrophobic patch which has been proposed as an initiator for aggregation in mechanistic studies (Jiang et al., 2013). An  $\alpha$ -helical segment from 321-330 and a  $\beta$ -hairpin structure with a turn separating two Asn/Gln regions from residues 341-357 are two sections that have been defined by NMR and MD simulations. It has also been determined that residues 358-367 form a defined  $\beta$ -turn, which is followed by a random coil which makes up the flexible tail (Mompeán et al., 2015)(Fig. 10).

The CTD of TDP-43 is highly aggregation prone and there have been many experiments done on this domain to determine its function in TDP-43 pathology. Mechanism prediction experiments show that the residues from 311-360, which normally

adopted an  $\alpha$ -helix structure in the CTD goes through a structural transformation that triggers unfolding to a  $\beta$ -sheet rich structure, which spontaneously associates and can serve as a “nucleus” for aggregation (Jiang et al., 2013). MD simulations show an amyloid like  $\beta$ -topology where the  $\beta$ -hairpins laterally associates in parallel contributing 2  $\beta$ -strands to one  $\beta$ -sheet (Mompeán et al., 2015)(Fig.11). It is proposed that oligomerization occurs through this  $\beta$ -strand expansion (Mompeán et al., 2014).

As mentioned earlier, the majority of disease mutations in ALS are found in the CTD of TDP-43. The specific mutation Q331K, where the 331 glutamate is mutated to a lysine is found in the middle of the  $\alpha$ -helical structure. This mutation disrupts the  $\alpha$ -helix structure and is toxic to the cell (Arnold et al., 2013; Conicella et al., 2017).

#### *1.3.4 Domain Contributions to Aggregation*

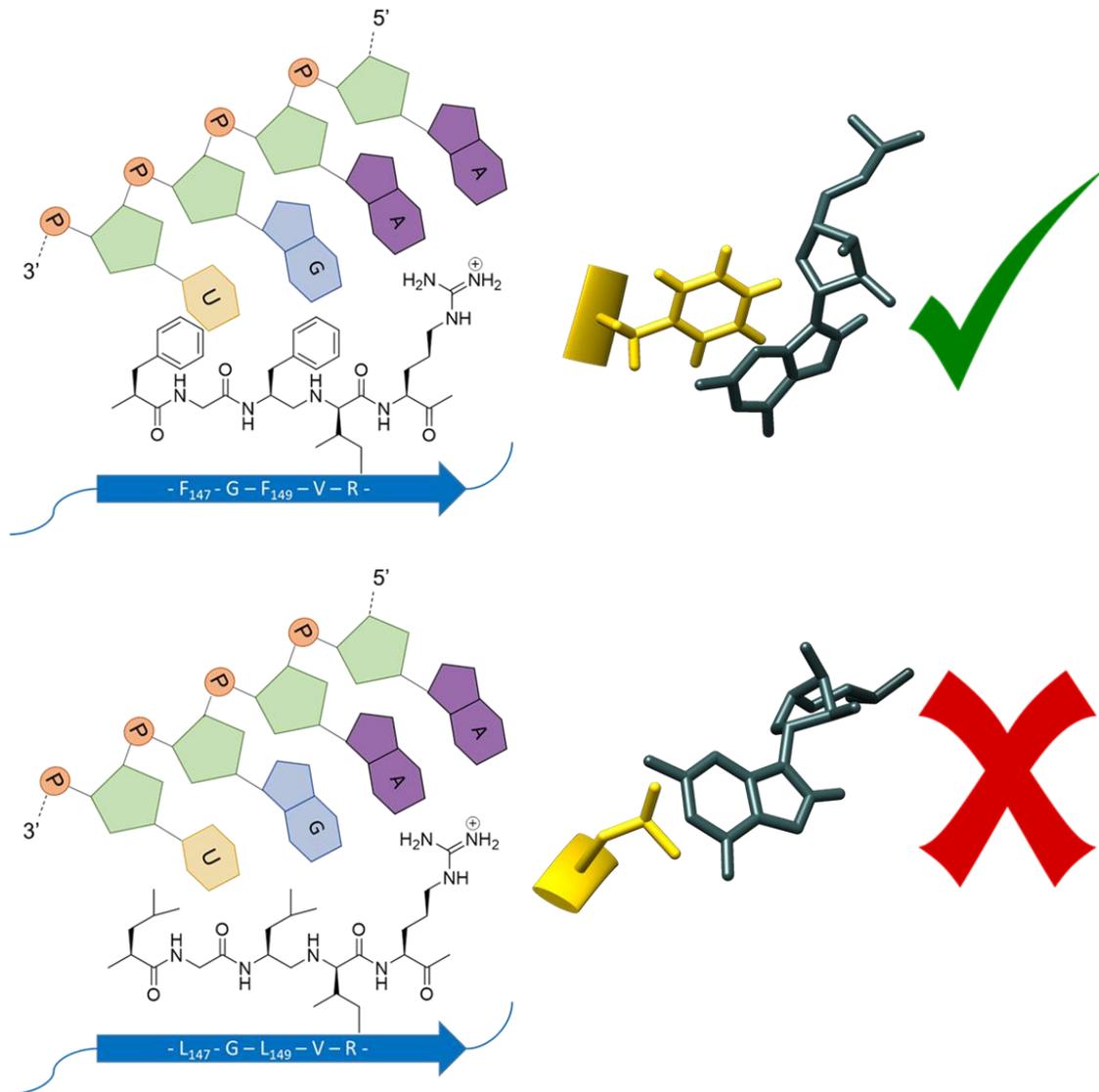
TDP-43 is known to have aggregation prone properties, however, the molecular mechanisms by which TDP-43 aggregates form is still largely unknown (Yang et al., 2010). Studies have identified regions of the different domains of TDP-43 to have individual factors that can contribute to aggregation of the full length protein. The NTD of TDP-43 has been shown to oligomerize in a concentration dependent manner with circular dichroism data (CD) suggesting that the oligomerization is mediated by the  $\beta$ -sheets on the surface of the NTD (Chang et al., 2012). A short 10 residue region on RRM2 has been shown to have polymorphic capabilities that are apparent in pathological amyloid aggregation (Guenther et al., 2018). In addition to the MD simulations that hypothesize  $\beta$ -strand expansion as an inducer for aggregation, the CTD of TDP-43 has

glycine and Q/N-rich regions which have been shown to form fibrillar amyloid-like aggregates *in vivo* (Prasad, Sivalingam, Bharathi, Girdhar, & Patel, 2018).

#### *1.4 Nucleic Acid Binding*

TDP-43 is a nucleic acid binding protein known to bind specifically to poly-UG repeats by its RRM1s in the 3' to 5' direction. (E Buratti et al., 2001; Sun & Chakrabartty, 2017). As stated previously, RRM1 has more involvement in the nucleic acid binding process while RRM2 is more of a support system. This is confirmed in pull down experiments where, when independent of each other, RRM2 did not bind nucleic acid while RRM1 had high affinity for single stranded nucleic acid. When in a complex, the RRM1s bound UG/TG repeats with high affinity suggesting that RRM2 plays an important role in sequence specificity (Furukawa et al., 2016).

A nucleic acid binding model suggests that in the presence of RNA the RRM1 will bind and dissociate from the UG nucleotides with  $K_d$  values of ~110 nM (Bhardwaj, Myers, Buratti, & Baralle, 2013). This supports the need for longer oligonucleotide repeat sequences so that the TDP-43 RRM1 can continuously bind the subsequent U and G in the sequence. A crystal structure of RRM1 has revealed the specific interactions of nucleic acid with TDP-43 where a T-G dinucleotide is recognized by Loop 1 and a G-T-G sequence interacts with a  $\beta$ -sheet (Kuo et al., 2014). Two aromatic residues in RRM1, Phe147 and Phe149, are known to make  $\pi$  stacking interactions with uracil (U) and guanidine (G) on RNA (Ayala et al., 2005). As mentioned previously, when these aromatic residues are mutated to aliphatic leucine they are no longer able to participate in stacking interactions with RNA, and binding is diminished (Cohen et al., 2015) (Fig.12).



**Figure 12. Phe $\rightarrow$ Leu mutation in RRM1 disrupts RNA binding.** Phe147 and Phe149 form  $\pi$ -tacking interactions with U and G in RNA (Kuo et al., 2014). When Phenylalanine is mutated to a Leucine binding of RNA can no longer occur.

#### *1.4.1 Uncertainty in RNA Binding Models*

TDP-43 is known to bind UG/TG repeats of RNA/DNA through its RRM's however, other probable processes for RNA binding with TDP-43 have been discussed. Anisotropy measurements monitoring binding affinity of the NTD in complex with the RRM's had a 3-fold higher binding affinity than the RRM's alone suggesting that the NTD contributes to nucleic acid binding in the full length TDP-43 (Chang et al., 2012). The NLS of TDP-43 is also cationic and has the ability to bind TT and TG repeat DNA (Mompean et al., 2016). Another study has shown that direct binding of RNA to the glycine rich region in the CTD of TDP-43 was able to prevent aggregation of cleaved C-terminal fragments of TDP-43 (Kitamura, Shibasaki, Takeda, Suno, & Kinjo, 2018). A better understanding of the binding models of RNA/DNA to TDP-43 could aid in identifying possible interventions to the neurodegenerative diseases in which aggregated TDP-43 is present.

#### *1.5 Ribonucleoprotein Complexes*

Ribonucleoprotein (RNP) Complexes form when RNA and RNA-binding proteins associate. These RNA-protein complexes have been shown to increase the solubility of aggregation prone proteins by improving protein-folding pathways (Choi et al., 2008). TDP-43 is a RNP known to have a high specificity for poly TG/UG repeats. While TDP-43 is known to aggregate in a variety of environmental stressors, nucleic acid binding has been shown to suppress aggregation of TDP-43 in *in vitro* heat stress conditions (Huang et al., 2013). In an experiment testing the propensity of nucleic acid to inhibit aggregation, the presence of TG<sub>12</sub> repeats showed substantially less aggregated species

than AC<sub>12</sub> repeat and a no RNA control species (Sun, Arslan, Won, Yip, & Chakrabartty, 2014).

Aggregated TDP-43 has been found in inclusions located in the brain matter of patients with CTE. It has been speculated that the violent shaking of neuronal cells may serve as an inducer for TDP-43 aggregation in these inclusions (McKee & Gavett, 2010). Given that TG<sub>12</sub> repeats attenuate TDP-43 aggregation in heat stress conditions we speculate that the addition of TG<sub>12</sub> repeats to agitated TDP-43 may attenuate aggregation as well. In this study, with turbidity measurements using agitation induced aggregation, fluorescence assays, and sedimentation analysis, we observe the aggregation potential of dimeric TDP-43 and investigate the propensity of DNA/RNA sequences to inhibit TDP-43 aggregation when shaken vigorously.

## CHAPTER 2: MATERIALS AND METHODS

### *2.1 LIC Cloning*

Human wild type (WT) TDP43, Q331K, and 5F→L ORFs were cloned into the pET His6 GFP TEV LIC cloning vector (1GFP) (Addgene) or pET Biotin His6 mCitrine LIC cloning vector (mCitrine) using a ligation independent cloning protocol (LIC) (QB Berkley) by a previous graduate student. LIC cloning is based on making DNA strands single stranded and complimentary to each other. Primers for 1GFP and mCitrine cloning were designed as described by QB3 MacroLab LIC Expression Vectors cloning manual to anneal to the desired digestion site at 55°C and to include a C terminal stop codon in the amplified ORF. The digestion site was created by linearizing the 1GFP and mCitrine vectors with the enzyme SspI at 37°C for 1 hour. Primers were created by designing a 3' end complementary to the insert target 20-25 bp long with a T<sub>m</sub> of ~60 °C ending with a 3' G or C (Table S1). The TDP43 (Addgene; #27458), Q331K (Addgene; #27461) or 5F→L(Addgene;#29608) GFP or mCitrine plasmid construct was transformed into Rosetta or BL21 competent cells, plated on LB plates containing 50ug/mL Kanamycin (GoldBio) and 25ug/mL Chloramphenicol (GoldBio) for Rosetta's or 50ug/mL Kanamycin (GoldBio) for BL21's, and incubated at 37°C overnight.

## *2.2 Site Directed Mutagenesis*

Q5<sup>®</sup> Site Directed Mutagenesis Kit (New England Biolabs, Ipswich, MA, USA) was used to substitute GAG (glutamate) for TCT (serine) in the pET His6 GFP-TEV-TDP-43. Primers were designed using NEBase changer from New England Biolabs (Ipswich, MA, www.neb.com). The substitution from serine to glutamate was incorporated in the center of the forward primer with nucleotides of complementarity downstream from the mutation point GAATCCAGTGGAGCAGTGTATGAGAGGTG - 3'. The reverse primer was designed with the 5' ends of the two primers annealing back to back: 5'- CTGTAGCGAAGCCCACAC -3'. Thus the entire pET His6 GFP-TEV-TDP-43 plasmid was PCR amplified and transferred to a Kinase, Ligase, Dpn1 (KLD) reaction. The newly synthesized GFP-TEV-S48E plasmid construct was then transformed into BL21 competent cells, plated on LB plates containing 50ug/mL Kanamycin, and incubated at 37°C overnight. Mutations were verified by DNA sequencing (Quintara Bio). Expression and purification was performed using the methods described below.

## *2.3 Expression and Purification*

Starter cultures were prepared by inoculating single bacterial colonies into 50mL of LB medium containing appropriate antibiotic and incubated at 37°C overnight. The overnight culture was used to inoculate 1L of LB medium containing the appropriate antibiotic. The cells were then grown to an OD<sub>600</sub> ~ 0.4, cooled to 19°C and induced by a final concentration of 1 mM IPTG (isopropyl-β-thiogalactopyranoside) and then incubated for 6 hours. The cell culture was centrifuged at 8000 RCF for 30 minutes. The pellet was suspended in a minimum amount of Lysis Buffer (40mM HEPES, 300mM

NaCl, 1mM MgCl<sub>2</sub>, 50mM Imidazole, 5mM βMe, 10% Glycerol, and complete EDTA free protease inhibitor, pH 7.5). CHAPS (1%), DNase (20μg/μL), and Lysozyme (0.25mg/mL) were added to the suspension and incubated with stirring at 4°C for 1 hour. The suspension was centrifuged at 31000 RCF for 75 minutes and the lysate was filtered through a 0.22μm bottle top filter (Genesee). Lysate was then loaded using a sample pump onto an 8mL nickel immobilized metal affinity chromatography (IMAC) column (Biorad) in an NGC System (BioRad) at 4°C. The column was washed with 5 column volumes (CV) of Lysis Wash Buffer followed by a modified Elution Buffer (40mM HEPES, 300mM NaCl, 1mM MgCl<sub>2</sub>, 500mM Imidazole, 5mM βMe, and 10% Glycerol, pH 7.5). The proteinaceous fractions were collected and immediately loaded onto a pre-equilibrated Superdex 200 16/600 Size exclusion Column (GE Life Sciences) to get a homogenous protein sample. Protein was eluted using an SEC Buffer (40mM HEPES, 300mM NaCl, and 5mM DTT, pH 7.5). Purified fractions were collected and immediately used for analysis.

#### *2.4 Denaturing Expression and Purification*

Starter cultures were made by inoculating single colonies into 50mL of LB medium containing appropriate antibiotic and incubated at 37°C overnight. The overnight culture was used to inoculate 1L of LB medium spiked with appropriate antibiotic. The cells were then grown to an OD<sub>600</sub> ~ 0.4, cooled to 19°C, induced by a final concentration of 1 mM IPTG (isopropyl-β-thiogalactopyranoside), and then incubated for 6 hours. The cell culture was centrifuged at 8000 RCF for 30 minutes. The pellet was suspended in a minimum amount of Lysis Buffer (40mM HEPES, 300mM NaCl, 1mM MgCl<sub>2</sub>, 50mM

Imidazole, 5mM  $\beta$ Me, 10% Glycerol, and complete EDTA free protease inhibitor, pH 7.5). CHAPS (1%), DNase (20 $\mu$ g/ $\mu$ L), and Lysozyme (0.25mg/mL) were added to the suspension and incubated with stirring at 4°C for 1 hour. The suspension was centrifuged at 31000 RCF for 75 minutes and the lysate was filtered through a 0.22 $\mu$ M bottle top filter (Genesee). Lysate was then loaded onto an 8mL nickel Immobilized metal affinity chromatography (IMAC) column (Biorad) in an NGC System (BioRad). The column was washed with 5 column volumes (CV) of Lysis Wash Buffer, then washed with a Denaturing Wash Buffer (40mM HEPES, 300mM NaCl, 1mM MgCl<sub>2</sub>, 500mM Imidazole, 5mM  $\beta$ Me, 10% Glycerol, and 6M Urea; pH 7.5) followed by a modified Elution Buffer (40mM HEPES, 300mM NaCl, 1mM MgCl<sub>2</sub>, 500mM Imidazole, 5mM  $\beta$ Me, 10% Glycerol, and 6M Urea; pH 7.5). The proteinaceous fractions were collected and immediately loaded onto a Superdex 200 16/600 Size exclusion Column (GE Life Sciences). Protein was eluted using an SEC Buffer (40mM HEPES, 300mM NaCl, 6M Urea, and 5mM DTT; pH 7.5). Purified fractions were collected and immediately used for analysis.

### *2.5 In vitro Aggregation Studies*

Purified protein typically eluted off the size exclusion column at  $\sim$ 7 $\mu$ M and was further diluted to a final concentration of  $\sim$ 3 $\mu$ M in a buffer containing 25mM HEPES, 170mM NaCl, and 3mM DTT at pH 7.5. Diluted protein was incubated at 25°C with agitation at 1400 RPM in an Eppendorf Thermomixer for 4 hours in the presence or absence of oligonucleotides. Turbidity measurements taken using a Tecan Infinite®

M1000 plate reader at an absorbance of 395nm were used to assess aggregation.

Measurements were taken at 25°C.

### *2.6 Fluorescence Assays*

Purified protein samples (5 $\mu$ M) in SEC buffer (40mM HEPES, 300 mM NaCl, and 5mM DTT; pH 7.5) were subjected to fluorescence intensity scans with GFP excitation at 488nm in the presence or absence of oligonucleotide (IDT). Emission spectra were measured from 500-700nm using a Tecan Infinite® M1000 plate reader. Protein was then subjected to agitation as described above. After agitation samples were spun down at max speed (21120 RCF) to pellet aggregated protein. Fluorescence spectra were measured on the supernatant as described above.

### *2.7 Sedimentation Analysis of Aggregation Experiments*

We performed a sedimentation analysis as was described in a previous study (Johnson et al., 2009). After *in vitro* agitation experiments the insoluble fraction (pellet) was separated soluble fraction (Supernatant) by centrifugation at 21130 x g for 60 min at 25°C. A denaturing SDS PAGE gel was used to analyze protein species in the supernatant and pellet. To accomplish this supernatant was decanted and pellet resuspended in equal volume ( ~50 $\mu$ L) of Laemmli Loading Buffer (BioRad). 15 $\mu$ L of both supernatant and pellet fractions were loaded onto an SDS PAGE gel and analyzed for comparison.

### *2.8 Dynamic Light Scattering*

Dynamic light scattering (DLS) measurements of the hydrodynamic radius ( $R_h$ ) were taken at 25°C with a DynaPro® Plate Reader II (Wyatt Technologies) using a 96

well half area plate (Corning). Proteinaceous fractions ( $\sim 7\mu\text{M}$ ) in an in SEC buffer (40mM HEPES, 300 mM NaCl, and 5mM DTT; pH 7.5) were filtered through 0.22 $\mu\text{M}$  PES filters prior to analysis to remove large species. Regularization analysis was determined using DYNAMICS (Wyatt) software with 5 consecutive measurements taken over a 5 second time frame. Autocorrelation decay curves were used to calculate the particle translation diffusion coefficient. The Stokes-Einstein equation ( $D_t = \frac{kT}{6\pi\eta R_h}$ ;  $D_t$  is the diffusion coefficient,  $k$  is the Boltzmann constant,  $T$  is absolute temperature in Kelvin,  $\eta$  is the viscosity of the solvent, and  $R_h$  is the hydrodynamic radius.) was then applied to determine the  $R_h$  of the particles in solution and the DYNAMICS software was used to create histograms of % mass vs.  $R_h$  for analysis.

### *2.9 Determining Concentration of TDP-43*

A Pierce™ BCA Protein Assay kit was used for the total protein quantification. Following the protocol described in the Pierce™ BCA Protein Assay manual, 2mg/mL Bovine Serum Albumin (BSA) protein standards (Thermo Scientific) were serially diluted to a working range which were then used to create a standard curve to determine protein concentration of WT TDP-43 and mutants. To 25 $\mu\text{L}$  of BSA standard of a known concentration, 175  $\mu\text{L}$  of working reagent (WR) was added to create a final concentration of 1:8 (v/v) protein to working reagent in a 96 well plate (Corning). The same was carried out for the unknown concentration of TDP-43 or mutant. The solution was thoroughly mixed for 30 seconds and immediately incubated at 37°C for 30 minutes. After incubation the solution was cooled to room temperature (RT) and the absorbance was

measured at 562nm. The resulting BSA standard curve was used to extrapolate protein concentrations of the samples.

### *2.10 AUG12 RNA*

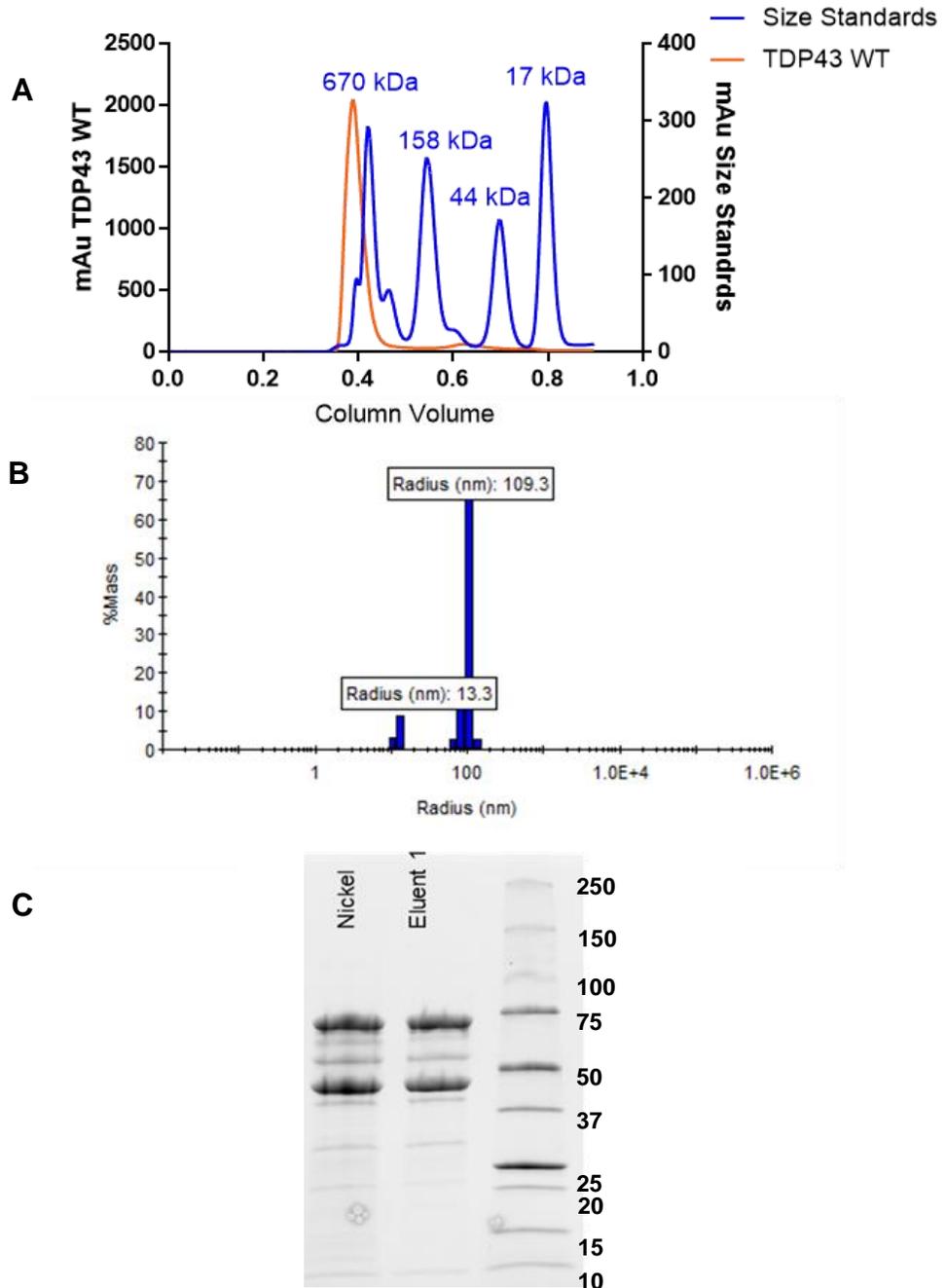
The 12-mer RNA in this study was transcribed using a recently published protocol (Chapman, Moon, Wilusz, & Kieft, 2014). Oligomers purchased from IDT containing a T7 promoter site were annealed together by heating to 95°C and snap cooled to 4°C for transcription by T7 RNA Polymerase (Table S2). Following transcription RNA was purified using a PLRP-S Chromatography Column on an Infinity II HPLC system (Agilent). Purity of RNA was analyzed using 10% dPAGE and concentration was determined on a Nanodrop One<sup>C</sup> (Thermo Scientific). Due to the small size of the RNA buffer exchange and desalting was accomplished by ethanol precipitation where 1 volume of RNA solution was added to 3 volumes of 100% ethanol and 1/10 volume of sodium acetate (NaOAc) and precipitated at -80°C overnight. The precipitated solution was then spun at >21,000 x g for 1 hour at 4°C. The supernatant was decanted and the pellet was suspended in minimal amount of HyPure™ Molecular Biology Grade water (GE Life Sciences) flash frozen in liquid nitrogen aliquoted and stored at -80°C.

## CHAPTER 3: RESULTS

### 3.1 WT TDP-43 Readily Forms Aggregates

We set out to obtain purified monomeric TDP-43 protein in order to complete *in vitro* aggregation experiments. TDP-43 protein has been reported to be inherently aggregation prone in previous *in vitro* studies (Johnson et al., 2009; Sun et al., 2014). The TDP-43 open reading frame (ORF) gene was cloned into an expression vector containing an N-Terminal monomeric enhanced green fluorescent protein (mEGFP) tag to act as a fluorescent probe. mEGFP is a derivative of green fluorescent protein which fluoresces more intensely and is less susceptible to unfolding (Gambotto et al., 2000). This construct was expressed in *Escherichia coli* using buffers adapted from previous studies (Sun et al., 2014).

The quality of protein purification was assessed using Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) gels, size exclusion chromatography (SEC) and dynamic light scattering (DLS) measurements. SDS-PAGE of the eluted sample shows the presence of a band at 75kDa which agrees with the size of the recombinant GFP-TDP-43 monomer (Fig. 13C). However, size exclusion chromatography shows the WT protein eluting in the void volume of the column



**Figure 13. Characterization of WT GFP-TDP-43.** A) Purification and analysis by SEC. Overlay of GFP-TDP-43 purified under standard conditions with Size Standards (BioRad). The single elution peak before 0.4CV reveals an aggregated species B) Size determination by DLS using  $\sim 7\mu\text{M}$  protein confirms large species.  $R_h$  is calculated by %mass using DYNAMICS software (Wyatt). C) SDS-PAGE of purified GFP-TDP-43. Lanes contain eluted fractions from Nickel affinity purification (Nickel) and SEC purification (Eluent 1) compared against Precision Plus Protein Ladder (Bio-Rad). Samples were run on a 4-20% polyacrylamide gel for analysis.

suggesting an aggregated species larger than 670kDa. This was assessed by overlaying elution peaks of GFP-TDP-43 with BioRad size standards (Fig. 13A).

DLS was used to assess the particle size of the species in solution by measuring the translational diffusion coefficient ( $D_t$ ) from which a hydrodynamic radius can be calculated using the Stokes-Einstein relation (See materials and methods). The  $D_t$  is calculated from an autocorrelation function which plots the average overall intensity of scattered light within a given time interval. DLS measurements give two species. One species has a radius of 13.3 nm which constituted 12% of the sample by mass and the second species is 109.3nm constituting 88% of the species by mass (Fig. 13B). Regarding absolute stoichiometry's, DLS measures the diffusion of particles in solution and does not deliver a proper molar mass. Molar mass is estimated by the DYNAMICS software using models that assume relationships between  $R_h$  and mass ( $M(r) \propto r^3$ ) where the DYNAMICS software assumes a spherical shape for the  $R_h$ . Conversion of radius to mass for the first species gives an estimated mass of 1428 kDa and the second species 198338 kDa. These results are inconsistent with previous studies that were able to purify the dimer from nickel affinity purification methods (Sun et al., 2014).

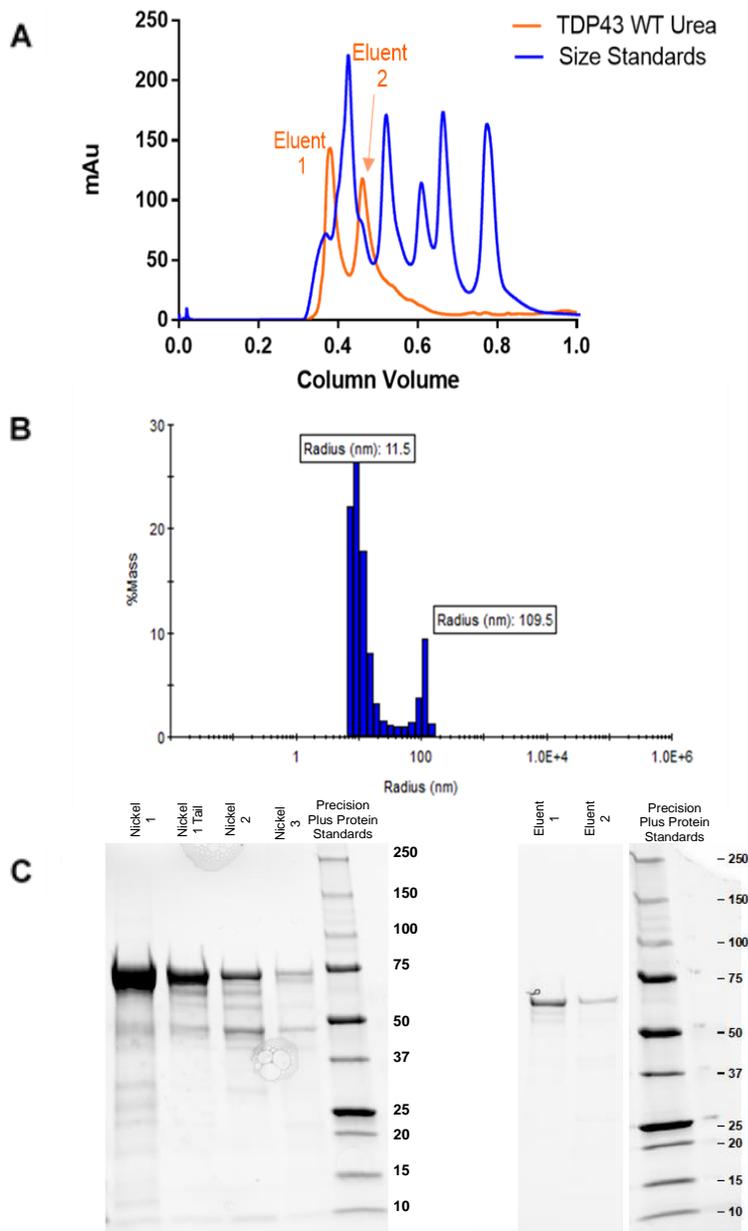
TDP-43 was also cloned into a vector containing a C-terminal mCitirine (mCit) fluorescent tag. mCit is a variant of YFP that has been modified to be more highly expressed (Griesbeck, Baird, Campbell, Zacharias, & Tsien, 2001). Purification of the TDP-43-mCit vector was almost identical to that of the GFP-TDP-43 variant. SDS-PAGE confirms the presence of the protein and DLS measurements and SEC analysis both confirm the presence of a multimer.

Our results suggest that purification of the WT GFP-TDP-43 plasmid with the N-terminal EGFP and the TDP-43-mCit plasmid with a C-terminal mCitrine in adapted buffers does not yield a dimeric TDP-43 protein.

### *3.2 Urea Denaturing Purification yields an aggregated TDP-43*

Difficulties in purifying full length monomeric TDP-43 have been previously reported (A. Wang et al., 2018). To improve purification, aggregated proteins have been unfolded in urea then subsequently refolded in a buffer, absent of urea back to their active form (Yamaguchi & Miyazaki, 2014). Denaturing purification of WT TDP-43 using 6M urea in wash and elution steps was used after native/ non-denaturing purification failed to produce a monomeric protein species for further analysis. The GFP-TDP-43 construct was expressed in *Escherichia coli* using lysis wash and elution buffers, outlined in the methods section, supplemented with 6M urea during purification. Initial purification methods where modified lysis wash buffer with 6M urea was used to suspend cell pellet and equilibrate the Ni/NTA column resulted in no protein binding to the column (Fig. S1).

Alternatively, resuspension of the cell pellet and equilibration of the Ni/NTA column was performed in lysis wash buffer without urea with subsequent purification steps completed with the modified lysis wash buffer with 6M urea. A single species at ~75kDa was identified by SDS-PAGE confirming the TDP-43 was present in the eluent (Fig 14C). SEC purification generated a single species that eluted in the void volume (Fig 14A). DLS measurements showed a multimeric species that constituted 95% of the



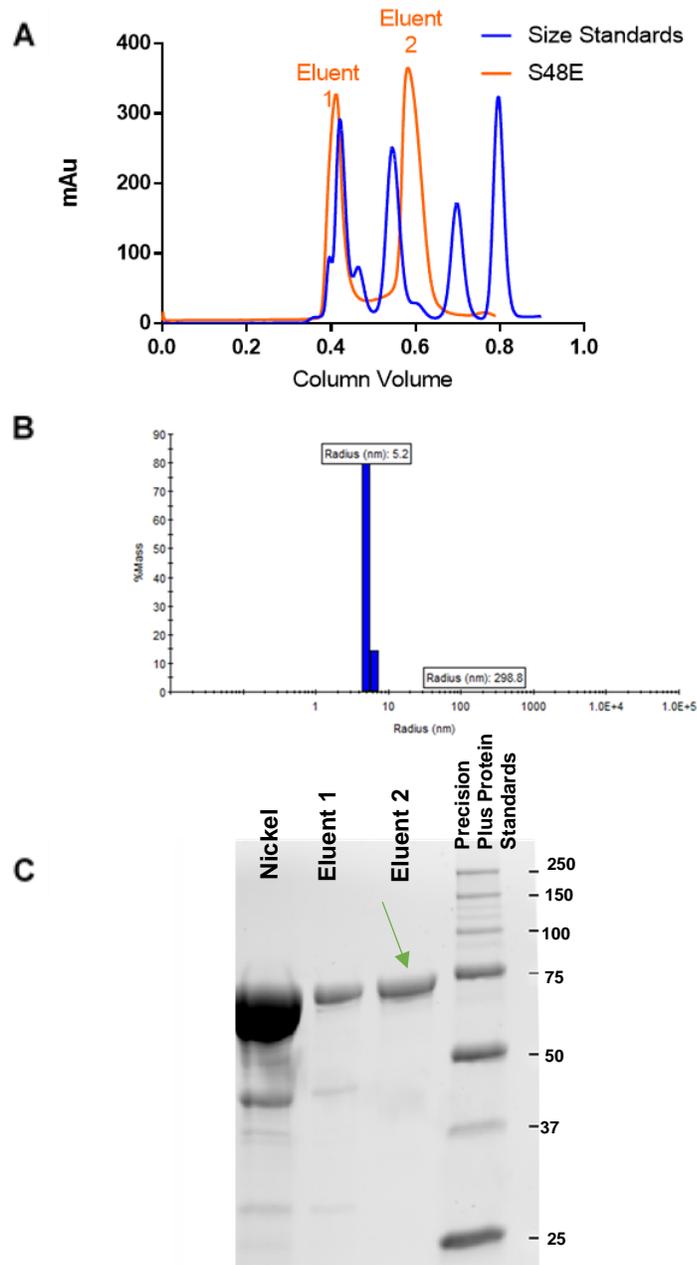
**Figure 14. Characterization of GFP-TDP-43 in a native purification followed by Urea.** A) Size exclusion chromatogram of TDP-43 WT Urea (orange) and size standards (blue). The short retention time of “eluent1” in TDP-43 WT Urea is consistent with a large aggregate. Retention time of “eluent 2” is smaller than the oligomer in “eluent 1” but still a larger oligomer than desired. B) Size determination by DLS using ~7 $\mu$ M of “eluent 2” confirms large species.  $R_h$  is calculated by mass with DYNAMICS software (Wyatt). C) SDS-PAGE of purified GFP-TDP-43. Lanes contain eluted fractions from Nickel affinity purification (Nickel 1-3) and SEC purification (Eluent 1-2) compared against Precision Plus Protein Ladder (Bio-Rad). Samples were run on a 4-20% polyacrylamide gel for analysis.

sample by mass with a radius of 11.5 nm (Fig 14B). This verifies the species that eluted was aggregated WT GFP-TDP-43. Our data present WT GFP-TDP-43 as a urea soluble multimer.

### *3.3 A S48E Phosphomimic Mutant Maintains Dimer Conformation after Purification*

Recent studies have identified a residue in the NTD involved in post-translational modification which, when phosphorylated, disrupted NTD self-assembly (A. Wang et al., 2018). Previous reports confirm that the NTD mediates TDP-43 oligomerization (Afroz et al., 2017). We corroborate the results that the phosphomimetic mutation, S48E, would favor in producing a species viable for aggregation studies.

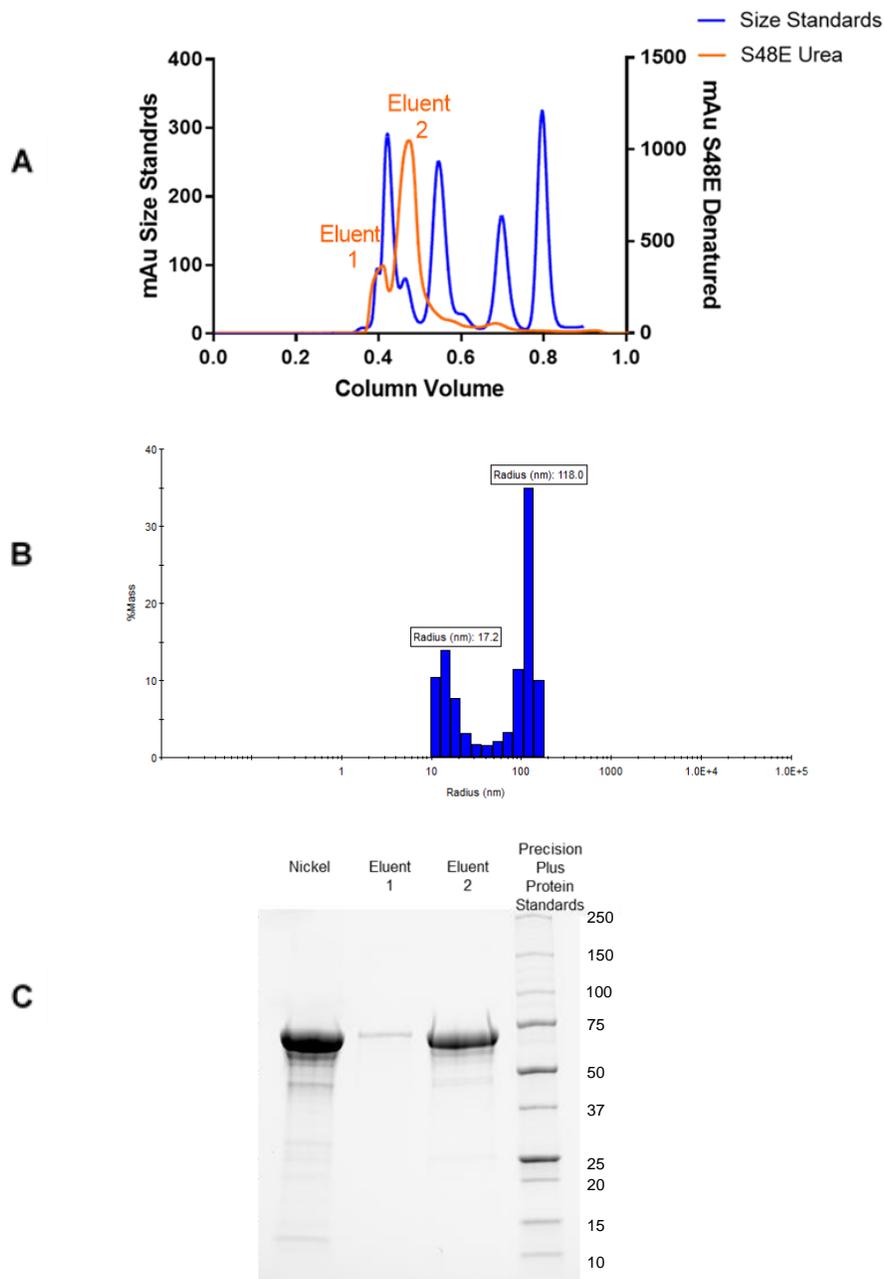
Using site directed mutagenesis, we created a mutated TDP-43 construct with a glutamate substituted for serine at the 48th residue in the NTD, S48E. The new construct, GFP-S48E-TDP-43, was expressed and purified under standard nondenaturing conditions. SEC purification resulted in two species, an aggregated species eluting in the void volume and another species eluted at ~0.6 CV (eluent #2) (Fig 15A). This second species was of extreme interest as it eluted with a peak consistent with the retention time of  $\gamma$ -globulin (bovine) in the BioRad Gel Filtration Size Standards, which has a molecular weight of 158kDa which is consistent with the dimer of TDP-43 (Fig 13A).



**Figure 15. Characterization of GFP-S48E-TDP-43.** A) Size exclusion chromatogram of S48E (orange) and size standards (blue). The short retention time of “eluent1” in S48E is consistent with a large aggregate. “Eluent 2” had a symmetric peak at a longer retention time consistent with a species closer to that of a dimer. B) Size determination by DLS using ~7uM of “eluent 2” sample. Rh is calculated by mass with DYNAMICS software (Wyatt). C) SDS-PAGE of purified GFP-TDP-43. Lanes contain eluted fractions from Nickel affinity purification (Nickel) and SEC purification (Eluent 1-2) compared against Precision plus Protein Ladder (Bio-Rad). Samples were run on a 4-20% polyacrylamide gel for analysis. Green arrow depicts purified dimer species.

Analysis by DLS identified the second species with 96% of the sample by mass to have a radius of 5.2 nm (Fig 15B). Conversion of radius to molecular weight gives an approximate mass of 160 kDa. SDS-PAGE revealed a single species at 73kDa confirming “elution #2” to be the GFP-S48E-TDP-43 protein (Fig 15C). Purification of the GFP-S48E-TDP-43 phosphomimetic mutant resulted in a purified dimer protein. We observe that the phosphomimetic mutant was able to delay formation of aggregated TDP-43 protein.

Denaturing purification of the GFP-S48E-TDP43 mutant was completed in an attempt to attain a monomeric protein species. The mutant construct was expressed and purified in the same fashion as the wild type. SDS-PAGE confirmed the presence of the mutant construct in the elution with a single band at ~75kDa. SEC elution profile and DLS both confirm the presence of an aggregated species (Fig. 16A and B). The SEC elution profile shows a shifted second peak to a longer retention time indicating a smaller but still oligomeric species. DLS analysis reveals a mutlimeric species with a species populating 33% of the sample by mass and another populating 53% of the species by mass. These species displayed radii of 17.2nm and 118.0nm respectively estimating masses of 2665 kDa and 236841 kDa. These results are consistent with the wild type denaturing purification suggesting that urea denaturation yields an oligomeric species.

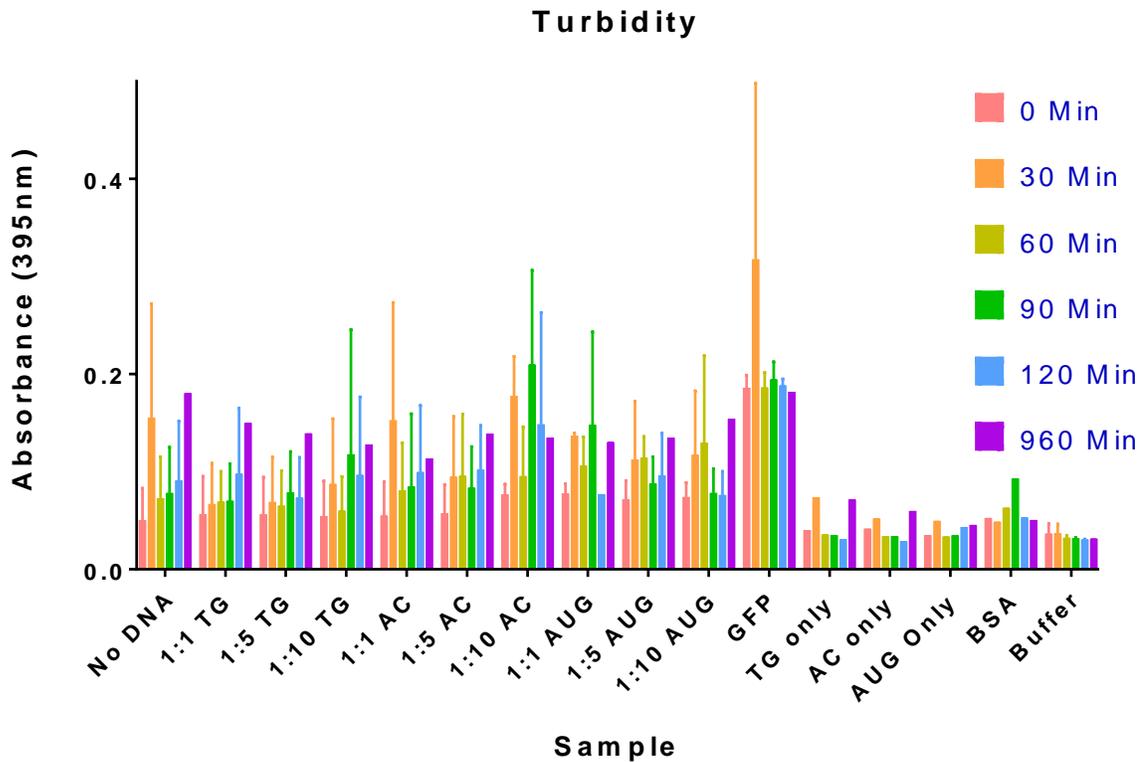


**Figure 16. Characterization of GFP-S48E-TDP-43 in denaturing conditions.** A) Size exclusion chromatogram of S48E Urea (orange) and size standards (blue). The short retention time of “eluent1” in S48E is consistent with a large aggregate. Retention time of “eluent 2” is smaller than the oligomer in “eluent 1” but still a larger oligomer than desired. B) Size determination by DLS using ~7 $\mu$ M of “Eluent 2” protein. Rh is calculated by mass with DYNAMICS software (Wyatt). C) SDS-PAGE of purified GFP-S48E Urea. Lanes contain eluted fractions from Nickel affinity purification (Nickel) and SEC purification (Eluent 1-2) compared against Precision plus Protein Ladder (Bio-Rad). Samples were run on a 4-20% polyacrylamide gel for analysis.

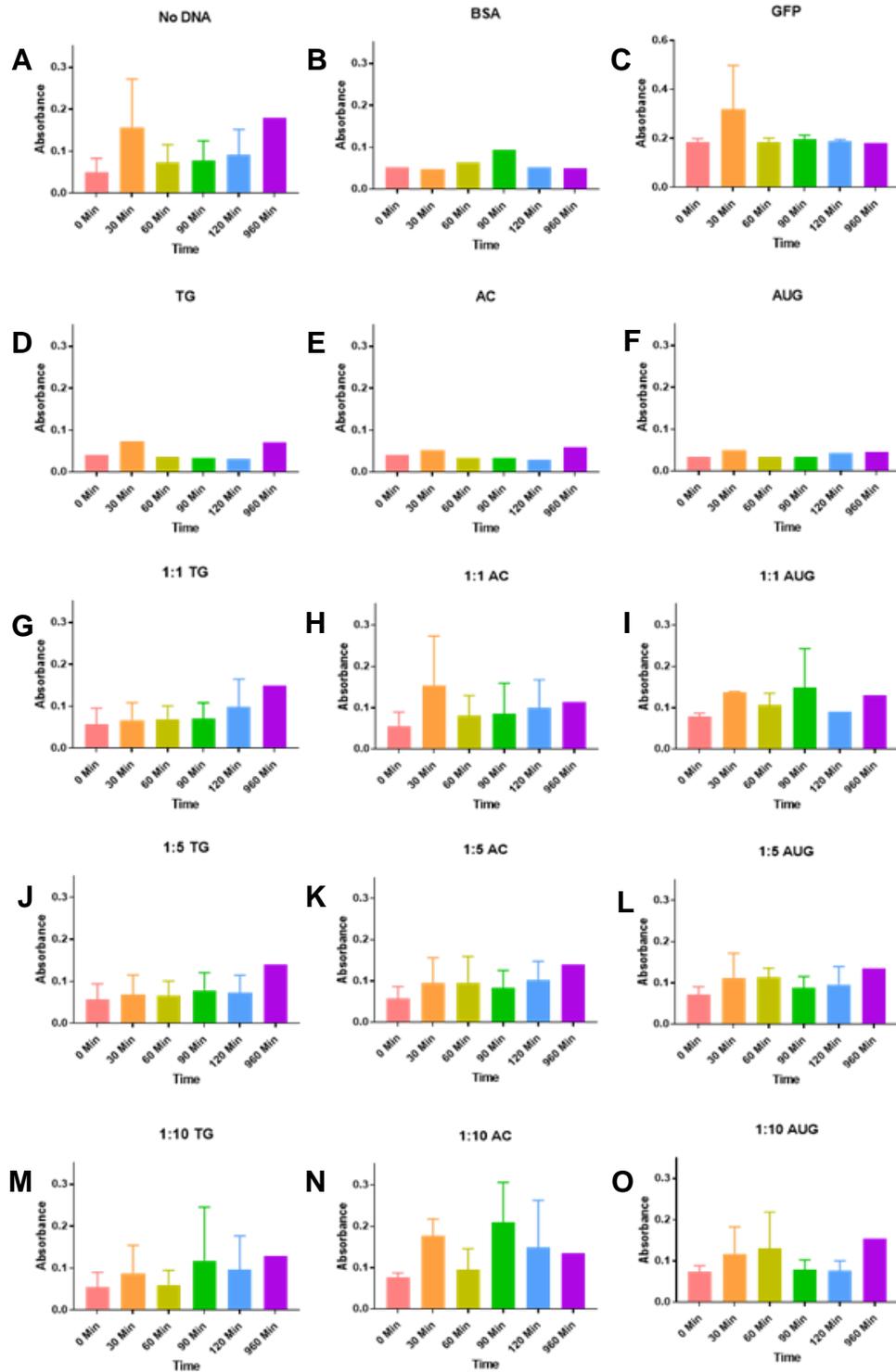
### *3.4 Nucleic Acids show minimal rescue effect when agitated*

Previous studies have shown that TDP-43 readily aggregates when agitated (Johnson et al., 2009). It has also been determined that decreasing the salt concentration of the buffer will lead TDP-43 to aggregate. The addition of TG<sub>12</sub> DNA to TDP-43 has been shown to suppress aggregation of the protein at room temperature (Sun et al., 2014).

To test whether aggregation of TDP-43 could be prevented during agitation, several different nucleic acids were combined with the dimeric TDP-43 and vigorously shaken at 1400RPM in a thermomixer. All samples containing S48E-TDP-43 aggregated at 25°C with agitation as can be seen by the increase in turbidity starting at 30 min and spanning a time frame of 16 hours (Fig 17). This data is congruent with previous reports that TDP-43 is inherently aggregation prone. TDP-43, when supplemented with TG<sub>12</sub> (Fig. 18 G,J,M) and AUG<sub>12</sub> (Fig 18 I,L,O) DNA and RNA, shows a slight lag in increasing turbidity compared with no DNA (Fig. 18A) and AC<sub>12</sub> (Fig 18 H,K,N) controls (See Fig 18 A vs G,J,M, and O). This lag could account for suppression of aggregation by nucleic acid binding. Several controls including GFP, BSA, and nucleic acid only solutions did not aggregate under identical conditions (Fig 17 and 18).



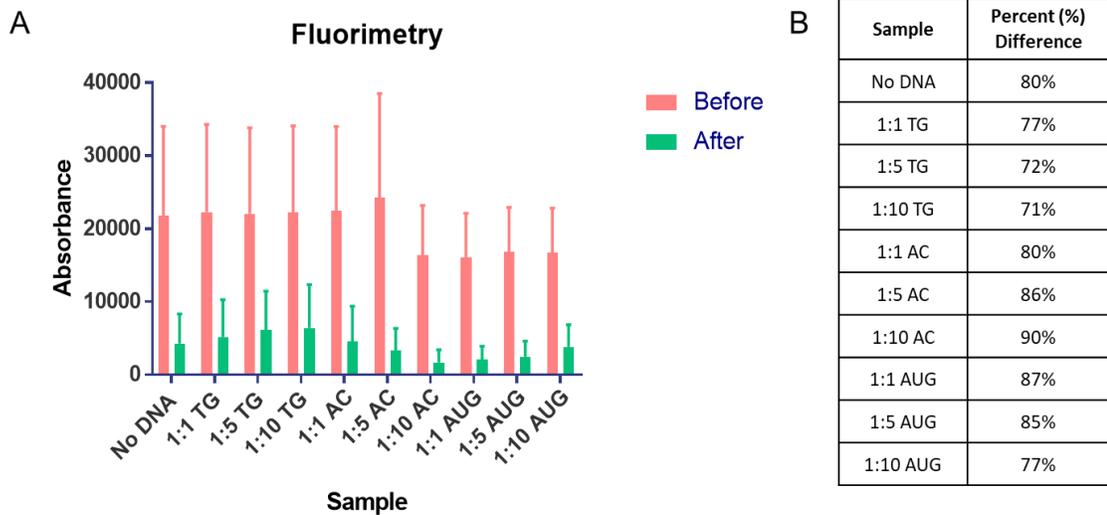
**Figure 17. Aggregation of GFP-S48E-TDP-43 *in vitro*.** GFP-S48E-TDP-43 and GFP-S48E-TDP-43+ RNA/DNA variants were incubated at 25°C with agitation for 0-960min in a buffer of ~170mM NaCl, 36mM HEPES, and 3mM  $\beta$ Me. Aggregation was determined by turbidity at 395nm. Nucleic acid in 1, 5, and 10 x stoichiometric excess to TDP-43 protein. Values with error bars are means  $\pm$ SD (n=2-3). RNA/DNA sequences in Table S5 in the appendix



**Figure 18. Individual aggregation data of TDP-43 aggregation *in vitro*.** Separated data from Figure 17 for better representation of turbidity. Absorbance is read at 396nm. Colors are consistent with Figure 17. Values are represented as mean  $\pm$ SD (n=2-3)

Fluorescence assays were used to assess the amount of TDP-43 protein in the pellet at TDP-43 “end-point” of aggregation. Similarly, we hypothesized that the change in GFP fluorescence in the supernatant would correspond to the amount of non-aggregated TDP-43 still left in the supernatant. Visual comparisons of no-DNA added TDP-43 and 1:10 TDP-43 to TG<sub>12</sub> show substantially increased fluorescence of supernatant with added nucleic acid (Fig. S3). Using EGFP’s previously reported emission/absorbance spectrum (Kennis et al., 2004), the data obtained was in parallel to data received by turbidity measurements. Addition of TG<sub>12</sub> and AUG<sub>12</sub> DNA and RNA showed a change in fluorescence which is associated with aggregation (Fig 19).

Addition of RNA/DNA does appear to substantially influence results however, statistical significance is lacking in results obtained to date. The results show no



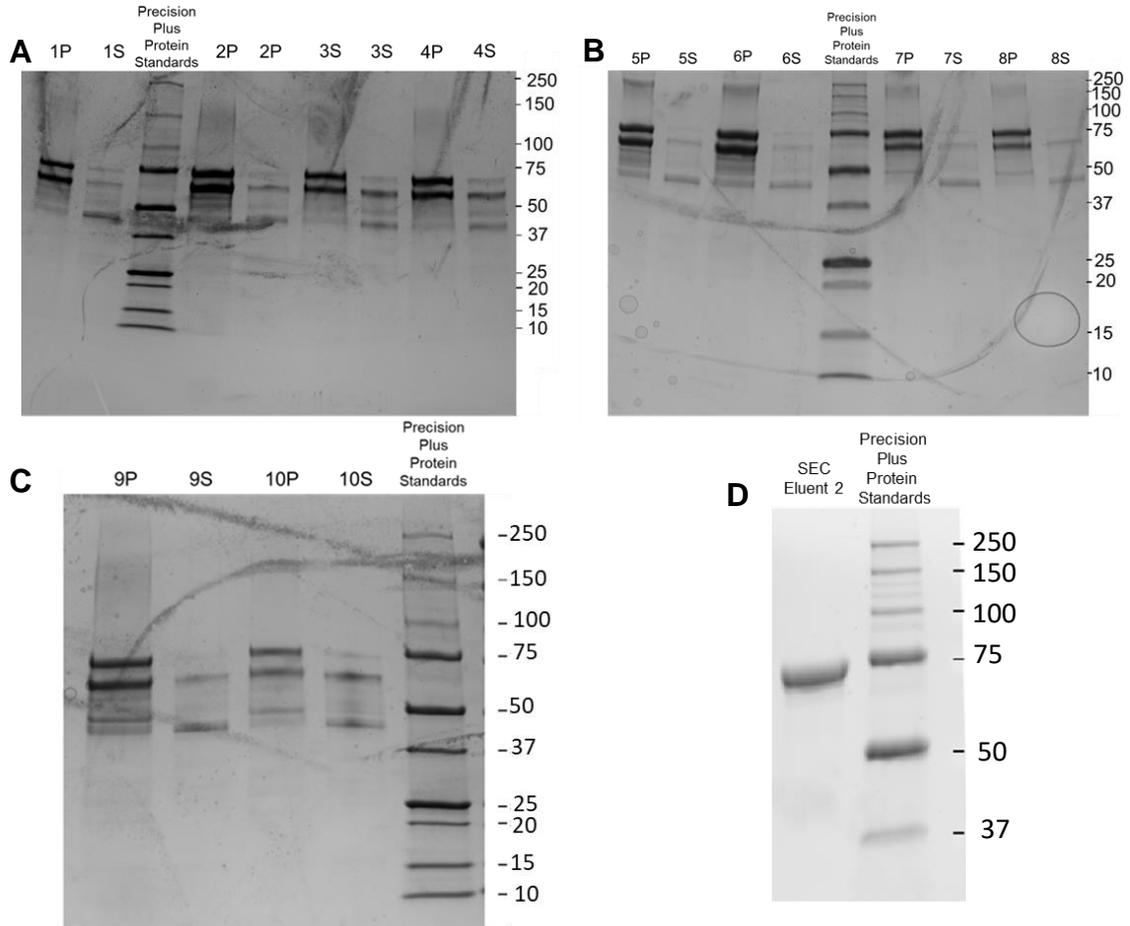
**Figure 19. Aggregation of S48E assessed by sedimentation analysis.**

A) Fluorescence at 511nm take before and after aggregation of GFP-S48E-TDP43 and GFP-S48E-TDP-43 + RNA/DNA variants as in Figure 17. B) Percent difference of before and after sample means  $\pm$ SD (n=2-3)

conclusive evidence for correlation of attenuation of aggregation during agitation with the addition of a nucleic acid substrate.

### *3.5 Cleaved protein visible in aggregates*

TDP-43 C-terminal truncations of 25kDa and 35kDa have been found in inclusions of patients with ALS and FTD (Arai et al., 2010). After 16 hours of agitation, analysis of GFP-S48E-TDP43 and GFP-S48E-TDP-43 + RNA/DNA variant samples were sedimented and supernatant removed from pellet. SDS-PAGE of supernatant and pellet species shows two species present, one at 75kDa and a smaller species at ~60kDa (Fig 20). All pellet species show bands at ~75kDa confirming the presence of TDP-43 in the pellets, probably aggregated. Lanes with the supernatant for the 1:1, 1:5, and 1:10 TG<sub>12</sub> DNA added species show faint bands increasing in concentration from 1 to 10 suggesting that TDP-43 is present in the supernatant as well. Lanes with AC<sub>12</sub> added appear to be similar to that of the No DNA added with distinct bands at ~75kDa in pellet analysis and minimal to no band in the supernatants species. Lanes with AUG<sub>12</sub> added are similar to the AC<sub>12</sub> and No DNA with the exception of the 1:10 AUG<sub>12</sub> which shows a band at ~75kDa in the supernatant fraction. These results conclude that addition of nucleic acid has a minimal effect on attenuation of aggregation when agitated for long periods of time (16 hours).



**E**

Name	Sample	Name	Sample
1P	No DNA Pellet	1S	No DNA Supernatant
2P	1:1 TG Pellet	2S	1:1 TG Supernatant
3P	1:5 TG Pellet	3S	1:5 TG Supernatant
4P	1:10 TG Pellet	4S	1:10 TG Supernatant
5P	1:1 AC Pellet	5S	1:1 AC Supernatant
6P	1:5 AC Pellet	6S	1:5 AC Supernatant
7P	1:10 AC Pellet	7S	1:10 AC Supernatant
8P	1:1 AUG Pellet	8S	1:1 AUG Supernatant
9P	1:5 AUG Pellet	9S	1:5 AUG Supernatant
10P	1:10 AUG Pellet	10S	1:10 AUG Supernatant

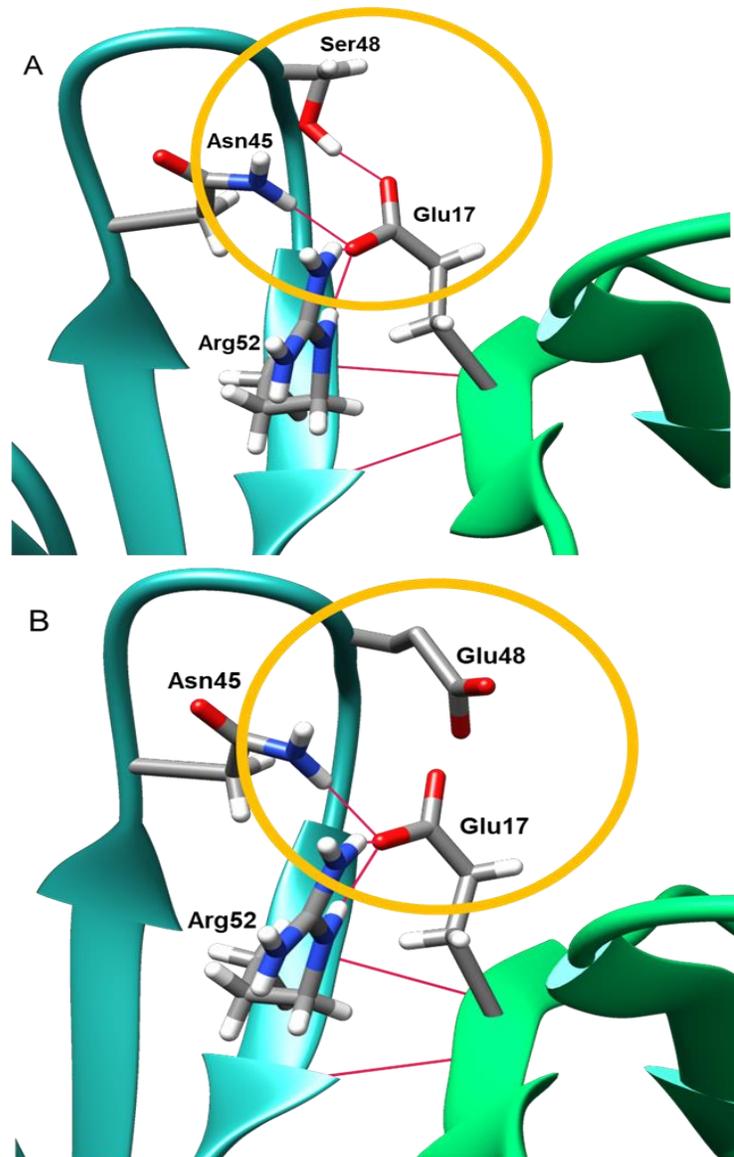
**Figure 20. Supernatant/ Pellet analysis of S48E aggregates after in vitro aggregation studies.** After aggregation analysis of GFP-S48E-TDP43 and GFP-S48E-TDP-43 + RNA/DNA variants samples were sedimented and supernatant removed from pellet. 10  $\mu$ L of each sample was run on an SDS 4-20% polyacrylamide gel for analysis. Protein compared against Precision plus Protein Ladder (Bio-Rad)A) Samples 1-4. B) Samples 5-8. C) Samples 9-10. D) SDS-PAGE of purified GFP-S48E before agitation E) Table of sample names

## CHAPTER 4: DISCUSSION AND SUMMARY

TDP-43 is a 414 residue protein that is known to bind DNA and RNA, has been shown to repress HIV-1 transcription, and to regulate the splicing of the CFTR gene (Morimoto & Cuervo, 2009). TDP-43 has been found ubiquitinated, phosphorylated, and cleaved in inclusion bodies in brain tissue of patients with ALS and FTD (Neumann et al., 2006). Many researchers have undertaken investigating the structure of this protein in an effort to gain a understanding about how TDP-43 aggregates, however, the mechanisms by which TDP-43 aggregates in these disease pathologies is still unknown (Morgan & Orrell, 2016).

Native TDP-43 exists as a dimer with a structured N-terminal domain, two RNA Recognition motifs, and a flexible C-terminal domain. Each of these domains play important roles in TDP-43's structure and function and while the entire structure of TDP-43 has yet to be resolved, much is known about the individual domains. The NTD of TDP-43 is known to facilitate dimer association in a head-tail-fashion through hydrogen bonds and electrostatic interactions (Afroz et al., 2017). A crystal structure of this domain has recently been published showing the NTD adopting both a ubiquitin- and axin-1-like fold (Afroz et al., 2017). TDP-43 contains two RRM which have been resolved by crystallography and show the RRMs to bind UG/TG repeats with high specificity (Chiang et al., 2016; Furukawa et al., 2016; Kuo et al., 2009). The highly flexible CTD of TDP-43

is intrinsically aggregation prone and the majority of ALS disease associated mutations are in this region. Due to its highly disordered nature not as much is known about the structure of the CTD compared with the other domains. An NMR structure was resolved



**Figure 21. S48E Mutation.** Mutation of serine 48 to glutamate depicted in a cartoon as a stick model with hydrogen bonding (red). When serine is mutated the hydrogen bond in the intermolecular interaction is lost. A) Normal intermolecular interaction with Ser48 and Glu17. B) No hydrogen bond interaction when mutated to glutamate.

showing an alpha-helical segment of the CTD which is suspected to drive aggregation of the full length protein.

Within each of these domains are notable mutations that affect TDP-43s structure and function. A phosphomimetic mutation in the NTD, where the serine at residue 48 is mutated to a glutamate, has been shown to disrupt intramolecular association of the dimer configuration (Fig. 21) (A. Wang et al., 2018). In the RRM1s there are a series of phenylalanines that play an important role in RNA binding. When these phenylalanines are mutated to leucines RNA binding is abolished, disrupting the native function of this protein (Gopal et al., 2017). The C-terminal domain is home to a number of mutations that are involved in disease pathogenesis. One of these mutations, Q331K, is located in the  $\alpha$ -helical segment of the domain. This mutation disrupts the  $\alpha$ -helix and drives aggregation of the protein.

TDP-43 is known as an RNA binding protein with preference for TG/UG rich sequences. Binding models suggest that when in proximity with nucleic acid the RRM1s will bind ssDNA/RNA with a rapid on/off rate (Bhardwaj et al., 2013). TDP-43 remains soluble in the presence of RNA as can be demonstrated by studies where depletion of nuclear RNA by RNase A generated aggregates (Huang et al., 2013). Previous work used these findings that RNA plays an important role in TDP-43 solubility and demonstrated the ability to attenuate aggregation with the addition of TG rich sequences (Sun et al., 2014). Our work tested the applicability of these findings to the S48E phosphomimetic mutant by looking at the ability of nucleic acid to mitigate aggregation

in agitation experiments. Full length TDP-43 is inherently aggregation prone and has been shown to aggregate upon agitation (Johnson et al., 2009).

Due to its high aggregation tendencies, full length TDP-43 is difficult to purify in its monomeric form (A. Wang et al., 2018). Previous studies have reported the ability to purify their specific constructs using nickel affinity (Sun et al., 2014), GST columns (Kitamura et al., 2018), and size exclusion chromatography (Johnson et al., 2009). We originally attempted to purify our protein construct using methods described by Sun et al. however, were not able to obtain a purified construct but rather a complex of proteins. We surmise that this could be a factor of using a different vector absent of a fusion construct, specifically one labeling a maltose binding protein (MBP) tag, which has been described in other studies to help solubilize the aggregation prone C terminal tail (A. Wang et al., 2018). Different purification strategies and adaptations from other studies that were able to purify monomeric TDP-43 were used. Differences between this study's successful purification approach and the original purification are described in supplementary table 4 (Table S4).

Due to the failure to produce a pure monomeric WT protein using conventional purification methods, a denaturing purification method using 6M urea was attempted. Lysis and subsequent column purification using a background of 6M urea was not able to be performed as no protein bound to the Ni/NTA column but instead eluted in the flow through (Fig S1). It is suspected that this occurred as a result of the urea unfolding the TDP-43 in such a way that the protein refolded and oligomerized creating an aggregate

with the buried affinity tag. This tag binds the protein to the column, gets buried within the hydrophobic core of the protein making the ligands inaccessible by the Ni/NTA resin.

A second purification method where cells were lysed in the absence of urea then bound to the column and were washed off the column with a supplemented wash buffer containing 6M urea was also pursued. Elution profiles and SDS-PAGE gels identified the majority of protein being washed off in the presence of the 6M urea wash buffer with lower imidazole concentration (Fig. S2). This protein fraction was further analyzed via SEC and DLS and presented as an oligomer with a radius  $\geq 10$ nm. Interestingly, other species can be observed on the elution profile of the denaturing purification during imidazole elution (Fig S.2, “elution 2”). Elution 2 was not used for analysis due to the majority of protein eluting in the first profile. However, it cannot be ruled out that the protein in elution 2 is the monomeric or dimeric TDP-43 species we were hoping to purify, as the protein is still present in this second species as observed by SDS-PAGE (Fig. 14).

Purification of a TDP-43 variant was accomplished using the phosphomimetic mutation described to disrupt dimer formation of the NTD of TDP-43 (A. Wang et al., 2018). A crystal structure of the TDP-43 NTD dimer shows an intermolecular interaction of Ser48 on one strand of the dimer and hydrogen bonding with Glu17 on the other (Fig 21) (Afroz et al., 2017). SEC purification and analysis of the GFP-S48E-TDP-43 mutation resulted in two elution peaks. The first elution peak in the void volume of the column suggests the presence of aggregates and a second peak eluted at a retention time consistent with a protein whose molecular weight is similar to that of the TDP-43 dimer

(Fig. 15). DLS measurements confirmed the radius of the species eluted to be consistent with previous reports of the dimer formation at 5.3nm (Sun et al., 2014).

Addition of nucleic acid to the purified GFP-S48E-TDP-43 dimer in agitation experiments began to show a trend of slight delay in aggregation as determined by turbidity measurements. The results however are not conclusive because they are not statistically significant within  $1\sigma$ . TDP-43 has previously been reported to have a specificity for UG/TG rich sequences but is also known to bind other nucleic acid sequences. There have been several reports that show TDP-43s inherent propensity to aggregate (Johnson et al., 2009). There have also been reports that prove nucleic acid's ability to hinder aggregation of TDP-43 protein (Sun et al., 2014). However, to the best of our knowledge, there have been no reports on the addition of nucleic acid to TDP-43 during agitation, mimicking the effects that may cause MND in CTE. Our results illustrate that addition of TG<sub>12</sub> and AUG<sub>12</sub> begin to show an effect of minimal mitigation of aggregation during agitation. While the effects are observable they are not as pronounced as was hypothesized and the experiments can be improved upon more to reach statistical significance for conclusive data. One possible explanation for the minimal mitigation effect is the presence of pre-formed, possibly terminally aggregated, aggregates in solution. TDP-43 has been reported to self-associate into aggregates in a time frame of 30 minutes at room temperature (Sun et al., 2014). An additional speculation for why the inhibitory effect is not so pronounced is that the vigorous shaking of solution isn't allowing for nucleic acid to bind to the RRM. As was mentioned before, DNA binds and dissociates from the TDP-43 in a rather fast manner. The rapid agitation

of the protein/nucleic acid solution may account for faster dissociation or otherwise provide an environment that is not suitable for binding. The phosphomimetic mutation, S48E, could also contribute to the mitigated attenuation effects. There is no conclusive evidence that the S48E-TDP-43 is binding nucleic acid in our studies. Future work could utilize a gel shift assay in order to visualize binding activity with the mutation.

Fluorescence assays were used to observe the amount of protein remaining in the supernatant vs. the aggregate with the assumption that aggregated protein would localize in the pellet. These sedimentation analysis results observed the fluorescence of GFP in the supernatant with the hypothesis that large aggregates sedimented into the pellet fraction which non-aggregated TDP-43 remained in the supernatant. These fluorescence assays of the supernatant are congruent with turbidity results which tell us that the addition of nucleic acid with protein in agitation experiments has minimal attenuation effects. Rescue from agitation induced aggregation with the nucleic acid was even less pronounced in this fluorescence assay of supernatant vs. the pellet which could be an effect of the measurements being taken at the “end-point” (~16 hours) of agitation. Turbidity results indicated that even with the addition of nucleic acid, aggregates were still being formed. The data received from fluorescence assays report mostly complete aggregate formation of agitated samples. The presence of two species as resolved by SDS-PAGE of supernatant vs. pellet of the agitated samples agrees with the finding that TDP-43 is cleaved in aggregates.

Further work streamlining the methods by which the TDP-43 protein is subjected to agitation studies will be important in determining the effect by which nucleic acid may

attenuate aggregation during agitation. In addition, modification of the vector by adding a MBP tag as is described in other studies may assist in the purification of full length proteins that may be used in agitation experiments as well.

Overall, we were not able to produce a purified monomeric species of WT protein using the methods described by previous studies, and our data suggest that urea denaturation does not produce a purified monomeric protein but rather a large, soluble oligomeric species. It can also be concluded that the addition of a phosphomimic mutant to the NTD of TDP-43 delays aggregation, perhaps disrupting self-association (A. Wang et al., 2018). Addition of nucleic acid to TDP-43 showed substantial decrease in aggregation propensity of species in heat stress environments in previous studies; however, the ability to mitigate aggregation while agitated was less pronounced in this study.

## REFERENCES

- Afroz, T., Hock, E. M., Ernst, P., Foglieni, C., Jambeau, M., Gilhespy, L. A. B., ... Polymenidou, M. (2017a). Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. *Nature Communications*, 8(1), 1–14. <https://doi.org/10.1038/s41467-017-00062-0>
- Afroz, T., Hock, E. M., Ernst, P., Foglieni, C., Jambeau, M., Gilhespy, L. A. B., ... Polymenidou, M. (2017b). Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. *Nature Communications*, 8(1), 1–14. <https://doi.org/10.1038/s41467-017-00062-0>
- Aguzzi, A., & Calella, A. M. (2009). Prions: Protein Aggregation and Infectious Diseases. *Physiological Reviews*, 89(4), 1105–1152. <https://doi.org/10.1152/physrev.00006.2009>
- Arai, T., Hasegawa, M., Nonoka, T., Kametani, F., Yamashita, M., Hosokawa, M., ... Akiyama, H. (2010). Phosphorylated and cleaved TDP-43 in ALS, FTLN and other neurodegenerative disorders and in cellular models of TDP-43 proteinopathy. *Neuropathology*, 30(2), 170–181. <https://doi.org/10.1111/j.1440-1789.2009.01089.x>
- Arnold, E. S., Ling, S.-C., Huelga, S. C., Lagier-Tourenne, C., Polymenidou, M., Ditsworth, D., ... Cleveland, D. W. (2013). ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43. *Proceedings of the National Academy of Sciences*, 110(8), E736–E745. <https://doi.org/10.1073/pnas.1222809110>
- Ayala, Y. M., Pantano, S., D'Ambrogio, A., Buratti, E., Brindisi, A., Marchetti, C., ... Baralle, F. E. (2005). Human, Drosophila, and C. elegans TDP43: Nucleic acid binding properties and splicing regulatory function. *Journal of Molecular Biology*, 348(3), 575–588. <https://doi.org/10.1016/j.jmb.2005.02.038>
- Bhardwaj, A., Myers, M. P., Buratti, E., & Baralle, F. E. (2013). Characterizing TDP-43 interaction with its RNA targets. *Nucleic Acids Research*, 41(9), 5062–5074. <https://doi.org/10.1093/nar/gkt189>
- Bonafede, R., & Mariotti, R. (2017). ALS Pathogenesis and Therapeutic Approaches: The Role of Mesenchymal Stem Cells and Extracellular Vesicles. *Frontiers in Cellular Neuroscience*, 11(March), 1–16. <https://doi.org/10.3389/fncel.2017.00080>
- Borroni, B., Stanic, J., Verpelli, C., Mellone, M., Bonomi, E., Alberici, A., ... Gardoni, F. (2017). Anti-AMPA GluA3 antibodies in Frontotemporal dementia: A new molecular target. *Scientific Reports*, 7(1), 1–10. <https://doi.org/10.1038/s41598-017-06117-y>
- Brange, J., Havelund, S., & Hougaard, P. (1992). Chemical Stability of Insulin. 2. Formation of Higher Molecular Weight Transformation Products During Storage of Pharmaceutical Preparations. *Pharmaceutical Research: An Official Journal of the American Association of Pharmaceutical Scientists*. <https://doi.org/10.1023/A:1015887001987>
- Brown, R. H., & Al-Chalabi, A. (2017). Amyotrophic Lateral Sclerosis. *New England Journal of Medicine*, 377(2), 162–172. <https://doi.org/10.1056/NEJMra1603471>
- Budini, M., Romano, V., Quadri, Z., Buratti, E., & Baralle, F. E. (2015). TDP-43 loss of

- cellular function through aggregation requires additional structural determinants beyond its C-terminal Q/N prion-like domain. *Human Molecular Genetics*, 24(1), 9–20. <https://doi.org/10.1093/hmg/ddu415>
- Buratti, E., & Baralle, F. E. (2001). Characterization and Functional Implications of the RNA Binding Properties of Nuclear Factor TDP-43, a Novel Splicing Regulator of CFTR Exon 9. *Journal of Biological Chemistry*, 276(39), 36337–36343. <https://doi.org/10.1074/jbc.M104236200>
- Buratti, E., Dörk, T., Zuccato, E., Pagani, F., Romano, M., Baralle, F. E., ... Green, M. (2001). Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. *The EMBO Journal*, 20(7), 1774–1784. <https://doi.org/10.1093/emboj/20.7.1774>
- Chang, C. ke, Wu, T. H., Wu, C. Y., Chiang, M. hui, Toh, E. K. W., Hsu, Y. C., ... Huang, J. J. T. (2012). 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>
- Chapman, E. G., Moon, S. L., Wilusz, J., & Kieft, J. S. (2014). RNA structures that resist degradation by Xrn1 produce a pathogenic dengue virus RNA. *ELife*, 2014(3), 1–25. <https://doi.org/10.7554/eLife.01892>
- Chaturvedi, S. K., Siddiqi, M. K., Alam, P., & Khan, R. H. (2016). Protein misfolding and aggregation: Mechanism, factors and detection. *Process Biochemistry*, 51(9), 1183–1192. <https://doi.org/10.1016/j.procbio.2016.05.015>
- Chi, E., Krishnan, S., & Kendrick, B. (2003). Roles of conformational stability and colloidal stability in the aggregation of recombinant human granulocyte colony-stimulating factor. *Protein Science*, 12, 903–913. <https://doi.org/10.1110/ps.0235703.mer>
- Chi, E. Y., Krishnan, S., Randolph, T. W., & Carpenter, J. F. (2003). Physical stability of proteins in aqueous solution: Mechanism and driving forces in nonnative protein aggregation. *Pharmaceutical Research*, 20(9), 1325–1336. <https://doi.org/10.1023/A:1025771421906>
- Chiang, C. H., Grauffel, C., Wu, L. S., Kuo, P. H., Doudeva, L. G., Lim, C., ... Yuan, H. S. (2016). Structural analysis of disease-related TDP-43 D169G mutation: Linking enhanced stability and caspase cleavage efficiency to protein accumulation. *Scientific Reports*, 6(September 2015), 1–14. <https://doi.org/10.1038/srep21581>
- Chiti, F., & Dobson, C. M. (2006). Protein Misfolding, Functional Amyloid, and Human Disease. *Annual Review of Biochemistry*, 75(1), 333–366. <https://doi.org/10.1146/annurev.biochem.75.101304.123901>
- Choi, S. Il, Han, K. S., Kim, C. W., Ryu, K. S., Kim, B. H., Kim, K. H., ... Seong, B. L. (2008). Protein solubility and folding enhancement by interaction with RNA. *PLoS ONE*, 3(7), 1–10. <https://doi.org/10.1371/journal.pone.0002677>
- Cléry, A., Blatter, M., & Allain, F. H. T. (2008). RNA recognition motifs: boring? Not quite. *Current Opinion in Structural Biology*, 18(3), 290–298. <https://doi.org/10.1016/j.sbi.2008.04.002>
- Cleveland, D. W., & Rothstein, J. D. (2001). From Charcot to Lou Gehrig. *Nature Reviews Neuroscience*, 2(November), 806–819.

- Cohen, T. J., Hwang, A. W., Restrepo, C. R., Yuan, C., John, Q., & Lee, V. M. Y. (2015). Propensity. <https://doi.org/10.1038/ncomms6845>. An
- Conicella, A. E., Mittal, J., Fawzi, N. L., Biology, C., States, U., States, U., & States, U. (2017). HHS Public Access, *24*(9), 1537–1549. <https://doi.org/10.1016/j.str.2016.07.007>.ALS
- D'Ambrogio, A., Buratti, E., Stuani, C., Guarnaccia, C., Romano, M., Ayala, Y. M., & Baralle, F. E. (2009). Functional mapping of the interaction between TDP-43 and hnRNP A2 in vivo. *Nucleic Acids Research*, *37*(12), 4116–4126. <https://doi.org/10.1093/nar/gkp342>
- Day, R., Bennion, B. J., Ham, S., & Daggett, V. (2002). Increasing temperature accelerates protein unfolding without changing the pathway of unfolding. *Journal of Molecular Biology*, *322*(1), 189–203. [https://doi.org/10.1016/S0022-2836\(02\)00672-1](https://doi.org/10.1016/S0022-2836(02)00672-1)
- De Baets, G., Van Doorn, L., Rousseau, F., & Schymkowitz, J. (2015). Increased Aggregation Is More Frequently Associated to Human Disease-Associated Mutations Than to Neutral Polymorphisms. *PLoS Computational Biology*, *11*(9), 1–14. <https://doi.org/10.1371/journal.pcbi.1004374>
- Du, K., Karp, P. H., Ackerley, C., Zabner, J., Keshavjee, S., Cutz, E., & Yeger, H. (2015). Aggregates of mutant CFTR fragments in airway epithelial cells of CF lungs: New pathologic observations. *Journal of Cystic Fibrosis*, *14*(2), 182–193. <https://doi.org/10.1016/j.jcf.2014.09.012>
- Fink, A. L. (1998). Protein aggregation: Folding aggregates, inclusion bodies and amyloid. *Folding and Design*, *3*(1), 9–23. [https://doi.org/10.1016/S1359-0278\(98\)00002-9](https://doi.org/10.1016/S1359-0278(98)00002-9)
- Furukawa, Y., Suzuki, Y., Fukuoka, M., Nagasawa, K., Nakagome, K., Shimizu, H., ... Akiyama, S. (2016). A molecular mechanism realizing sequence-specific recognition of nucleic acids by TDP-43. *Scientific Reports*, *6*(September 2015), 1–12. <https://doi.org/10.1038/srep20576>
- Gambotto, A., Dworacki, G., Cicinnati, V., Kenniston, T., Steitz, J., Tü Ting, T., ... Deleo, A. (2000). Immunogenicity of enhanced green fluorescent protein (EGFP) in BALB/c mice: identification of an H2-K d -restricted CTL epitope. *Gene Therapy*, *7*, 2036–2040. Retrieved from <https://search.proquest.com/openview/eed3bb534d1f8c190ddb47048ef91adb/1?pq-origsite=gscholar&cbl=34384>
- Gopal, P. P., Nirschl, J. J., Klinman, E., & Holzbaur, E. L. F. (2017). Amyotrophic lateral sclerosis-linked mutations increase the viscosity of liquid-like TDP-43 RNP granules in neurons. *Proceedings of the National Academy of Sciences*, *114*(12), E2466–E2475. <https://doi.org/10.1073/pnas.1614462114>
- Griesbeck, O., Baird, G. S., Campbell, R. E., Zacharias, D. A., & Tsien, R. Y. (2001). Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *Journal of Biological Chemistry*, *276*(31), 29188–29194. <https://doi.org/10.1074/jbc.M102815200>
- Guenther, E. L., Ge, P., Trinh, H., Sawaya, M. R., Cascio, D., Boyer, D. R., ... Eisenberg, D. S. (2018). Atomic-level evidence for packing and positional amyloid

- polymorphism by segment from TDP-43 RRM2. *Nature Structural and Molecular Biology*, 25(4), 311–319. <https://doi.org/10.1038/s41594-018-0045-5>
- Haass, C., & Neumann, M. (2016). Frontotemporal dementia: from molecular mechanisms to therapy. *Journal of Neurochemistry*, 138, 3–5. <https://doi.org/10.1111/jnc.13619>
- Hachiya, N. S., Kozuka, Y., & Kaneko, K. (2008). Mechanical stress and formation of protein aggregates in neurodegenerative disorders. *Medical Hypotheses*, 70(5), 1034–1037. <https://doi.org/10.1016/j.mehy.2007.06.043>
- Hillgren, A., Lindgren, J., & Aldén, M. (2002). Protection mechanism of Tween 80 during freeze-thawing of a model protein, LDH. *International Journal of Pharmaceutics*, 237(1–2), 57–69. [https://doi.org/10.1016/S0378-5173\(02\)00021-2](https://doi.org/10.1016/S0378-5173(02)00021-2)
- Höhfeld, J. (2017). Preserving protein function through reversible aggregation. *Nature Cell Biology*, 19, 1142. Retrieved from <http://dx.doi.org/10.1038/ncb3620>
- Huang, Y. C., Lin, K. F., He, R. Y., Tu, P. H., Koubek, J., Hsu, Y. C., & Huang, J. J. T. (2013). Inhibition of TDP-43 Aggregation by Nucleic Acid Binding. *PLoS ONE*, 8(5). <https://doi.org/10.1371/journal.pone.0064002>
- Jiang, L. L., Che, M. X., Zhao, J., Zhou, C. J., Xie, M. Y., Li, H. Y., ... Hu, H. Y. (2013). Structural transformation of the amyloidogenic core region of TDP-43 protein initiates its aggregation and cytoplasmic inclusion. *Journal of Biological Chemistry*, 288(27), 19614–19624. <https://doi.org/10.1074/jbc.M113.463828>
- Johnson, B. S., Snead, D., Lee, J. J., McCaffery, J. M., Shorter, J., & Gitler, A. D. (2009). TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. *Journal of Biological Chemistry*, 284(30), 20329–20339. <https://doi.org/10.1074/jbc.M109.010264>
- Kameoka, D., Masuzaki, E., Ueda, T., & Imoto, T. (2007). Effect of buffer species on the unfolding and the aggregation of humanized IgG. *Journal of Biochemistry*, 142(3), 383–391. <https://doi.org/10.1093/jb/mvm145>
- Katayama, D. S., Nayar, R., Chou, D. K., Valente, J. J., Cooper, J., Henry, C. S., ... Manning, M. C. (2006). Effect of buffer species on the thermally induced aggregation of interferon-tau. *Journal of Pharmaceutical Sciences*, 95(6), 1212–1226. <https://doi.org/10.1002/jps.20471>
- Kennis, J. T. M., Larsen, D. S., van Stokkum, I. H. M., Vengris, M., van Thor, J. J., & van Grondelle, R. (2004). Uncovering the hidden ground state of green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America*, 101(52), 17988–17993. <https://doi.org/10.1073/pnas.0404262102>
- Kiernan, M. C., Vucic, S., Cheah, B. C., Turner, M. R., Eisen, A., Hardiman, O., ... Zoing, M. C. (2011). Amyotrophic lateral sclerosis. *The Lancet*, 377(9769), 942–955. [https://doi.org/10.1016/S0140-6736\(10\)61156-7](https://doi.org/10.1016/S0140-6736(10)61156-7)
- Kiese, S., Pappengerger, A., Friess, W., & Mahler, H. C. (2008). Shaken, not stirred: Mechanical stress testing of an IgG1 antibody. *Journal of Pharmaceutical Sciences*, 97(10), 4347–4366. <https://doi.org/10.1002/jps.21328>
- Kitamura, A., Shibasaki, A., Takeda, K., Suno, R., & Kinjo, M. (2018). Analysis of the substrate recognition state of TDP-43 to single-stranded DNA using fluorescence correlation spectroscopy. *Biochemistry and Biophysics Reports*, 14(March), 58–63.

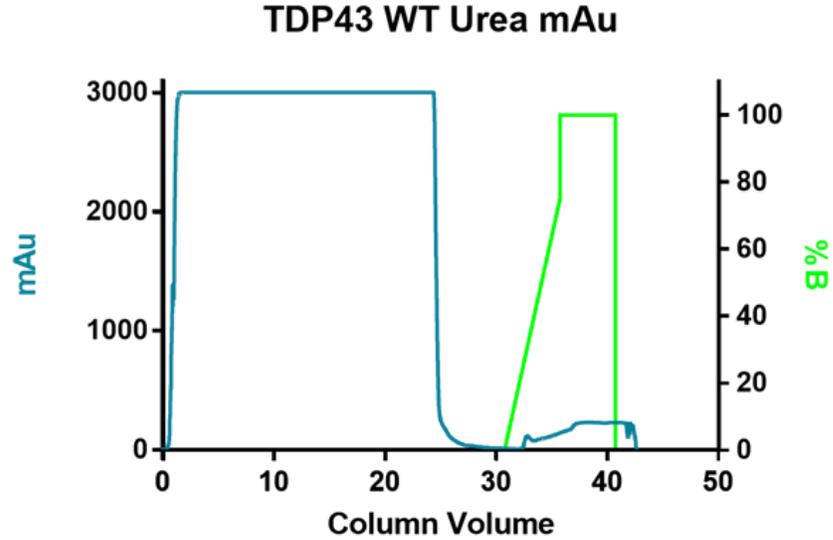
- <https://doi.org/10.1016/j.bbrep.2018.03.009>
- Krielgaard, L., Jones, L. S., Randolph, T. W., Frokjaer, S., Flink, J. M., Manning, M. C., & Carpenter, J. F. (1998). Effect of tween 20 on freeze-thawing- and agitation-induced aggregation of recombinant human factor XIII. *Journal of Pharmaceutical Sciences*, *87*(12), 1597–1603. <https://doi.org/10.1021/js980126i>
- Krishnan, S., Chi, E. Y., Webb, J. N., Chang, B. S., Shan, D., Goldenberg, M., ... Carpenter, J. F. (2002). Aggregation of granulocyte colony stimulating factor under physiological conditions: Characterization and thermodynamic inhibition. *Biochemistry*, *41*(20), 6422–6431. <https://doi.org/10.1021/bi012006m>
- Kueltzo, L. A., Wang, W., Randolph, T. W., & Carpenter, J. F. (2008). Effects of solution conditions, processing parameters, and container materials on aggregation of a monoclonal antibody during freeze-thawing. *Journal of Pharmaceutical Sciences*, *97*(5), 1801–1812. <https://doi.org/10.1002/jps.21110>
- Kumar, S., Mary Thangakani, A., Nagarajan, R., Singh, S. K., Velmurugan, D., & Michael Gromiha, M. (2016). Autoimmune responses to soluble aggregates of amyloidogenic proteins involved in neurodegenerative diseases: Overlapping aggregation prone and autoimmunogenic regions. *Scientific Reports*, *6*(October 2015), 1–13. <https://doi.org/10.1038/srep22258>
- Kuo, P. H., Chiang, C. H., Wang, Y. T., Doudeva, L. G., & Yuan, H. S. (2014). The crystal structure of TDP-43 RRM1-DNA complex reveals the specific recognition for UG- and TG-rich nucleic acids. *Nucleic Acids Research*, *42*(7), 4712–4722. <https://doi.org/10.1093/nar/gkt1407>
- 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>
- Landsman, D. (1992). RNP-1, an RNA-binding motif is conserved in the DNA-binding cold shock domain. *Nucleic Acids Research*, *20*(11), 2861–2864. <https://doi.org/10.1093/nar/20.11.2861>
- Lange, A., Mills, R. E., Lange, C. J., Stewart, M., Devine, S. E., & Corbett, A. H. (2007). Classical nuclear localization signals: Definition, function, and interaction with importin  $\alpha$ . *Journal of Biological Chemistry*, *282*(8), 5101–5105. <https://doi.org/10.1074/jbc.R600026200>
- Lee E. Goldstein, Andrew M. Fisher, Chad A. Tagge, Xiao-Lei Zhang, L., Velisek, John A. Sullivan, Chirag Upreti, Jonathan M. Kracht, Maria Ericsson, M. W., Wojnarowicz, Cezar J. Goletiani, Giorgi M. Maglakelidze, Noel Casey, J. A., Moncaster, Olga Minaeva, Robert D. Moir, Christopher J. Nowinski<sup>8</sup>, R. A., Stern<sup>2, 8</sup>, Robert C. Cantu<sup>8, 9</sup>, James Geiling, Jan K. Blusztajn, B. L. W., Tsuneya Ikezu, Thor D. Stein, Andrew E. Budson, Neil W. Kowall, D., ... Cleveland, Rudolph E. Tanzi, Patric K. Stanton, and A. C. M. (2012). Chronic Traumatic Encephalopathy in Blast-Exposed Military Veterans and a Blast Neurotrauma Mouse Model. *Science Translational Medicine*. <https://doi.org/10.1126/scitranslmed.3003716>
- Lim, L., Wei, Y., Lu, Y., & Song, J. (2016). ALS-Causing Mutations Significantly Perturb the Self-Assembly and Interaction with Nucleic Acid of the Intrinsically Disordered Prion-Like Domain of TDP-43. *PLoS Biology*, *14*(1), 1–35.

- <https://doi.org/10.1371/journal.pbio.1002338>
- Mackenzie, I. R. A., & Neumann, M. (2016). Molecular neuropathology of frontotemporal dementia: insights into disease mechanisms from postmortem studies. *Journal of Neurochemistry*, 138, 54–70. <https://doi.org/10.1111/jnc.13588>
- Mahler, H. C., Müller, R., Frieß, W., Delille, A., & Matheus, S. (2005). Induction and analysis of aggregates in a liquid IgG1-antibody formulation. *European Journal of Pharmaceutics and Biopharmaceutics*, 59(3), 407–417. <https://doi.org/10.1016/j.ejpb.2004.12.004>
- Maris, C., Dominguez, C., & Allain, F. H. T. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS Journal*, 272(9), 2118–2131. <https://doi.org/10.1111/j.1742-4658.2005.04653.x>
- Martland, H. S. (1928). Punch drunk. *Journal of the American Medical Association*, 91(15), 1103–1107. <https://doi.org/10.1001/jama.1928.02700150029009>
- McKee, A. C., Cantu, R. C., Nowinski, C. J., Hedley-whyte, T., Gavett, B. E., Budson, A. E., ... Stern, R. A. (2009). Chronic Traumatic Encephalopathy in Athletes: Progresive Tauopathy following Repetitive Head Injury. *Journal of Neuropathology and Experimental Neurology*, 68(7), 709–735. <https://doi.org/10.1097/NEN.0b013e3181a9d503>
- McKee, A., & Gavett, B. (2010). TDP-43 proteinopathy and motor neuron disease in chronic traumatic encephalopathy. *Journal of ...*, 69(9), 918–929. <https://doi.org/10.1097/NEN.0b013e3181ee7d85>
- Miyazaki, Y., Mizumoto, K., Dey, G., Kudo, T., Perrino, J., Chen, L. C., ... Wandless, T. J. (2016). A method to rapidly create protein aggregates in living cells. *Nature Communications*, 7(May), 1–7. <https://doi.org/10.1038/ncomms11689>
- Mompeán, M., Buratti, E., Guarnaccia, C., Brito, R. M. M., Chakrabarty, A., Baralle, F. E., & Laurents, D. V. (2014). Structural characterization of the minimal segment of TDP-43 competent for aggregation. *Archives of Biochemistry and Biophysics*, 545, 53–62. <https://doi.org/10.1016/j.abb.2014.01.007>
- Mompeán, M., Hervás, R., Xu, Y., Tran, T. H., Guarnaccia, C., Buratti, E., ... Laurents, D. V. (2015). Structural Evidence of Amyloid Fibril Formation in the Putative Aggregation Domain of TDP-43. *Journal of Physical Chemistry Letters*, 6(13), 2608–2615. <https://doi.org/10.1021/acs.jpcllett.5b00918>
- Mompean, M., Romano, V., Pantoja-Uceda, D., Stuani, C., Baralle, F. E., Buratti, E., & Laurents, D. V. (2016). The TDP-43 N-terminal domain structure at high resolution. *The FEBS Journal*, 283(7), 1242–1260. <https://doi.org/10.1111/febs.13651>
- 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>
- Morimoto, R. I., & Cuervo, A. M. (2009). Protein homeostasis and aging: Taking care of

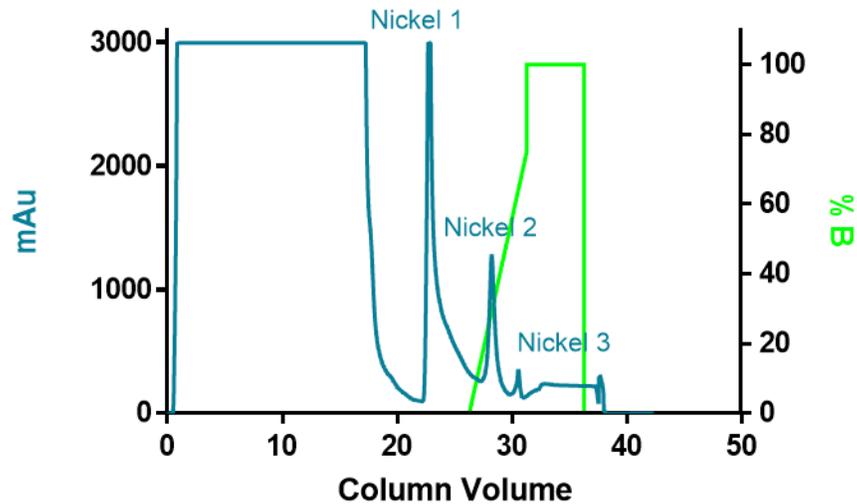
- proteins from the cradle to the grave. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, 64(2), 167–170.  
<https://doi.org/10.1093/gerona/gln071>
- Neumann, M., Sampathu, D. M., & Kwong, L. K. (2006). Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*, 314(5796), 130–133. <https://doi.org/10.1126/science.1134108>
- Ou, S. H., Wu, F., Harrich, D., García-Martínez, L. F., Gaynor, R. B., Garcia-Martinez, L., & Gaynor, R. B. (1995). Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *Journal of Virology*, 69(6), 3584–3596.
- Philo, J. S., & Arakawa, T. (2009). Mechanisms of protein aggregation. *Current Pharmaceutical Biotechnology*, 10(4), 348–351.  
<https://doi.org/10.2174/138920109788488932>
- Prasad, A., Sivalingam, V., Bharathi, V., Girdhar, A., & Patel, B. K. (2018). The amyloidogenicity of a C-terminal region of TDP-43 implicated in Amyotrophic Lateral Sclerosis can be affected by anions, acetylation and homodimerization. *Biochimie*, 150, 76–87. <https://doi.org/10.1016/j.biochi.2018.05.003>
- Puvenna, V., Engeler, M., Banjara, M., Brennan, C., Schreiber, P., Dadas, A., ... Janigro, D. (2016). Is phosphorylated tau unique to chronic traumatic encephalopathy? Phosphorylated tau in epileptic brain and chronic traumatic encephalopathy. *Brain Research*, 1630, 225–240. <https://doi.org/10.1016/j.brainres.2015.11.007>
- Ratti, A., & Buratti, E. (2016). Physiological functions and pathobiology of TDP-43 and FUS/TLS proteins. *Journal of Neurochemistry*, 138, 95–111.  
<https://doi.org/10.1111/jnc.13625>
- Ross, C. A., & Poirier, M. A. (2004). Protein aggregation and neurodegenerative disease. *Nature Medicine*, 10(7), S10. <https://doi.org/10.1038/nm1066>
- Rubinsztein, D. C. (2006). The roles of intracellular protein-degradation pathways in neurodegeneration. *Nature*, 443(7113), 780–786.  
<https://doi.org/10.1038/nature05291>
- Scheper, G. C., Van Der Knaap, M. S., & Proud, C. G. (2007). Translation matters: Protein synthesis defects in inherited disease. *Nature Reviews Genetics*, 8(9), 711–723. <https://doi.org/10.1038/nrg2142>
- Scotter, E. L., Chen, H. J., & Shaw, C. E. (2015). TDP-43 Proteinopathy and ALS: Insights into Disease Mechanisms and Therapeutic Targets. *Neurotherapeutics*, 12(2), 352–363. <https://doi.org/10.1007/s13311-015-0338-x>
- Shihong Li, Christian Schoneich, and R. T. B. (1995). Chemical Instability of Protein Phosphorylation: Mechanisms of Oxidation and Strategies for Stabilization, 48, 490–500.
- Stein, T. D., Alvarez, V. E., Mckee, A. C., & Traumatic, C. (2015). Concussion in Chronic Traumatic Encephalopathy. *Curr Pain Headache Rep*, 19(10).  
<https://doi.org/10.1007/s11916-015-0522-z>
- Strambini, G. B., & Gonnelli, M. (2007). Protein stability in ice. *Biophysical Journal*, 92(6), 2131–2138. <https://doi.org/10.1529/biophysj.106.099531>
- 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., & Chakrabarty, A. (2017). Phase to Phase with TDP-43. *Biochemistry*, 56(6), 809–823. <https://doi.org/10.1021/acs.biochem.6b01088>
- Treuheit, M. J., Kosky, A. a, & Brems, D. N. (2002). Inverse Relationship of Protein Concentration and Aggregation, 19(4).
- Van Buren, N., Rehder, D., Gadgil, H., Matsumura, M., & Jacob, J. (2009). Elucidation of two major aggregation pathways in an IgG2 antibody. *Journal of Pharmaceutical Sciences*, 98(9), 3013–3030. <https://doi.org/10.1002/jps.21514>
- Wang, A., Conicella, A. E., Schmidt, H. B., Martin, E. W., Rhoads, S. N., Reeb, A. N., ... Fawzi, N. L. (2018). A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. *The EMBO Journal*, e97452. <https://doi.org/10.15252/embj.201797452>
- Wang, W. (2005). Protein aggregation and its inhibition in biopharmaceutics. *International Journal of Pharmaceutics*, 289(1–2), 1–30. <https://doi.org/10.1016/j.ijpharm.2004.11.014>
- Wang, W., Nema, S., & Teagarden, D. (2010). Protein aggregation-Pathways and influencing factors. *International Journal of Pharmaceutics*, 390(2), 89–99. <https://doi.org/10.1016/j.ijpharm.2010.02.025>
- Westerlund, I., von Heijne, G., & Emanuelsson, O. (2009). LumenP-A neural network predictor for protein localization in the thylakoid lumen. *Protein Science*, 12(10), 2360–2366. <https://doi.org/10.1110/ps.0306003>
- Woollacott, I. O. C., & Rohrer, J. D. (2016). The clinical spectrum of sporadic and familial forms of frontotemporal dementia. *Journal of Neurochemistry*, 138, 6–31. <https://doi.org/10.1111/jnc.13654>
- Yang-hartwich, Y., Bingham, J., Garofalo, F., Alvero, A. B., & Mor, G. (2015). Apoptosis and Cancer, 1219, 75–86. <https://doi.org/10.1007/978-1-4939-1661-0>
- Yang, C., Tan, W., Whittle, C., Qiu, L., Cao, L., Akbarian, S., & Xu, Z. (2010). The C-terminal TDP-43 fragments have a high aggregation propensity and harm neurons by a dominant-negative mechanism. *PLoS ONE*, 5(12). <https://doi.org/10.1371/journal.pone.0015878>
- Zhang, Y. J., Caulfield, T., Xu, Y. F., Gendron, T. F., Hubbard, J., Stetler, 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>

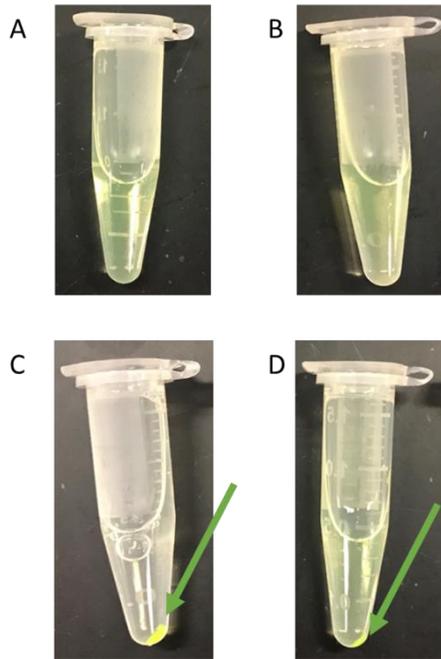
APPENDIX  
Supplemental Information



**Figure S1. Nickel affinity chromatogram of a denaturing urea purification with WT TDP-43.** First urea purification attempt reveals no binding to the column. mAu of sample presented in teal. %B (mixture of lysis wash and elution wash) presented in green.



**Figure S2. Nickel affinity chromatogram of modified denaturing urea purification.** Modified purification method (Materials and Methods) succeeded in binding recombinant protein vector and washing off the Ni/NTA column.



**Figure S3. Visual Fluorescence Assay.** Visual comparison of fluorescence intensity of GFP-TDP-43 samples with and without TG<sub>12</sub> DNA. A) Before *in vitro* aggregation study of GFP-TDP-43 No DNA B) Before *in vitro* aggregation study of GFP-TDP-43 + TG<sub>12</sub> DNA C) After *in vitro* aggregation study of GFP-TDP-43 No DNA D) After *in vitro* aggregation study of GFP-TDP-43 + TG<sub>12</sub> DNA. Green arrow identifies pellet after centrifugation.



<b>Chakrabartty</b>	<b>Experiment</b>	<b>This Lab</b>
pET30 Vector with vYFP-TDP43	<b>Protein Purification</b>	LIC 1-GFP Vector with TDP43WT ORF inserted
BL21 Cells		Rosetta Cells
Lysis Buffer: 40mM HEPES-KCL; 500mM KCL; 20mM Imidazole; 20mM MgCl <sub>2</sub> ; 2mM βMe; 10% Glycerol; Protease inhibitor		Lysis Buffer: 40mM HEPES; 300mM NaCl, 10mM MgCl <sub>2</sub> ; 5mM βMe; 50mM Imidazole; 10% Glycerol; Protease Inhibitor
Lysed by sonication		Lysed by 10% CHAPS
Batch purification with Ni-NTA beads		FPLC Purification with 8mL Ni-NTA Column
Eluted with 250mM Imidazole		Eluted with 500mM Imidazole
Final Product: Purified Protein		Final Product: "Dirty" protein sent for further purification on a Size Exclusion Column
20°C		<b>DLS</b>
DynaPro DLS Instrument (Quartz Cuvette, assume NanoStar)	DynaPro Plate Reader 2	
10 sec x 10 measurements	5 seconds x 5 measurements	
Used for Analysis	<b>Size Exclusion Chromatography</b>	Used for Purification and Analysis
40uL of sample		5mL of Sample
Superdex75 10/300		Superdex 16/200
20uM Protein Sample Loaded On		unknown protein concentration
10 fold Dilution of 20μM Stock	<b>In Vitro Aggregation</b>	Dilution of stock to 3μM
Original Buffer: 40mM HEPES-KCL; 500mM KCL; 250mM Imidazole; 20mM MgCl <sub>2</sub> ; 2mM βMe; 10% Glycerol		Original Buffer: 40mM HEPES; 300mM NaCl; 5mM DTT
Final Buffer: 36mM HEPES-KCL; 170mM KCL; 25mM Imidazole; 18mM MgCl <sub>2</sub> ; 1.8mM βMe; 1% Glycerol		Final Buffer: 36mM HEPES; 170mM NaCl; 2mM DTT
Incubated at 25°C		Incubated at 25°C
Time: 4 hours		Time: 960 Min
		Agitation at 1400RPM

**Table S4. Comparison of previous experiments and my experimental method.**