10-22-2015 12:00 AM

Interdomain interactions of the transactive response DNA binding protein 43 kDa (TDP-43)

Karen M. Dunkerley, *The University of Western Ontario*

Supervisor: Dr. Stanley D. Dunn, *The University of Western Ontario*
Joint Supervisor: Dr. Michael J. Strong, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Biochemistry

Interdomain interactions of the transactive response DNA binding protein 43 kDa (TDP-43)

(Thesis Format: Monograph)

by

Karen M. Dunkerley

Graduate Program in Biochemistry

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

Transactive response DNA binding protein, 43 kDa (TDP-43) is 416-residue RNA processing and transport protein, observed in insoluble cytoplasmic aggregates within affected neurons in neurodegenerative diseases. TDP-43 has three domains: the N-terminal (N), RNA-binding (R) and unstructured C-terminal domains (G). Unstructured domains often form intramolecular interactions regulating other domains; our goal was to determine if such an interaction occurs in TDP-43. In Far Western blots, tagged NR was observed to bind to G. A ten residue C-terminal truncation of G virtually abolished binding and introduction of phosphomimetics at Ser409/Ser410 also reduced binding. Sedimentation velocity ultracentrifugation with tagged NR and G also revealed interaction, observed by a shift in sedimentation coefficients when compared to those of the individual polypeptides. In vivo colocalization studies confirmed a cellular interaction between fluorescently labeled NR and G. This interaction has potential implications for the regulation of TDP-43 and the mechanism of generation of aggregative forms.

Keywords: TDP-43, intrinsically disordered protein, analytical ultracentrifugation, protein interactions, Far Western blot, amyotrophic lateral sclerosis

Dedication

For my parents.

Their love and support means the world to me. Thank-you for always believing in me and having delicious leftovers in the fridge for me after a late day at the lab.

I love you.

Acknowledgments

I would like to thank my supervisors, Dr. Stanley Dunn and Dr. Michael Strong. Their support, guidance and advice throughout my Master's degree have been an immense help through planning experiments, interpreting data and completing my written thesis. I have learned so much from their leadership as a scientist and will carry it through my future endeavours.

Many thanks to my advisory committee members, Dr. Brian Shilton and Dr. James Choy for their understanding and input during meetings. Their suggestions aided in shaping my thesis through focus and encouragement.

I would also like to thank Dr. Kathryn Volkening for her assistance in the lab and her efforts in editing my thesis. Her suggestions and insight greatly improved the final message of my thesis and emphasized the importance to the field.

To Yumin Bi, whose assistance in the lab was irreplaceable through maintaining the necessary components to keep the Dunn Lab going and for always being available for questions. Her endless contributions through cloning, purifying proteins and always bringing a pile of spring rolls to lab potlucks, I will never forget.

I would like to thank Cheryl Leystra-Lantz and Dr. Cristian Droppelmann for their efforts in performing the cloning and colocalization work that provided critical biological evidence for my conclusions. Thank-you Lee-Ann Briere for your training and critical eye in interpreting the complicated world of ultracentrifugation data. This one was not easy.

This work would also not be complete without the help of past and present undergraduate researchers who dedicated their time to solubilizing the disordered

domain. To Christina Chung, Thamiya Vasanthakumar and Jethro Kwong, I thank-you.

The graduate experience would not be complete without networking and developing bonds with fellow graduate students and lab colleagues. Thank-you Safee Mian for listening when my experiments were troubling and offering advice on how to make it better. Thank-you also to Alex Moszczynski, Kevin Cheung, Michael Tavolieri, Sali Farhan, Dr. Danae Campos-Melo and Wendy Strong for their observations during lab meetings and excellent company at conferences.

To my friends in graduate school and beyond, you know who you are, thank-you for making this time unforgettable.

To everyone who has provided me guidance and assistance throughout this work, thank-you and I am ever grateful.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2xYT: yeast tryptone extract media

3’ UTR: 3’ untranslated region

ALS: amyotrophic lateral sclerosis

NH_4SO_4: ammonium sulfate

Arg (R): arginine

Asn (N): asparagine

Asp (D): aspartic acid

AUC: analytical ultracentrifugation

BRCA1: breast cancer 1

BSA: bovine serum albumin

CaMKI: calcium/calmodulin-dependent protein kinase type 1

CD: circular dichroism

CDC7: cell division cycle 7-related protein kinase

CDK: cyclin-dependent kinase

CF: cystic fibrosis

CFTR: cystic fibrosis transmembrane conductance regulator

Cys (C): cysteine

ddH_2O: double deionized H_2O

DNA: deoxyribonucleic acid

DTT: dithiothreitol

EDTA: ethylenediaminetetraacetic acid

ETF1: eukaryotic translation termination factor 1

FERM domain: 4.1, ezrin, radixin, moesin domain

FOXO: forkhead box protein O1, 3a, 4, 6

FTD: frontotemporal dementia

G: residues 262-414 of TDP-43, the C-terminal, glycine-rich domain

G3BP1: GTPase activating protein (SH3 domain) binding protein 1

Gln (Q): glutamine

Glu (E): glutamic acid

Gly (G): glycine

GSK-3β: glycogen synthase kinase 3 beta

GuHCl: guanidine hydrochloride

HDAC6: histone deacetylase 6

HEK293T: human embryonic kidney 293 T antigen cells

His (H): histidine

HIV: human immunodeficiency virus

hNFL: human low molecular weight neurofilament

hNEFL: gene encoding the human low molecular weight neurofilament protein

hnRNP: heterogeneous nuclear ribonucleoprotein

HPLC: high performance liquid chromatography

HT: His_6-thioredoxin

IDPR: intrinsically disordered protein region

Ile (I): isoleucine

IPTG: isopropyl β-D-1-thiogalactopyranoside

IX: ion exchange

Leu (L): leucine

Lys (K): lysine

Met (M): methionine

N: residues 1-101 of TDP-43, the N-terminal domain

NES: nuclear export signal

NLS: nuclear localization signal

NMR: nuclear magnetic resonance

PABP: polyadenylate-binding protein

Phe: phenylalanine

PMSF: phenylmethylsulfonyl fluoride

Pro (P): proline

PTPIP51: Protein Tyrosine Phosphatase-Interacting Protein 51

PVDF: polyvinylidene fluoride

R: residues 102-261 of TDP-43, the RNA-binding domain

RE: restriction endonuclease

RNA: ribonucleic acid

RNP: ribonucleoprotein

RRM: RNA recognition motif

RRS: ribosome reinitiation site

RXRG: retinoid X receptor, gamma

SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis

SE: sedimentation equilibrium

Ser (S): serine

SH-SY5Y: neuroblastoma derived, neuron-like cell line

Src: Proto-oncogene tyrosine-protein kinase Src

ssDNA: Single stranded DNA

SV: sedimentation velocity

SxI: Protein sex-lethal

TCEP: tris (2-carboxyethyl) phosphine

TDP-43: transactive response DNA binding protein 43 kDa

TFA: trifluoroacetic acid

Thr (T): threonine

TIA-1: cytotoxic granule-associated RNA binding protein

Tris base: tris (hydroxymethyl) aminomethane

Tris-HCl: tris (hydroxymethyl) aminomethane hydrochloride

Trp (W): tryptophan

Tyr (Y): tyrosine

VABP: vesicle-associated membrane protein-associated protein B

UPS: ubiquitin proteasome system

Val (V): valine

Chapter 1: Introduction

In devastating neurodegenerative diseases including amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), transactive response DNA binding protein 43 kDa (TDP-43) has been identified as a hallmark component of pathological intracellular inclusions found in patients (1–3). These TDP-43 positive inclusions are typically found in both upper and lower motor neurons in the spinal cord, cortex and hippocampus of affected individuals (4–6). Initially discovered as bound to transactive response sequences in HIV DNA (7), TDP-43 has since been implicated in multiple other diseases including cystic fibrosis (8, 9), Alzheimer's disease (10, 11), Parkinson's disease (10) and Alexander's disease (12). Incomplete characterization of the structure and function of TDP-43 has thus far prevented a thorough understanding of its roles in healthy and disease states.

TDP-43 is a member of the heterogeneous nuclear ribonucleoprotein (hnRNP) family of proteins (7). hnRNPs have a characteristic RNA-binding domain, comprised of highly structured nucleotide binding motifs called RNA recognition motifs (RRMs), followed by an unstructured, glycine-rich C-terminal domain (8, 13–15). The domain organization of TDP-43 is closely related to that of the hnRNP A/B type, as demonstrated in the sequence alignment in Figure 1A&B (14, 15). In addition to the RNA-binding and unstructured domains, TDP-43 has a structured domain N-terminal to the RNA-binding domain, called the N-terminal domain, as outlined in Figure 1A.

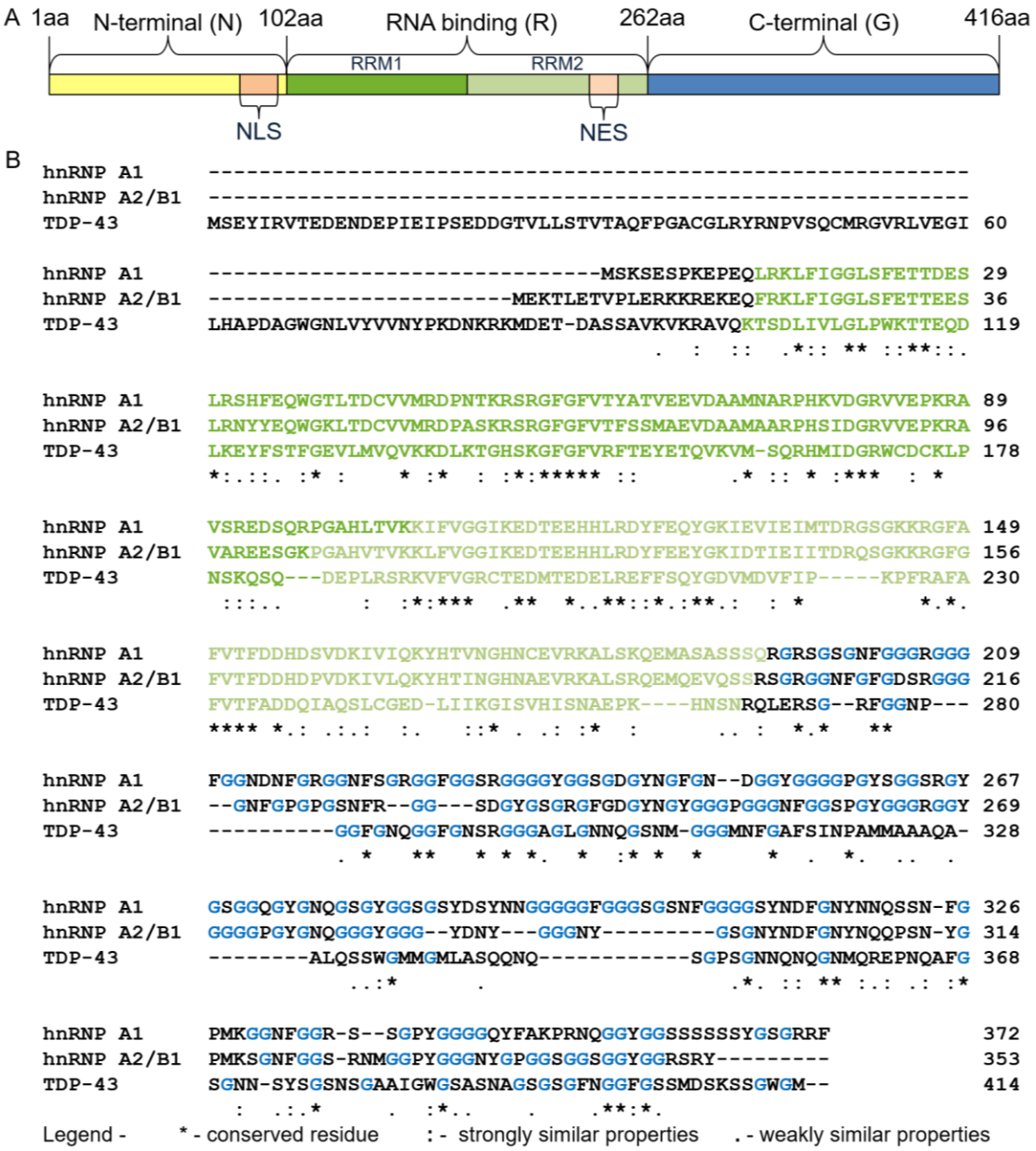


Figure 1. Sequence illustration of the hnRNP family. A) Schematic diagram of TDP-43 domain boundaries. Each domain is highlighted and named. The colours will be carried throughout this work to represent each domain in further diagrams. NLS-Nuclear localization signal; NES-Nuclear export signal; RRM-RNA recognition motif. B) Sequence alignment of two hnRNPs, A1 and A2/B1, and the 414 residue TDP-43. The two RRM domains are highlighted in dark and light green. The glycine residues are coloured blue.

The human *TARDBP* gene is located on chromosome 1 at 1p36.21 and contains 6 exons that can be alternatively spliced, encoding at least five TDP-43 variants in human cells (13). Orthologous TDP-43 genes have been identified in other eukaryotes including mouse, *Caenorhabditis elegans* (*C. elegans*) and *Drosophila melanogaster* (13). Human TDP-43 has high sequence identity with its orthologous proteins in *Drosophila* (59%) and *C. elegans* (38%) (16). The greatest sequence variation in orthologous TDP-43 proteins is in the C-terminal domain which may cause variation in the functions of TDP-43 across species such as RNA-binding but this is still not well understood (13). The most commonly reported isoform of TDP-43 is 414 amino acids in length and it is recorded as isoform 1 by the UniProt database (17). A second isoform of 416 amino acids was the primary species isolated from healthy human brain (18) by this lab and it is the variant used throughout this work.

TDP-43 is a primarily nuclear protein and is involved in many aspects of RNA processing including transcription, stability, splicing and transport. Through both nuclear export (NES) and localization signals (NLS) (Figure 1A), TDP-43 moves between the nucleus and the cytoplasm (19). The RNA-binding domain of TDP-43 binds $(UG)_n$-rich sequences in its target mRNA species; these sequences are most commonly located in the 3'UTR or at splicing junctions (8, 20). Similar to other hnRNPs, the C-terminal domain of TDP-43 has been demonstrated to bind with many other proteins and has multiple sites for post-translational modification (21). The C-terminal domain is intrinsically disordered and may also modulate function through intramolecular interactions. Many studies have revealed the

intrinsic aggregative properties of TDP-43 through both *in vitro* and *in vivo* methods that replicate the TDP-43 positive inclusions found in neurodegenerative diseases (22–28). These inclusions contain hyperphosphorylated and polyubiquitinated, full-length and C-terminal fragments of TDP-43, as well as RNA species and proteins found in stress granules (19).

Whether some degree of self-interaction/aggregation is functional or strictly pathological is not known. If self-interaction is not functionally relevant, then pathological conditions may arise from natural protein aggregation due to excess protein in the cell and the aggregation prone nature of the C-terminal domain. However, if self-interaction is functional for increasing RNA stability or transport, then it would have to be tightly regulated to avoid excessive interactions which may lead to aggregation. Developing a more basic understanding of TDP-43 is necessary for understanding how TDP-43 correctly forms protein interactions and whether mutations or modifications to TDP-43 that are associated with disease-related aggregation alter this process. Our overall hypothesis is that the domains of TDP-43 do participate in intramolecular interactions.

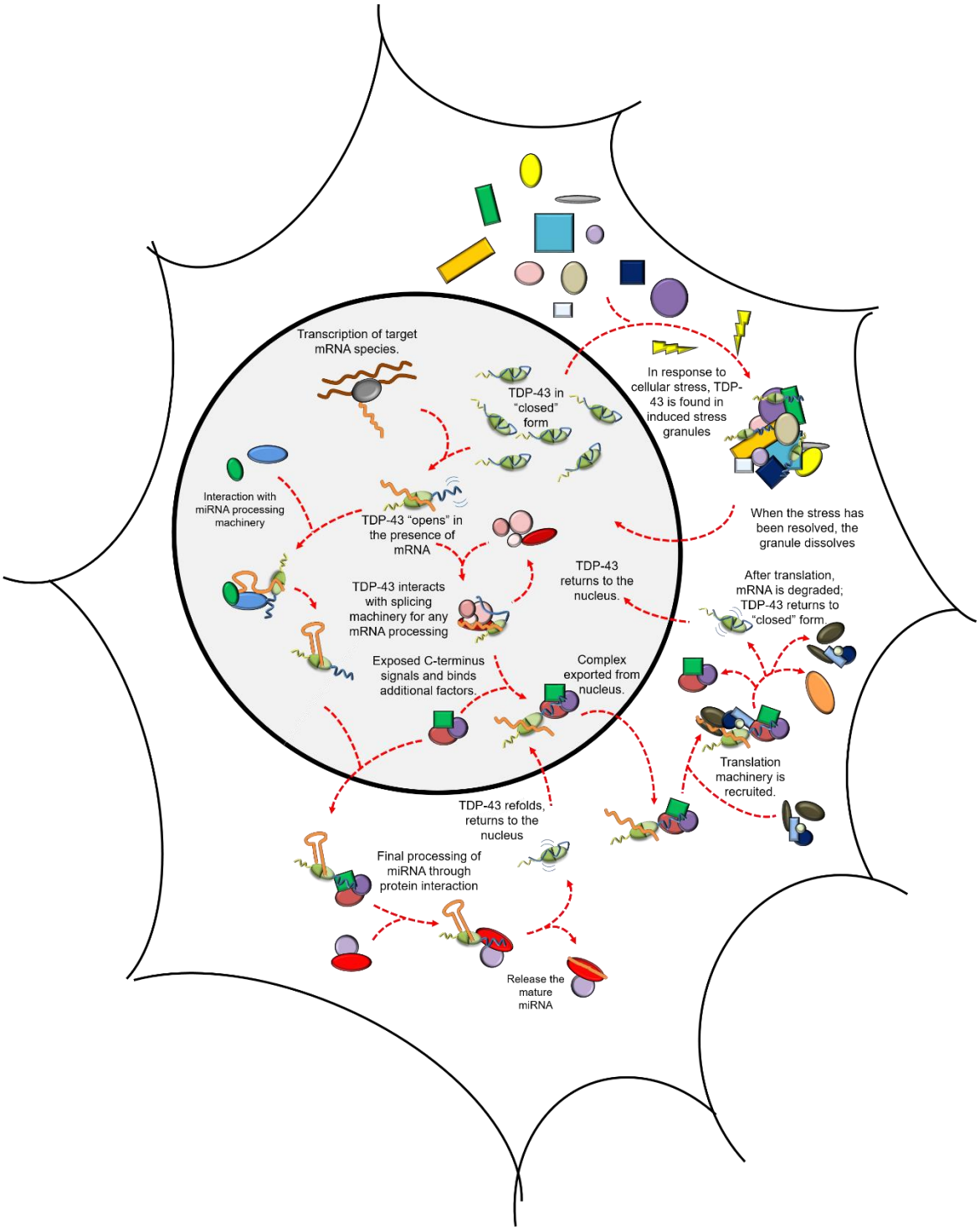
1.1 Cellular Functions of TDP-43

TDP-43 was initially described as repressing the expression of HIV genes through binding the TAR DNA sequence and blocking the transcription factors that would initiate expression of HIV-1 genes (7). Since then, TDP-43 has been shown to be associated with splicing regulation, mRNA stability, transcriptional regulation, microRNA processing and the cellular stress response (Figure 2) (reviewed in 29).

When TDP-43 becomes aggregated in disease states, it is plausible to assume that normal function is reduced or lost. Gaining further understanding of the normal functions of TDP-43 is critical to discovering the cellular basis of TDP-43 aggregation that leads to toxicity because very little is known about the regulation of cell localization or of protein interactions in a functionally normal state.

1.1.1 TDP-43 localization—In a global analysis of the TDP-43 interactome, two protein networks were strongly represented in the interactions with TDP-43: Nuclear/transcriptional proteins and cytoplasmic/translational proteins (30). In the first group, interacting proteins include transcription factors, hnRNPs, splicing factors and RNA transport proteins. This does not come as a surprise since TDP-43 has already been observed to play a critical role in many of these nuclear RNA processes, and the experimental protocol lacked ribonuclease. The discovery of interactions with the cytoplasmic translation machinery and ribosomal subunits suggests TDP-43 stays associated with its target mRNA and continues to have a role outside the nucleus (30). With roles in different compartments of the cell, the NLS and NES within TDP-43 are necessary for TDP-43 distribution within the cell (31). Perturbing either of these signals by mutations causes TDP-43 to aggregate in the nucleus or cytoplasm of primary neurons, sequestering endogenous TDP-43 with it (31). If mutations to the localization signals can have such a strong effect on TDP-43, control over cellular location must be critical to its physiological function.

Figure 2. Schematic diagram of the various functions of TDP-43 within a cell. TDP-43 is primarily found within the nucleus (shaded area), where it participates in many functions in mRNA processing including mRNA stability, splicing and miRNA processing. TDP-43 can also be transported from the nucleus to the cytoplasm where it is involved in further processes including miRNA processing and mRNA translation. TDP-43 is also a component of stress granules. After its cytoplasmic function is complete, TDP-43 can then be transported back into the nucleus. Shapes not corresponding to TDP-43 are protein placeholders to illustrate the functions described by this figure. The functions illustrated include the hypothesis of this thesis.


Transcription of target mRNA species.
TDP-43 in "closed" form
In response to cellular stress, TDP-43 is found in induced stress granules
Interaction with miRNA processing machinery
TDP-43 "opens" in the presence of mRNA
When the stress has been resolved, the granule dissolves
TDP-43 returns to the nucleus.
After translation, mRNA is degraded; TDP-43 returns to "closed" form
TDP-43 interacts with splicing machinery for any mRNA processing
Exposed C-terminus signals and binds additional factors.
Complex exported from nucleus.
Translation machinery is recruited.
TDP-43 refolds, returns to the nucleus
Final processing of miRNA through protein interaction
Release the mature miRNA

1.1.2 Splicing—A well-characterized function of TDP-43 is its effect on splicing of the cystic fibrosis transmembrane conductance regulator (*CFTR*) mRNA (8). This is mediated by TDP-43 binding to a UG repeat rich region at the intron 8/exon 9 junction of the *CFTR* gene, where TDP-43 binding inhibits splicing. When TDP-43 binds at a junction, the spliceosome is blocked from acting on this junction and skips to the next intron, effectively removing exon 9 and producing a non-functional CFTR protein (8). Deletion of the TDP-43 C-terminal domain prevented both human and *Drosophila* TDP-43 from properly interacting with, and splicing, the target *CFTR* mRNA (13). *C. elegans* TDP-43 does not have the same disordered C-terminal domain and also does not participate in splicing (13). Therefore, the C-terminal domain of TDP-43 is likely important for interaction with the splicing machinery (8, 13). TDP-43 has also been found to have an effect on splicing of mRNA for other genes including *BRCA1, ETF1* and *RXRG* (32). The ability of TDP-43 to interact with hnRNPs and other nuclear factors is required for proper splicing function (33, 34). Disruption of interactions, including induction of cytoplasmic aggregates, removes the functional pool of nuclear TDP-43 and reduces the ability of TDP-43 to aid in splicing (35).

1.1.3 mRNA stability—Along with splicing, TDP-43 has a role in mRNA stability. Human neurofilament (*hNEFL*) mRNA was co-immunoprecipitated with TDP-43 (18). The binding region was mapped to the 3'UTR of *hNEFL*, where proper folding of the mRNA creates a $(UG)_6$ sequence for TDP-43 to bind (36). Binding and stabilization by TDP-43 is necessary for proper hNEFL expression through

interaction with other cellular RNA processing machinery (18). TDP-43 also binds with *HDAC6* mRNA and is necessary for maintaining *HDAC6* mRNA and protein levels *in vivo* (37). Knockdown of TDP-43 decreased the amount of HDAC6 in HEK293 and SH-SY5Y cells and in *Drosophila* (37). Knockdown of TDP-43 also induced instability of *Atg7* mRNA, which is the transcript for a protein critical to autophagy (38). If autophagy is compromised, aggregates of TDP-43 would not be efficiently cleared from cells, inducing further toxicity (38). Conversely, TDP-43 knockdown increased the mRNA and protein levels of Cdk6 (39). *CDK6* mRNA, like *hNEFL* mRNA, has TDP-43 recognition sequences in the 3'UTR but the opposite effect of TDP-43 expression on Cdk6 protein expression indicated a more complicated role for TDP-43. Perhaps the most important known mRNA interaction of TDP-43 is with the *TARDBP* transcript itself; TDP-43 negatively regulates its own expression by binding a recognition sequence in the 3'UTR and destabilizing the *TARDBP* mRNA (40). This negative feedback loop suggests the importance of closely regulating TDP-43 levels so that upon TDP-43 binding with a mRNA substrate, the subsequent actions are correct for the fate of the bound mRNA species.

1.1.4 microRNA processing—In a TDP-43 knockdown model in Hep-3B cells, many miRNAs were discovered to be significantly decreased (41). TDP-43 interaction with the Drosha and Dicer complexes is critical in miRNA biogenesis (42). Immunoprecipitation with a FLAG antibody (anti-FLAG) pulled down a FLAG-tagged Drosha and its complex where a direct interaction between Drosha and

TDP-43 was observed, regardless of the presence of RNA, through TDP-43's C-terminal domain (42). A similar direct interaction was found for cytoplasmic TDP-43 and FLAG-Dicer (42). In the reciprocal experiment, anti-FLAG pulled down FLAG-TDP-43 with Drosha or Dicer (42). Deletion of the C-terminal domain prevented interaction with the splicing machinery (42). When TDP-43 was knocked down in Neuro2A cells, miRNA production critical for neurite outgrowth was reduced (42).

1.1.5 Cellular stress response—Another area of increasing interest is a role for TDP-43 in the cellular stress response (43, 44). After exposure to various stressors including sodium arsenite, thapsigargin and heat shock, TDP-43 and hnRNP A2 were sequestered to stress granules (44). In TDP-43 depleted cells, stress granules took longer to form and were smaller in size (44). In addition to granule formation, loss of TDP-43 also reduced expression of the key stress response proteins, TIA-1 and G3BP (44). When the C-terminal domain of TDP-43 was deleted, stress granules failed to form after exposure to sorbitol (43). Another study has shown that the family of FOXO transcription factors are activated by TDP-43 after exposure to stress (45). The stress response triggers translocation of TDP-43 to the cytoplasm, where it relieves inhibition of FOXOs by 14-3-3. TDP-43 competitively interacted with 14-3-3, which allowed the FOXO proteins to return to the nucleus and as a result expression of stress response proteins occurred (45). TDP-43 and 14-3-3 then associated with stress granules to sequester unneeded mRNAs from translation (45). Additionally, TDP-43 may be involved in

disrupting endoplasmic reticulum/mitochondria contacts through dissociation of VAPB and PTPIP51, following activation of GSK-3β (46). Loss of the interaction between VAPB and PTPIP51 causes an imbalance in Ca^{2+} homeostasis and activation of a stress response (46).

1.2 Domain Structure and Function

As previously stated, TDP-43 has three domains: the N-terminal, RNA-binding and C-terminal domains. Figure 3 shows the amino acid sequence of TDP-43 and highlights the important regions and residues for structure and function as currently known.

1.2.1 N-terminal domain—The N-terminal domain (N) is classically defined as the first 101 amino acids of TDP-43, and includes a putative NLS at position 82-98 (47, 48). The NLS is critical to TDP-43 function as while TDP-43 is primarily found in the nucleus, it shuttles to and from the cytoplasm in response to cellular stress or injury (47). Another interesting region within the N-terminal domain is a highly conserved, highly acidic region (residues 9-23) but the exact function of this region is not known (8, 9). Online software such as PredictProtein (www.predictprotein.org) predicts the formation of β–sheets and α-helices within the N-terminal domain of TDP-43 (49).

Recent nuclear magnetic resonance (NMR) and circular dichroism (CD) studies on the N-terminal domain indicate the presence of secondary structures as predicted (48, 50, 51). The NMR analysis on a polypeptide of TDP-43 residues 1-

105 indicated the presence of a structured domain with some helices but the actual structure formed could not be determined, as the workable concentration was too low, however there was likely some presence of oligomers (50). CD analysis on the same polypeptide indicated the presence of both β–strands and α-helices (50). Similar studies on residues 1-102 of the N-terminal domain suggested the existence of two conformations in equilibrium, a folded and an unfolded form, that were not evident in a construct consisting only of residues 1-80 (51). The structure was solved by NMR and the folded form was suggested to have a ubiquitin-like fold (51). The N-terminal domain also did not significantly contribute to oligonucleotide binding (50).

Co-immunoprecipitation of differentially tagged TDP-43 proteins revealed the ability of TDP-43 to form dimers, where the N-terminus was necessary for dimerization (48, 52, 53). Further investigation through deletion mutants and *in vivo* crosslinking showed the first 10 amino acids of the N-terminal domain to be essential for dimer formation (48). *In silico* modeling also demonstrated that these amino acids may form a loop structure that positions the rest of the TDP-43 molecule for correct interaction with another TDP-43 molecule (48).

1.2.2 RNA-binding domain—The RNA-binding domain (residues 102-262) consists of two highly conserved RRMs: RRM1 and RRM2. The sequences are highlighted in green in Figure 3 and the solution structure is shown in Figure 4. RRMs are common in many well characterized nucleotide binding proteins,

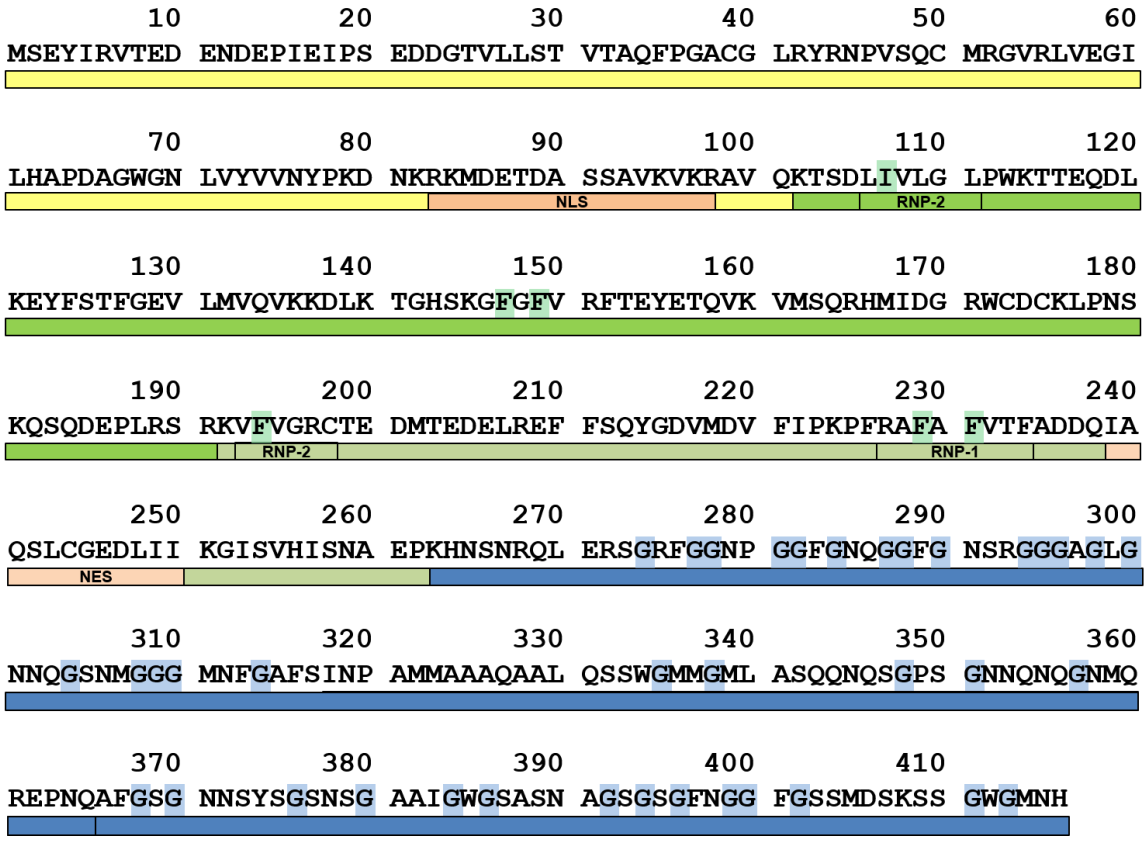


Figure 3. Amino acid sequence of the 416 residue TDP-43 isoform. Highlighted are some important sequences and residues that are mentioned throughout this work. The N-terminal domain (yellow) also contains the nuclear localization signal (NLS) (pink). The RNA-binding domain (green) is made up of two RNA recognition motifs (RRMs) in dark green, RRM1, and light green, RRM2. Within each RRM, are two highly conserved ribonucleoprotein sequences (RNP-1 and RNP-2) that contain the critical residues for RNA-binding (highlighted in green). The C-terminal domain (blue) is glycine rich (highlighted in blue) and contains the hydrophobic patch and the glutamine/asparagine rich (Gln/Asn-rich) region.

including hnRNPs, Sxl and PABP (54). Structurally, RRMs fold with a 4-strand β-sheet, connected by 3 α-helices Figure 4B (54). Tandem RRMs, as found in TDP-43, typically increase the binding affinity of the protein to mRNA (54). RRM2 also contains the putative NES (residues 239-250) (47). The two RRMs of TDP-43 appear to have different stabilities when exposed to denaturing conditions (55). Structural studies using CD revealed an intermediate state in the melting of RRM2 that was not evident in the CD spectrum of RRM1 during chemical denaturation. Most importantly, tethering the two RRMs with the natural linker sequence increased stability compared to RRM1 alone (55). When RRM1 and RRM2 are tethered, the hydrophobic residue clusters in RRM2 make contacts with RRM1, contributing to stability (55). Disrupting the native structure of RRM2 by a cleavage N-terminal to Arg208 increased the population of the intermediate state and this mutant had a higher affinity for RNA than full RRM2 (55). It is unclear whether the intermediate state is functionally relevant.

In 2009, the first crystal structure of TDP-43 was solved, with RRM2 only (residues 192-265) in complex with ssDNA (56). Since then, the solution NMR structure of the full RNA-binding domain, both RRM1 and RRM2 bound to an RNA species (5'- GUGUGAAUGAAU-3'), was solved (Figure 4) (57). This new structure was critical for gaining a better understanding of the RNA sequence specificity of TDP-43. The RNA molecule was bound across the β-sheet surfaces of both RRMs with the guanine in position 5 inserted between the two domains (Figure 4A). This base made contacts with both RRMs, appearing to hold them in a rigid structure.

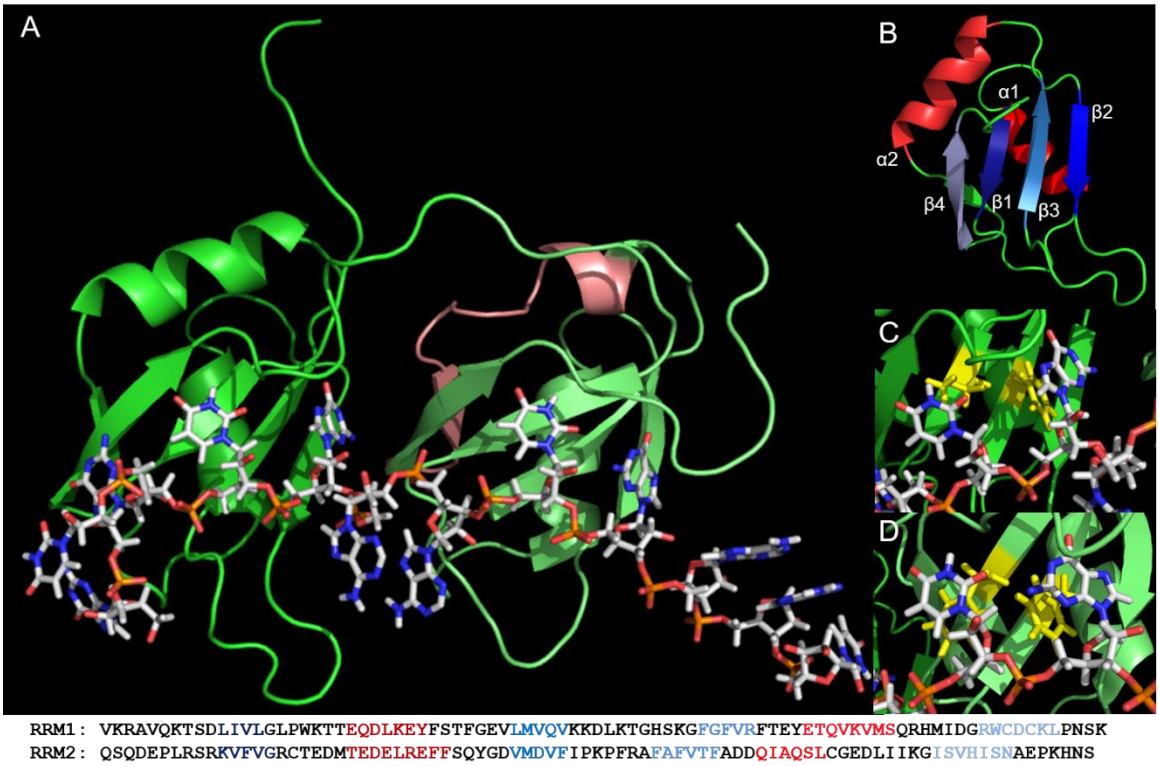


Figure 4. Solution NMR structure of the RNA-binding domain (residues 96-267) of TDP-43 (PDB: 4BS2). A) Full structure of this domain bound to an RNA species, sequence 5'-GUGUGAAUGAAU-3': RRM1 is displayed in darker green on the left, bound with the 5' end of the RNA, RRM2 is in light green. The NES is indicated by the pink region within RRM2. B) A single RRM and the significant secondary structures. The α-helices and β-sheets are colour coded in the corresponding sequences at the bottom. C) Close up of RRM1 and the residues that make base-specific contacts with the RNA species: I107 and F149 (yellow). D) Close up of RRM2 and the residues that make base-specific contacts with the RNA species: F194 and F231 (yellow). This figure was constructed with Pymol.

Only bases G1, G3, U4, G5, U8 and G9 made base-specific contacts with the RRMs. The 5' nucleotide base bound with RRM1, while the downstream nucleotides bound to RRM2; this is opposite to other proteins with tandem RRMs where the 5' base binds with the second RRM in the sequence (57).

Within each RRM of TDP-43, there are two conserved sequences called ribonucleoprotein sequences (RNP-1 and RNP-2). These have consensus sequences of approximately eight residues that are hydrophobic or positively charged to interact with nucleotides: KGFGFVRF for RNP-1 and LIVLGLPW for RNP-2 (29, 58). RNP-1 and RNP-2 are located in the β-sheet, creating a binding surface for ssDNA or RNA. Figure 4B and C highlights these residues and the direct contacts made with an RNA oligo, as observed in the NMR structure. In RRM1, Ile109 and Phe149 make direct contacts with bases U4 and G5, respectively. In RRM2, Phe194 and Phe231 make contacts with U8 and G9, respectively. These contacts appear like base-stacking interactions that precisely position the RNA and likely contribute to specificity (57). Introducing the mutations F147L and F149L to TDP-43 completely abolished binding with both $(TG)_6$ and $(UG)_6$ species (20, 59). Based on the NMR structure, the precise UG-rich consensus sequence for binding to TDP-43 was suggested to be 5'-GNGUGNNUGN-3' (16,17). Finally, RRM1 was crystallized alone with $(TG)_n$ ssDNA (59). The ssDNA bound to the β-sheet surface with the same orientation as observed in the NMR structure, where the nucleotides at G2, G4, C5 and G6 make direct contacts with the protein (59). This finding established a similar consensus sequence for ssDNA binding as with RNA.

Following the discovery of RRMs in TDP-43, oligonucleotide binding assays were conducted to determine the binding specificity of TDP-43. TDP-43 is capable of binding both $(UG)_n$ and $(TG)_n$ repeats in RNA and ssDNA, respectively (20, 56). Binding was observed with sequences of 3-12 repeats, where affinity increased as the oligonucleotide length increased. Using a truncated TDP-43 with only the RNA-binding domain, the K_d of $(UG)_n$ binding decreased from increasing oligonucleotide length from 3,060 nM (n=3) to 2.79 nM (n=8) (56). TDP-43 did not efficiently bind with non-$(UG)_n$ sequences that were tested (56). Deletion of RRM1, but not deletion of RRM2, abolished nucleotide binding and RRM1 alone bound $(UG)_n$/$(TG)_n$ repeats with a much higher affinity than RRM2 (20). Surprisingly, RRM2 bound more strongly with smaller oligonucleotides than larger, opposite to what was measured with RRM1 (56). In a crystal structure of RRM2 alone, dimers were observed but it is not clear whether this dimerization interface is biologically relevant or a crystal contact only (56). Crosslinking immunoprecipitation was used to find the binding sequences within the 3'UTR of *TARDBP* mRNA for a construct of the RNA-binding domain only (60). Of the two identified oligonucleotides, the longer sequence bound with higher affinity than the shorter, K_d = 112 nM and 262 nM, respectively (60). In addition, comparable affinities were measured for a similar construct with the residues Phe229 and Phe231 of RRM2 mutated to leucines. The Phe residues of RRM2 were not as critical to RNA binding, unlike the similar residues in RRM1 (60).

It appears that regions with a UG-rich sequence provide more favourable binding conditions for TDP-43. Protein binding is dynamic, therefore more

available binding sequences allow for regulation by TDP-43 to be maintained for a longer time (60). In addition, correct binding by RRM1 is most important for sequence recognition where as RRM2 merely contributes to binding.

1.2.3 C-terminal domain—Perhaps the most interesting domain of TDP-43 is the C-terminal domain, consisting of residues 263-414. This domain is glycine-rich (25%) and predicted to be intrinsically disordered by the online structure prediction software, DISOPRED (http://bioinf.cs.ucl.ac.uk/psipred/?disopred=1) (61, 62). Figure 5 shows the output from DISOPRED, where a probability of 1 is indicative of a likely region of disorder. The C-terminal domain, indicated by blue, is largely disordered. The region of predicted order within the C-terminal domain has been labelled a “hydrophobic patch” (residues 318-343; HP) and through solution NMR studies of a peptide encompassing resides 311-360, was determined to form a helix-turn-helix motif in solution (63). There is also a region rich in glutamine and asparagine (residues 344-365; Q/N) that has been implicated in nucleating TDP-43 aggregation (24, 64). These two regions are labelled in Figure 5.

The C-terminal domain has been observed as a key part of TDP-43 protein interactions. TDP-43 was first observed to interact with hnRNP A1, hnRNP A2/B1 and hnRNP C1/C2, while deletion of the C-terminal domain ablated the interaction (33). The region of interaction was more precisely mapped to residues 321-366 by truncating TDP-43 and measuring interaction with hnRNP A2 (65). Introducing a synthetic peptide corresponding to this region was even capable of disrupting TDP-43/hnRNP A2 interaction. This 45 residue region is highly conserved in

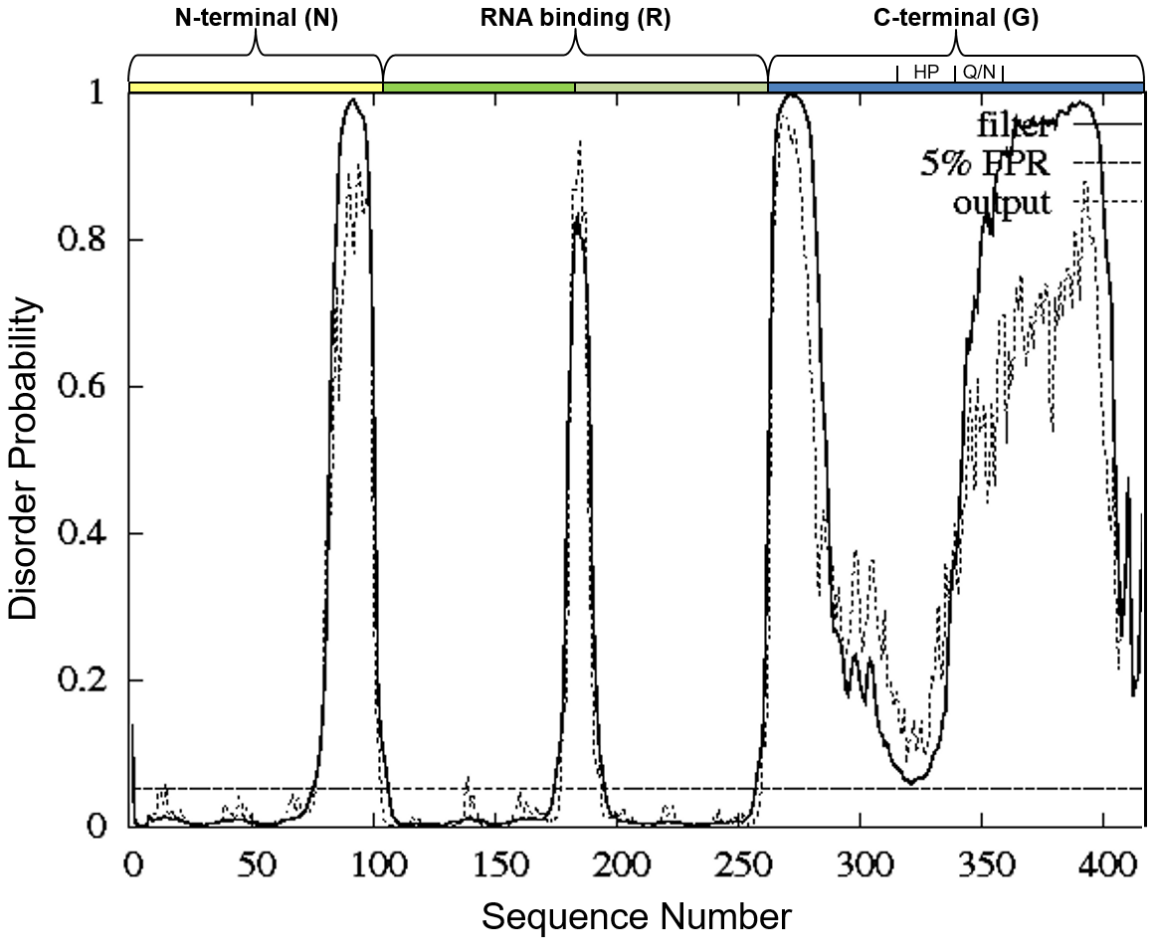


Figure 5. Disorder prediction and sequence outline of TDP-43 structure. A) The disorder prediction for the full TDP-43 sequence shows a high level of disorder within the C-terminal domain (blue). Calculated using DISOPRED software. Two regions of the C-terminal domain are indicated: the hydrophobic patch (HP), residues 318-343, and the Gln/Asn-rich region (Q/N), residues 344-365.

homologous TDP-43 proteins across species, demonstrating an importance for interaction with hnRNPs (65). Upon closer inspection, the C-terminal domain has a particularly glutamine/asparagine (Gln/Asn)–rich region (residues 344-365) located within the protein interacting region (24). Polyglutamine or Gln/Asn-rich regions are highly aggregation-prone and often nucleate toxic protein aggregates (66). C-terminal fragments of TDP-43 containing this region form insoluble inclusions in cells and even sequester endogenous TDP-43 in cells (24). The expression of 12 tandem repeats of residues 342-366 could form aggregates within cells and these aggregates could sequester endogenous TDP-43 (64).

In addition to the Gln/Asn-rich region, the aforementioned “hydrophobic patch” was also found to contribute to TDP-43 aggregation (63). A polypeptide of TDP-43 residues 311-360, which contained the hydrophobic patch, appeared to form some α-helical structure after analysis by CD and NMR spectroscopy (63). After incubation, the percentage of α-helical structure decreased while the β-sheet content increased, consistent with observations in other polyglutamine aggregate proteins (63, 67). Deletion of the hydrophobic patch decreased the amount of aggregation observed in cells (63). The combination of the hydrophobic patch and the Gln/Asn-rich region was named the “amyloidogenic core” for its resemblance to other β-amyloid aggregating proteins where α-helix to β-sheet structural transformation propagates aggregation (63, 68).

1.2.4 TDP-43 interactions and aggregation—TDP-43 protein is found aggregated in multiple diseases (2). The aggregates often contain C-terminal fragments of

TDP-43, cleaved at Arg^{208}, which is located in the first helix of RRM2 (2, 22). Overexpression of this fragment in cells triggered aggregation similar to that found in patient cases (22) while deletion of the C-terminal domain prevented it (69). Over 40 TDP-43 mutations, the majority of which are in the C-terminal domain (Figure 6), have been associated with disease (70). The different mutations had varying effects on aggregation by stabilizing or destabilizing structures that were considered relevant for aggregate formation (68).

While C-terminal mutations are not necessary for toxic inclusions to form, it is clear that the C-terminal domain is critical for normal structure and function of TDP-43. Its intrinsic tendency for protein interaction may lead to small oligomers that are biologically normal and pathogenic aggregation may only be due to dysregulation of this tendency. This is why a more basic understanding of the normal functions of TDP-43 will be critical for understanding its role in disease.

1.3 TDP-43 in Disease Pathology

As previously mentioned, TDP-43 was first discovered bound to TAR sequences of HIV-1 DNA, repressing transcription of these genes (7). Since then, the active involvement of TDP-43 in preventing HIV-1 replication in immune cells has been refuted (71) but evidence for TDP-43 involvement in the pathogenesis of other diseases has arisen.

In cystic fibrosis (CF), TDP-43 binds at the 3' splice site of exon 9 in *CFTR* mRNA, excluding exon 9 from the mature mRNA. The resulting protein is non-functional (3, 57). This splice site has a $(UG)_m(U)_n$ region in which polymorphisms

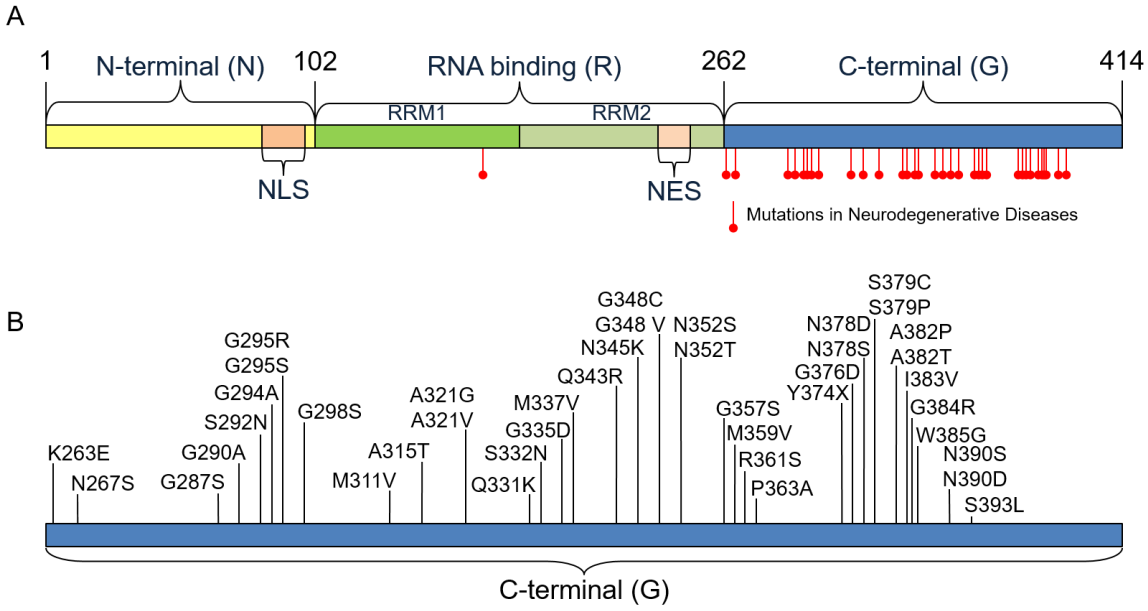


Figure 6. Mutations in TDP-43 in neurodegenerative diseases. A) Highlights the full sequence of TDP-43 where the identified mutations (|) are largely concentrated to the C-terminal domain. One mutation, D169G is in RRM1. B) Mutations currently associated with neurodegenerative diseases in the C-terminal domain (70).

introduce more or less favourable sites for regulatory factors to bind. Knockdown of TDP-43 in culture and patient cells, with or without unfavourable sequences, reduced the amount of misspliced *CFTR.* This finding suggests a critical role of TDP-43 in the regulation of splicing and a possible therapeutic target for patients with CF (72).

TDP-43 is also found in cellular inclusions in neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) (26, 53). These cytoplasmic inclusions contain TDP-43 that is often hyper-phosphorylated, ubiquitinated and N-terminally cleaved (1–3). TDP-43 cleavage, similar to that observed in patients, was induced *in vitro* after exposure to caspase-3 and caspase-7 (73, 74). There are three predicted cleavage sites within TDP-43 for caspase cleavage at $_{10}$DEND, $_{86}$DETD, $_{216}$DVMD with predicted cleavage products of 42, 35 and 25 kDa, respectively (73). Two fragments, 25-kDa and 35-kDa (73), of TDP-43 have been identified in patients with neurodegenerative diseases and similar polypeptides were observed experimentally to recapitulate the pathological aggregates observed in patients (26, 53). Caspase-3 and caspase-7 mediated cleavage was also induced in multiple cells lines including HeLa (cervical cancer) and THP-1 (acute monocytic leukemia) after a treatment to reduce the Ca^{2+} availability *in vivo* (74).

Further investigation into TDP-43 involvement in neurodegenerative diseases revealed cleaved and phosphorylated TDP-43 inclusions in the brains of Alzheimer's patients who developed cognitive impairment and patients with Lewy body dementia (11, 75). With over 40 TDP-43 mutations identified to date in

patients with neurodegenerative disease (Figure 6) (76–78), it is becoming evident that aggregation of TDP-43, and therefore a loss in critical RNA processing, has a profound effect on neuronal homeostasis, leading to disease.

1.4 Post-Translational Modifications of TDP-43

A critical step in TDP-43-related pathology is thought to be aberrant phosphorylation of multiple residues in the C-terminal domain including Ser379, Ser403, Ser404, Ser409 and Ser410 (79). The exact sites of phosphorylation were initially investigated through antibodies specific for phospho-epitopes of TDP-43. The 5 sites were confirmed to be phosphorylated in inclusions found in the brains of patients who were diagnosed with either ALS or FTD (79). Next, *in vitro* kinase assays were used to determine that casein kinase 1 (CK1) was likely responsible for phosphorylating TDP-43 and that aberrantly phosphorylated TDP-43 formed SDS-insoluble oligomers (79). Mass spectrometry confirmed the above 5 phosphorylated residues in recombinant TDP-43 treated with CK1, in addition to 24 other phosphorylated residues (80). This work involved *in vitro* incubation of CK1 with recombinant TDP-43 and therefore non-biologically relevant phosphorylation could have occurred. The kinase CDC7 may also be involved in phosphorylating TDP-43 (81).

Additional studies compared patient-derived inclusions with fresh rat brain and found phosphorylated Ser409 (pSer409) and phosphorylated Ser410 (pSer410) in the patient samples only (82). In *C. elegans* models of ALS (TDP-43 bearing a G290A, A315T, or M337V mutation) pSer409/pSer410 was detected, as

well as in the insoluble protein fraction of strains expressing wild-type TDP-43 (25). A phosphomimetic TDP-43 with Ser409Glu/Ser410Glu expressed in Neuro2A cells induced toxicity and cell death (83). These studies suggested that phosphorylation of these residues was not a normal post-translational modification of TDP-43.

There is currently no evidence of the level of *in vivo* phosphorylation in a healthy cellular environment. Without understanding this, it is difficult to understand exactly what sites are pathological during formation of TDP-43 aggregates or when this aggregation occurs. One theory is that as a result of disease pathogenesis, kinases become highly active and abnormal phosphorylation occurs thus leading to aggregation. In a neuroblastoma cell line, aggregates containing phosphorylated C-terminal fragments of TDP-43 (residues 220-414) were more slowly cleared by the ubiquitin-proteasome system (UPS) than unmodified fragments (84).

In an alternative theory, phosphorylation was proposed to occur after aggregation as a possible mechanism to facilitate clearance of the aggregates. In HEK293T and Neuro2A cells, C-terminal fragments (residues 252-414) formed cytoplasmic aggregates that were phosphorylated at the expected five Ser residues: 379, 403, 404, 409 and 410 (85, 86). The fragments were then mutated to contain either Asp/Glu (phosphomimetic) or Ala at the 5 phosphorylation sites. Fewer aggregates of the phosphomimetic C-terminal fragments were observed than the Ala fragments (86). The reduction in aggregate formation suggested phosphorylation occurs as a means of protection from larger aggregate formation

(85, 86). In a time course assay, more phosphorylated aggregates were detected at later than earlier time points (87, 88). However, phosphomimetics in full length TDP-43 did not show a significant reduction in aggregation (86).

Finally, disruption of the UPS with an inhibitor also led to increased phosphorylation and aggregate formation of mutant TDP-43 compared to cells with a functional UPS (89), suggesting that impaired protein degradation may also contribute to TDP-43 aggregation (88, 89).

Though there is significant conflicting information, it appears that phosphorylation of TDP-43, at least at Ser409/Ser410, is not a normal event for TDP-43 function. Elucidating the biochemical effect of phosphorylation would be an important piece of information in understanding the transition of soluble TDP-43 to aggregate.

1.5 Initial Evidence of Interactions

When work with TDP-43 began in the Dunn lab, it was observed that a GST-tagged, full-length TDP-43 protein (GST-NRG), expressed in *Escherichia coli*, was insoluble and spun down in the pellets of both low and high speed centrifugations (Figure 7A). Much of the TDP-43 literature has described the intrinsically disordered C-terminal domain as aggregation prone, suggesting this may be contributing to the insolubility of the full length TDP-43. Two models of aggregation were then considered for explaining the insolubility of full length TDP-43 as demonstrated in Figure 7B. The first suggests the C-terminal domain of one TDP-43 molecule interacting with the N-terminal or RNA-binding domain of another

TDP-43. The free C-terminal domain then binds with another TDP-43, causing a continually oligomerizing chain. The second depends only on the intrinsic aggregation of the C-terminal domain: many C-terminal domains interact, creating many punctate aggregates. Of the two models, the first would be more easily remedied by creating a split construct that would express the C-terminal domain as a separate polypeptide from the rest of the tagged TDP-43 molecule. This was achieved by inserting a ribosome re-initiation site (RRS) consisting of overlapping translational termination and initiation codons between the RNA-binding domain and C-terminal domain (Figure 7C). When this mRNA is expressed the ribosome would translate the GST-NR polypeptide and terminate at the stop codon, but remain bound to the mRNA due to the presence of a weak Shine-Delgarno site. Translation would then begin again, synthesizing the G polypeptide in a theoretical 1:1 ratio to GST-NR (Figure 7D). After induction of this plasmid and the same extraction conditions, the split in TDP-43 was observed to reduce the level of insoluble protein and both polypeptides remained in the high speed supernatant (Figure 7E). This dual expression allows for any interdomain interactions in TDP-43 while preventing the large aggregate formation (Figure 7F).

With soluble protein, the next step was to purify the two TDP-43 polypeptides separately. To achieve this, a number of column chromatography techniques were used, including ion exchange (DEAE-Sepharose; IX), to separate the polypeptides based on their differing intrinsic properties. What was observed was coelution of GST-NR and G, even after consecutive IX columns.

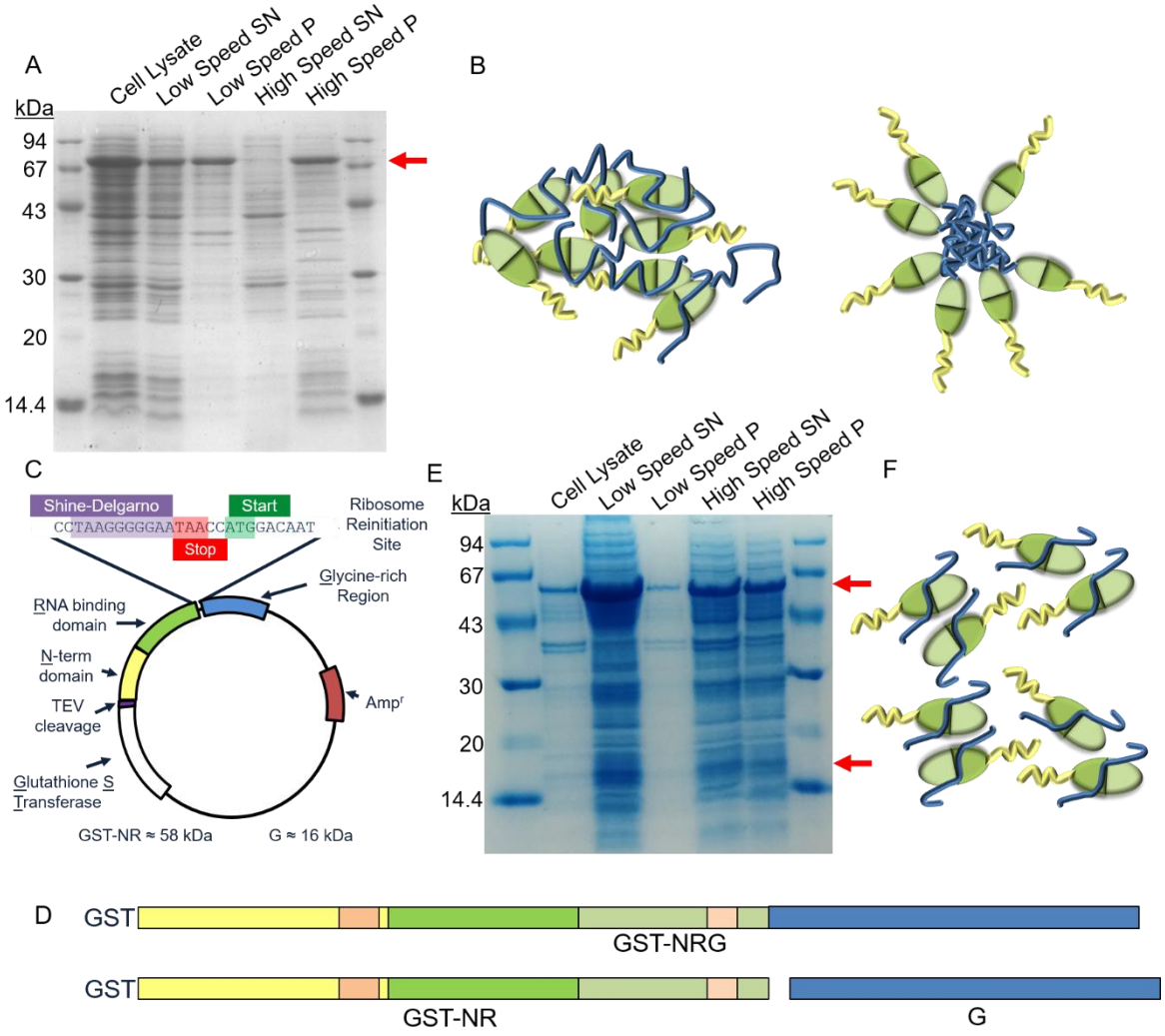


Figure 7. Justification for construction of the split TDP-43 construct, GST-NR_G. A) Coomassie stained gel showing the aggregated GST-TDP-43, located in the pellets of cell lysates (red arrow). B) Two proposed models of aggregation of TDP-43 (yellow: N-terminal domain; green: RRM domains; blue: C-terminal domain). C) Schematic diagram of the split TDP-43 plasmid. D) Schematic of the two polypeptides, GST-NR and G (bottom), created with the plasmid in C), compared to full length GST-TDP-43 (top). E) Coomassie stained gel of the now soluble GST-NR (top red arrow) and G (bottom red arrow) polypeptides. F) Model of how GST-NR and G can interact without forming any large aggregates.

Figure 8 shows the result of one such IX column, DEAE-Sepharose. Lanes 6-10 show GST-NR and G eluting in the same fractions through the applied NaCl gradient, though G was not expected to interact with the column. The two polypeptides were confirmed by mass spectrometry using a concentrated fraction from this analysis (Figure 8, Lane 16). This observation introduced an interesting theory that intra- or intermolecular interactions may form between the two polypeptides and the existence of such an interaction became the hypothesis of this thesis work.

1.6 Intrinsically Disordered Protein Regions

Intrinsically disordered protein regions (IDPRs) do not form rigid secondary protein structures, but rather form transient structures dictated by nearby binding partners or environmental factors (22, 23). IDPRs in some proteins have autoinhibitory functions through intramolecular interaction with a functional domain of the same polypeptide and inhibit functions such as subcellular localization, ligand interaction and enzymatic function (92, 93). Relieving inhibition may involve post-translational modification (commonly phosphorylation), proteolysis or ligand binding (92, 93). Known autoinhibited proteins include CaMKI, Src and moesin (93). For example, the C-terminal domain of moesin is an IDPR which binds intramolecularly to its FERM domain, inhibiting interactions with binding partners of moesin (93, 94). While bound, the C-terminal domain adopts a helical secondary structure. After phosphorylation of a threonine residue within this domain, it is thought that the

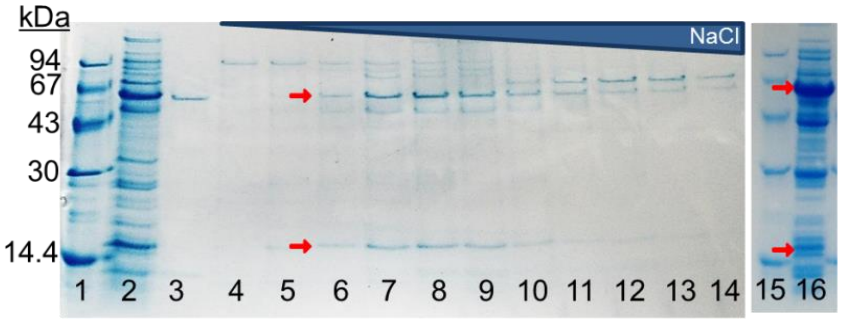


Figure 8. Coelution of the GST-NR (58 kDa) and G (16 kDa) polypeptides, indicated by the two red arrows. Lane 2 shows the sample loaded on the column. Lane 3 is the flow-through proteins. Lanes 4-14 show the eluted proteins through the NaCl gradient. Lane 16 shows the concentrated eluted protein run on a separate gel, the bands indicated by arrows were analyzed by mass spectrometry. The upper band was GST-NR and the lower band G.

helical structure is destabilized, causing a loss of binding and therefore inhibition (93, 94). We postulate that a similar interaction may exist in TDP-43.

Some models for intramolecular interaction of TDP-43 domains are shown in Figure 9. The C-terminal domain may interact with either the N-terminal or RNA-binding domain of the same TDP-43 molecule, a region of another TDP-43 molecule or involve a mediator protein. There is no current evidence for autoinhibition by the C-terminal domain, but there are regions of the C-terminal domain that participate in protein interactions and aggregation. As previously discussed, the C-terminal domain also contains at least 2 serine residues, S409 and S410, that are confirmed to be phosphorylated in many patients with neurodegenerative disease (21). These phosphorylation events could be evidence for autoinhibition because as previously mentioned, by means of relieving inhibition is post-translational modification. However, the physiological levels of phosphorylated TDP-43 and its functional role have yet to be determined.

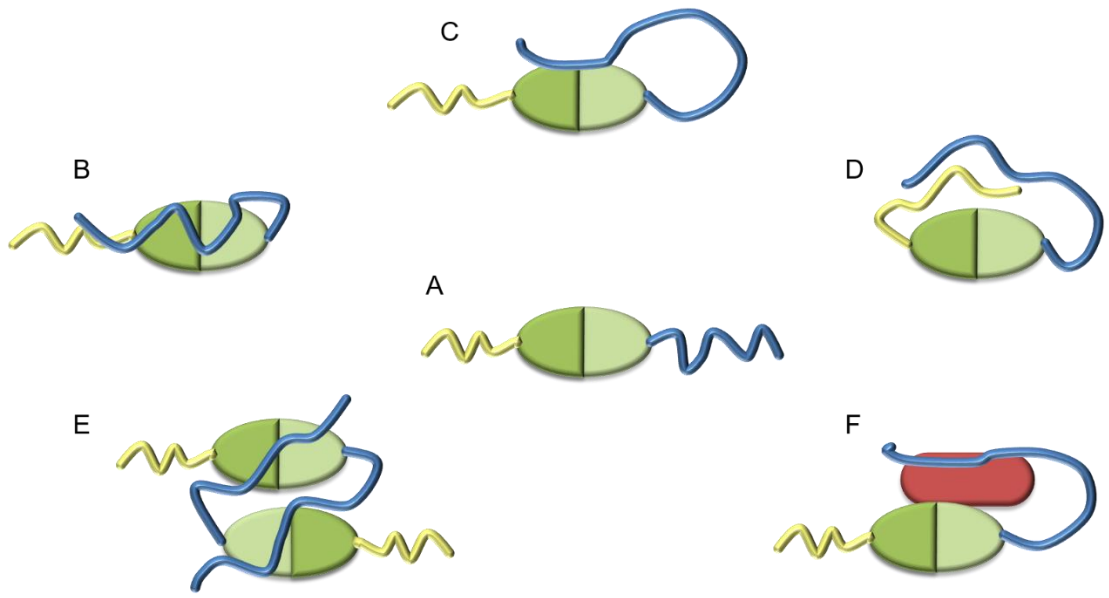


Figure 9. Different proposed models of intra- and intermolecular interactions by the TDP-43 C-terminal domain. A) The non-interacting form of TDP-43. The N-terminal domain is indicated in yellow, the RNA-binding domain is indicated in green and the C-terminal domain in blue. B) The C-terminal domain folds and makes many interactions with the N-terminal domain and RNA-binding domains. C) Only a small region of the C-terminal domain is involved in interactions with the RNA-binding domain. D) Only a small portion of the C-terminal domain is involved in interactions with the N-terminal domain. E) The interaction is intermolecular and involves two TDP-43 molecules. F) A potential mediator protein (indicated by red) serves as an intermediary between the domains of TDP-43.

1.7 Hypothesis

Developing a more thorough understanding of TDP-43 and its normal, biological structure and function is necessary for finding the critical events leading to disease onset. My hypothesis is that the C-terminal domain of TDP-43 is involved in intra- or intermolecular interactions (Figure 10). To test this hypothesis I will create constructs expressing specific regions of TDP-43 and subject them to various biophysical and cell biology techniques to elucidate possible interdomain interactions. The constructs will be studied in isolated conditions to reconstitute interactions and to confirm the important binding region within the C-terminal domain.

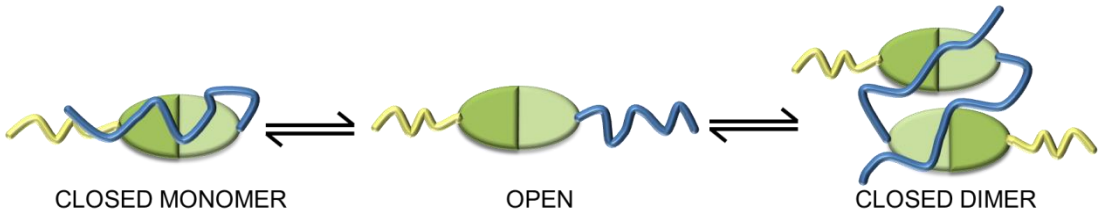


Figure 10. Cartoon diagram of the predicted binding of TDP-43 domains. When TDP-43 is in a monomeric “closed” state, the C-terminal domain (blue) forms an intramolecular interaction with the N-terminal domain (yellow) or RNA-binding domain (green). The dimeric “closed” state is included because its formation is also possible since TDP-43 has been observed as a dimer. In the “open” state, the intramolecular interaction has been relieved.

Chapter 2: Experimental Procedures

2.1 TDP-43 Construct Cloning

The TDP-43 coding construct used in this study was provided by Dr. Kathryn Volkening. This was a neuronal variant of TDP-43, cloned from human brain which encodes a 416 amino acid TDP-43 variant. Its isolation and expression was reported by Strong *et al* (2007) (18).

The proteins used in this study were all cloned into the parent plasmid, pEBT7, which contains a T7 promoter, ampicillin resistance gene and an N-terminal TEV cleavable His_6-Thioredoxin- (HT-) tag (95). Any plasmids were isolated by alkaline lysis or QIAprep Miniprep extraction from MM294 or BL21DE3 *E. coli* cells. After restriction endonuclease (RE) digestion, all DNA fragments were run on 0.8% agarose gels and purified by excision from the gel and repeated freezing at -20°C and thawing. The final thawed sample was centrifuged in a desktop centrifuge through fibre floss to extract the DNA. The plasmids pCB1, pSD600, pSD639, pSD747, pSD748 and pSD749 were made by Yumin Bi and Dr. Stan Dunn and were used throughout this work.

The plasmid pKD006 was made by ligating the 1257-bp BamHI/HindIII fragment from pCB1 with the 1320-bp HindIII/PstI and 4495-bp PstI/BamHI. fragments from pEBT7. pKD006 expressed full- length TDP-43 with an N-terminal HT tag. The plasmid pKD004 was made by ligating the 790-bp BamHI/HindIII fragment from pSD600 with the 1320-bp HindIII/PstI and 4495-bp PstI/BamHI fragments from pEBT7. pKD004 expressed the N-terminal and RNA-binding domains with an N-terminal HT tag.

Quickchange mutagenesis by PCR was used to create a number of mutations to pSD639, which expressed the C-terminal domain with an N-terminal HT tag. The reactions with the forward and reverse primers were carried out in separate tubes and utilized the KOD Hot Start DNA polymerase (EMD Millipore Corp). The thermocycler protocol started at 94°C for 5 min, then cycled [94°C for 40 s, 55°C for 40 s, 72°C for 5 min] 30 times. The forward and reverse reactions were combined and returned to the thermocycler for annealing—95°C for 5 min, 90°C for 1 min, 80°C for 1 min, 70°C for 30 s, 60°C for 30 s, 50°C for 30 s, 40°C for 30 s and 37°C hold. The DNA was then cleaned up with a PCR cleanup kit (BioBasic Canada Inc., Markham ON) and DpnI-digested to degrade the template DNA. Table 1 outlines the plasmids made by this method.

Further cloning by Quickchange mutagenesis was performed to construct a set of plasmids that express truncated versions of the C-terminal domain variants. Forward and reverse primers were designed to insert a second BamHI site within the C-terminal domain gene sequence. PCR reactions were run and digested as previously described. Purified plasmids were then digested with BamHI and the 5917-bp or 5878-bp fragment was ligated by itself, creating the truncated C-terminal fragment. Table 2 outlines the plasmids made by this method.

Plasmid	Name	Primers (5’ to 3’)
pKD011	HT-G-M337V	F:CAGTTGGGGTATGGTGGGCATGTTAGC R:GCTAACATGCCCACCATACCCCAACTG
pKD012	HT-G-Δ407	F:CTCAAGCATGGATTAAGTCTTCTGGCTG R:CAGCCAGAAGACTTAATCCATGCTTGAG
pKD013	HT-G-S409D	F:CTCAAGCATGGATTCTAAGGAATCTGGCTGGGGAATGAATC R:GATTCATTCCCCAGCCAGATTCCTTAGAATCCATGCTTGAG
pKD015	HT-G-S410D	F:GCATGGATTCTAAGTCTGAAGGCTGGGGAATGAATCAC R:GTGATTCATTCCCCAGCCTTCAGACTTAGAATCCATGC
pKD017	HT-G-SS409DD	F:GTTAGTGATTCATTCCCCAGCCkTCkTCCTTAGAATCCATGCTTGAGCCAAAGCCTCCATTAAAAC R:GTTTTAATGGAGGCTTTGGCTCAAGCATGGATTCTAAGGAkGAkGGCTGGGGAATGAATCACTAAC
pKD027	HT-G-ΔNH	F:GGCTGGGGAATGTAACCTAGGAAGC R:GCTTCCTAGGTTACATTCCCCAGCC
pKD028[α]	HT-G-S409D-ΔNH	F:GGCTGGGGAATGTAACCTAGGAAGC R:GCTTCCTAGGTTACATTCCCCAGCC

F-forward primer; R-reverse primer; k-G or C in sequence; [α]-parent plasmid was pKD013

Table 1. List of plasmids and primers constructed by Quickchange mutagenesis.

Plasmid	Name	Parent Plasmid	Primers (5' to 3')
pKD019	HT-$G_{386-416}$	pSD639	F:CAATTGGTTGGGGATCCGCATCCAATGCAGGG R:CCCTGCATTGGATGCGGATCCCCAACCAATTG
pKD020	HT-$G_{399-416}$	pSD639	F:CAGTGGTTTTAATGGATCCTTTGGCTCAAGCATGG R:CCATGCTTGAGCCAAAGGATCCATTAAAACCACTG
pKD021	HT-$G_{386-406}$	pKD012	F:CAATTGGTTGGGGATCCGCATCCAATGCAGGG R:CCCTGCATTGGATGCGGATCCCCAACCAATTG
pKD022	HT-$G_{399-406}$	pKD012	F:CAGTGGTTTTAATGGATCCTTTGGCTCAAGCATGG R:CCATGCTTGAGCCAAAGGATCCATTAAAACCACTG
pKD023	HT-$G_{386-416}$S409D	pKD013	F:CAATTGGTTGGGGATCCGCATCCAATGCAGGG R:CCCTGCATTGGATGCGGATCCCCAACCAATTG
pKD024	HT-$G_{399-416}$S409D	pKD013	F:CAGTGGTTTTAATGGATCCTTTGGCTCAAGCATGG R:CCATGCTTGAGCCAAAGGATCCATTAAAACCACTG
pKD025	HT-$G_{386-416}$DD	pKD017	F:CAATTGGTTGGGGATCCGCATCCAATGCAGGG R:CCCTGCATTGGATGCGGATCCCCAACCAATTG
pKD026	HT-$G_{399-416}$DD	pKD017	F:CAGTGGTTTTAATGGATCCTTTGGCTCAAGCATGG R:CCATGCTTGAGCCAAAGGATCCATTAAAACCACTG

F-forward primer; R-reverse primer

Table 2. List of plasmids and primers constructed by Quickchange mutagenesis, based on the parent plasmid, pSD639.

2.2 Protein Expression and Cell Lysis

The relevant expression plasmids for the various TDP-43 constructs were transformed into and expressed from the *E. coli* strains, MM294 and BL21DE3. For a large cell growth, one transformed *E. coli* colony was added to 25 mL of liquid LB^{amp} and incubated for 3 h (37°C, shaking) until a density of 0.1 absorbance at 600 nm (A_{600}) was reached. Ten µL of a 10x dilution of this growth was added to 1 litre $2xYT^{amp}$ media (Yeast Extract Tryptone media) and grown overnight at 30°C, shaking, to an A_{600} of 0.6. Protein expression was induced by adding 120 mg of IPTG and growth continued for 6 h (37°C, shaking). Cells were collected by low speed centrifugation in a Beckman JA-9.1 rotor (6k rpm, 10 min) and resuspended in 50 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8.0, buffer to a 50% suspension and stored at -80°C for later use.

The induced cells were lysed using a French Pressure Cell. Two mL of the 50% induced cell suspension were added to 30 mL of lysing buffer, containing 50 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8.0, plus 1 mM PMSF. The cell lysate was centrifuged at low speed in a Beckman JA-25.5 rotor (10k rpm, 10 min, 4°C), the supernatant was collected and centrifuged again at high speed (38k rpm, 90 min, 4°C) in a Beckman 70.1Ti rotor. If the TDP-43 polypeptide was soluble, the final supernatant was stored at -80°C for future experimentation. Some of the variants expressed as inclusion bodies and were initially stored as pellets.

2.3 Protein Expression by Mini-Induction

For some of the experiments, induced cell extracts were required in small amounts. A transformed colony was added to 5 mL of LB^{amp} and grown in a 37°C water bath, shaking. At an A_{600} of 0.1, the cultures were induced with 0.55 mM IPTG and allowed to grow to stationary phase. The induced cells were centrifuged and resuspended in 1x SDS-PAGE sample buffer with 50 mM DTT and boiled for 5 minutes. These extracts were frozen at -20°C for later use.

2.4 TDP-43 Construct Purification

TDP-43 is a highly aggregation-prone protein and the purification of the overexpressed constructs for this work was complicated. The following are the different procedures developed to prepare the TDP-43 variants that were purified for this work. Proteins were analyzed by SDS-PAGE using 15% polyacrylamide gels and a Tris-Tricine running buffer (96). SDS protein samples were prepared with denaturing sample buffer with 50 mM DTT (SB+). The protein bands were visualized by staining the gel with Coomassie Blue G. The protein concentration at various steps was determined by the Bio-Rad Protein Assay (reagent from Bio-Rad Laboratories) by comparing sample absorbance at 595 nm to a calibration curve of known BSA concentrations.

2.4.1 HT-NR purification—This protein was soluble after initial centrifugation. The first step in purification was a 40% saturated ammonium sulfate (NH_4SO_4) precipitation, achieved by adding 100% saturated NH_4SO_4 directly to the retrieved

supernatant. At 40% saturation, the majority of the HT-NR polypeptide precipitated and after centrifugation at 4°C in a Beckman JA-25.5 rotor, the NH_4SO_4 supernatant was removed and the pellet was dissolved in 50 mM Tris-HCl, 100 mM NaCl, 10 mM imidazole, pH 8.0. This buffer was selected because it is compatible with an immobilized metal affinity column. Before loading the resolubilized protein on the column, it was centrifuged at 13,000 rpm for 20 min in a desktop centrifuge to remove any insoluble particles.

A column containing 10 mL of Chelating Sepharose was charged with Ni^{2+} (Ni^{2+} affinity column) to utilize the His_6-tag on HT-NR. At room temperature, the HT-NR solution was loaded onto the column and the flowthrough (FT) was collected. Since the HT-NR polypeptide contains the RNA-binding domain, the column was washed with 10 column volumes of 200 mM $NaNO_3$ and 1 M NaCl to dissociate any bound nucleotides. To elute the bound protein, an elution buffer containing 50 mM Tris-HCl, 100 mM NaCl and 500 mM imidazole, pH 8.0, was added in 5-mL aliquots, with 5-mL fractions collected. The fractions containing the highest levels of HT-NR were identified by SDS-PAGE, pooled and adjusted again to 40% saturated NH_4SO_4 to precipitate the protein as before. After centrifugation, the resulting pellet was collected and redissolved in 2 mL of 750 mM arginine hydrochloride (ArgHCl), 50 mM Tris-HCl, 1 mM DTT, pH 8.0 to concentrate the HT-NR and prepare it for further purification. ArgHCl was added to disrupt any unwanted protein interactions that were maintained through the first column and separate the contaminating species through a second purification column, gel filtration.

A 200 mL column containing Sephacryl S-200 resin (S200) was equilibrated with the 750 mM ArgHCl, 50 mM Tris-HCl, 1 mM DTT, pH 8.0, buffer and the protein sample was loaded on the column and maintained at 4°C. The same buffer was used to run the column, collecting 1-mL fractions at 2 mL/h. As before, the fractions with the highest level of protein were identified by SDS-PAGE, pooled and precipitated with 40% saturated NH_4SO_4. The precipitate was centrifuged as before and redissolved in 50 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8.0, buffer and centrifuged again in preparation for the final purification column.

A 23 mL Q-Sepharose ion exchange (IX) column connected to an AKTAexplorer HPLC was equilibrated with 50 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8.0, at room temperature. The redissolved HT-NR protein was loaded onto the column and eluted with a 6-column volume 0-1 M NaCl gradient in 50 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8.0. The presence of eluted protein was monitored by reading absorbance at 280 nm (A_{280}) of each fraction, further identified by SDS-PAGE, pooled and dialyzed against 50 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8.0, for storage at -80°C as this final step resulted in pure HT-NR protein.

2.4.2 HT-G purification—The HT-G polypeptide was purified by Dr. Stan Dunn and as undergraduate project work by Christina Chung, Thamiya Vasanthakumar and Jethro Kwong (97, 98). Briefly, overexpression of HT-G in *E. coli* formed an inclusion body that sedimented in the low speed pellet after lysis. The inclusion body was washed and resolubilized in 4 M guanidine hydrochloride (GuHCl), 10

mM Tris-HCl, 0.1 mM EDTA, 10 mM imidazole, pH 8.0 then loaded onto a Ni^{2+} affinity column. To facilitate refolding of the denatured HT-G, the column was washed with 100 mM Tris-HCl, 30% dimethylformamide, 10 mM imidazole, pH 8.0, then with 100 mM Tris-HCl, 10 mM imidazole, pH 8.0. The column was washed again with 750 mM ArgHCl, 100 mM Tris-HCl, 10 mM imidazole, pH 8.0, and finally 750 mM ArgHCl, 100 mM Tris-HCl, 250 mM imidazole, pH 8.0, to elute the bound protein. The eluted protein was pooled and stored at -80°C.

HT-G was dialyzed at room temperature against 1 mM trifluoroacetic acid (TFA) in ddH_2O, then 0.1 mM TFA. Since TFA is acidic, a further dialysis was done with 25 mM Tris-HCl, 0.1 mM TFA and 1 mM TCEP, pH 7.5 or 25 mM Tris-HCl, 25 mM Tricine and 1 mM TCEP, pH 7.3, to neutralize the solution for further work. Dialysis against 25 mM Tris base, 50 mM Tricine, 1 mM TCEP, pH 7.5 also maintained HT-G solubility and was used in some work.

2.4.3 C-terminal domain mutant purification—Two of the deletion mutant polypeptides, HT-$G_{386-416}$ and HT-$G_{386-406}$, were purified. Much of the aggregation prone region of the C-terminal domain was deleted and therefore after cell lysis, the two constructs were soluble after the high speed spin. Both were purified by the same procedure as follows.

After centrifugation, the supernatants were 40% saturated NH_4SO_4 precipitated, centrifuged at 13,000 rpm at 4°C and the pellet was resuspended in 50 mM Tris-HCl, 100 mM NaCl and 10 mM imidazole, pH 8.0, for loading on a Ni^{2+} affinity column. The bound protein was washed with 5 column volumes of 50 mM

Tris-HCl, 100 mM NaCl, 10 mM imidazole, pH 8.0, then eluted as previously described in the HT-NR purification procedure. Following elution from the Ni^{2+} affinity column, the fractions that contained the most protein were pooled and dialyzed against 50 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8.0. Next, the samples were loaded onto the IX Q-column, as previously described. Following elution, the fractions containing the polypeptides of interest were pooled, 70% saturated NH_4SO_4 precipitated and the pellets were redissolved in 25 mM Tris base, 50 mM tricine, 0.5 mM TCEP, pH 7.5. The two HT-G variants were purified at this point and were stored for future analysis.

2.4.4 HT tag purification—A volume of cleaved and Ni^{2+} affinity column purified HT tag was attained from Jethro Kwong, a member of the Dunn Lab. This was dialyzed against 50 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8.0, and loaded onto the IX Q-column, as described in section 2.4.1. Following elution, the fractions containing HT were pooled, 70% saturated NH_4SO_4 precipitated and the pellet was redissolved in 25 mM Tris base, 50 mM tricine, 0.5 mM TCEP, pH 7.5, then stored for later use.

2.5 Western Blot

Through many stages of this work, Western blots were used to verify the presence of TDP-43 polypeptides. The standard methods for Western blotting were conducted as previously described (99).

The samples were run by SDS-PAGE long enough for the dye front to run off the 15% polyacrylamide gel. The proteins were then electrophoretically transferred to an Immobilon polyvinylidene fluoride (PVDF) membrane at 0.25 A for 1 h in a carbonate blot buffer (99). The blotted membrane was transferred to a blocking buffer containing 3% BSA in blot rinse buffer (10 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 0.1% Tween 20 and 0.04% sodium azide, pH 7.4) and incubated for at least 2 h at room temperature while shaking. The membrane was then rinsed 3 times in plain blot rinse buffer for 10 min. Next, the buffer was exchanged for blot rinse buffer containing 0.3% BSA, the primary antibody was added at a dilution of 1:5,000 and the membrane was incubated for 2h at 23°C or overnight at 4°C. Two polyclonal rabbit antibodies from Sigma-Aldrich were used to detect TDP-43. The first was an anti-N-terminal antibody, with an immunogen of a synthetic peptide composed of residues 75-90 of TDP-43 (Product Number: SAB4200006). The second was an anti-C-terminal antibody, with an immunogen of a synthetic polypeptide composed of residues 355-369 of TDP-43 (Product Number: T1580). After a 2 h incubation at room temperature, the membrane was washed again with rinse buffer, exchanged with 0.3% BSA buffer and the secondary antibody (alkaline phosphatase conjugated goat anti-rabbit IgG, Jackson ImmunoResearch Laboratories) was added at a dilution of 1:5,000. The membrane was incubated for 1h at 23°C then washed again with rinse buffer. Briefly, the alkaline phosphatase was reacted by incubating the membrane in 10 mL of 100 mM Tris-HCl, 100 mM NaCl, 5 mM $MgCl_2$, pH 9.5, with 80 µL of BCIP solution (20 mg/mL in 100% dimethylformamide) and 60 µL NBT solution (50

mg/mL in 70% dimethylformamide). The membrane was left to react until the desired level of detection of the alkaline phosphatase was reached, approximately 10 minutes.

2.6 Far Western Blot

In this method, the proteins to be blotted were expressed in 5-mL liquid cultures of BL21 *E.coli*. After induction by IPTG, samples were created by centrifuging a volume of the liquid cell culture and resuspending the cells in 1xSB+ and placing in a boiling water bath for 3-5 min. The cell extracts were analyzed by SDS-PAGE to begin the process of equalizing the loading of the desired proteins through subsequent dilutions and SDS-PAGE. Once the desired loading amounts were detected, the samples were stored at -20°C for use in this analysis.

Probe proteins were labelled with ^{125}I by Dr. Stan Dunn using the Iodogen procedure (100).

The blots were prepared as previously described using Immobilon membrane. Following blocking and rinsing, the membranes were transferred into 0.3% BSA blot rinse buffer and the radioactive probe was added at 50,000-100,000 dpm/mL. After an overnight incubation at room temperature, the blots were washed for 10 min to remove any unbound probe and air dried. Finally, the blots were visualized with a Storm 820 phosphorimager (Amersham Biosciences).

2.7 Sedimentation Equilibrium and Velocity

The sedimentation experiments were conducted in a Beckman XLA Analytical Ultracentrifuge in an An60 Ti rotor.

For sedimentation equilibrium (SE), samples were loaded in a 6-channel centerpiece and maintained at 4°C. The loaded protein concentrations had absorbances between 0.2-0.5 at A_{280} and were sedimented at 12,000-24,000 rpm, depending on the experiment. The time allotted for sedimentation equilibration after acceleration also depended on the experiment but all scans had radial intervals of 0.003 mm and 10 replicates.

The SE data was fit to a *single species* model using GraphPad Prism 5. The distribution is described by the equation:

$$C = C_o \cdot \exp\left[\frac{\omega^2}{2RT} \cdot MW_{exp}(1 - \bar{\upsilon}\rho) \cdot (x^2 - x_o^2)\right] + I_o \qquad \text{Equation 1}$$

where C is the concentration at radius x, with a concentration C_o at reference radius x_o, ω is the angular velocity of the rotor, the calculated molecular weight of the protein is MW_{exp}, $\bar{\upsilon}$ is the partial specific volume of the analyzed protein, ρ is the solvent density, R is the gas constant (8.314×10^7 $\text{erg} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$), T is temperature in Kelvin and I_o is the baseline correction.

For sedimentation velocity (SV), samples were loaded in a 2-channel centerpiece and maintained at 4°C. The loaded protein concentrations had A_{280}=0.2-0.5 and were sedimented at 45,000 rpm. Scans were taken every 10 minutes for a total of 30 scans, at a radial interval of 0.003 mm with 5 replicates. The sedimentation spectra were analyzed by the program Sedfit (downloaded from www.analyticalultracentrifugation.com) (101).

2.8 Transfections and Colocalization

To complement the *in vitro* studies, *in vivo* colocalization studies with the same TDP-43 domains were carried out by Cheryl Leystra-Lantz and Dr. Cristian Droppelmann in the Strong Lab. Briefly, four plasmids expressing fluorescently labelled NR and G polypeptides were constructed: GFP-NR, mCherry-NR, GFP-G and mCherry-G. The plasmids were isolated by miniprep (Invitrogen) for transfection into HEK293T cells.

To visualize *in vivo* binding of NR and G by confocal microscopy, cotransfections with GFP-NR/mCherry-G or mCherry-NR/GFP-G were performed. The cells were plated at $1.25x10^5$ cells/dish in 35 mm confocal dishes (MatTek). One µg of each plasmid was incubated with Lipofectamine 2000 at a 1:3 ratio and transfection was carried out as per the manufacturer's protocol. The cells were allowed to grow for 48 h then fixed with 4% paraformaldehyde paraformaldehyde in PBS for 10 minutes, washed in PBS and then visualized using the Meta510 Confocal Microscope (Robarts Confocal Imaging Facility) at 600x magnification.

Chapter 3: Results

3.1 TDP-43 Variant Plasmid Cloning

To generate materials for studying interdomain interactions in TDP-43, a plasmid library was created for this work that expressed variants of TDP-43 including individual domains, truncations or point mutations, as described in Chapter 2. The expressed TDP-43 has 416 amino acids and was isolated from human brain (36). All variants were N-terminally tagged with a His_6-Thioredoxin (HT) tag and were induced in the *E. coli* strain BL21DE3. Figure 11 schematically shows the library of expressed polypeptides. Throughout this work, the proteins will be referred to as their descriptive "Name" and not the plasmid number. The proteins HT-NRG, HT-NR, HT-RG, HT-N, HT-R and HT-G all express various combinations of the wild-type TDP-43 domains, as indicated by their names.

Many variants of the C-terminal domain were also created to see how mutations or phosphorylation affected possible interdomain interactions. HT-G-M337V expressed the C-terminal domain with the common familial ALS mutation, Met337Val.

The HT-G-Δ407 construct is a truncated C-terminal domain, missing residues 407-416. This construct was created to eliminate the extreme C-terminus, including the key Ser409/410 residues that were phosphorylated in aggregates.

The constructs HT-G-S409D, HT-G-S410D and HT-G-SS409DD express mutations to the C-terminal domain that introduce an Asp (D) residue at the

phosphorylatable Ser409/410. These mutations introduce a phosphomimetic residue that mimic the effects of phosphorylation on interaction.

The remaining constructs were made to introduce further alterations to the previously mentioned C-terminal domain proteins. Two truncations were made that deleted much of the C-terminal domain, expressing only residues 386-416 or 399-416. Similar deletions were also made to some of the previously mentioned mutants. Additionally, the last two residues, Asn415 and His416, were deleted from two of the C-terminal domain constructs.

To verify that the new TDP-43 constructs contained the expected domains of TDP-43 and would blot sufficiently, SDS cell extracts of most of the constructs were blotted and probed with antibodies for the N-terminus and C-terminus of TDP-43. The HT-G constructs pKD019-026 were not blotted because the immunogen site was within the deleted residues. Figure 12 shows the developed blots. All polypeptides were detectable by the appropriate antibody. HT-R was not detected by either antibody as expected. There was some background signal with the antibodies but this was likely due to sample loading, degradation products or non-specific binding by either the primary or secondary antibody.

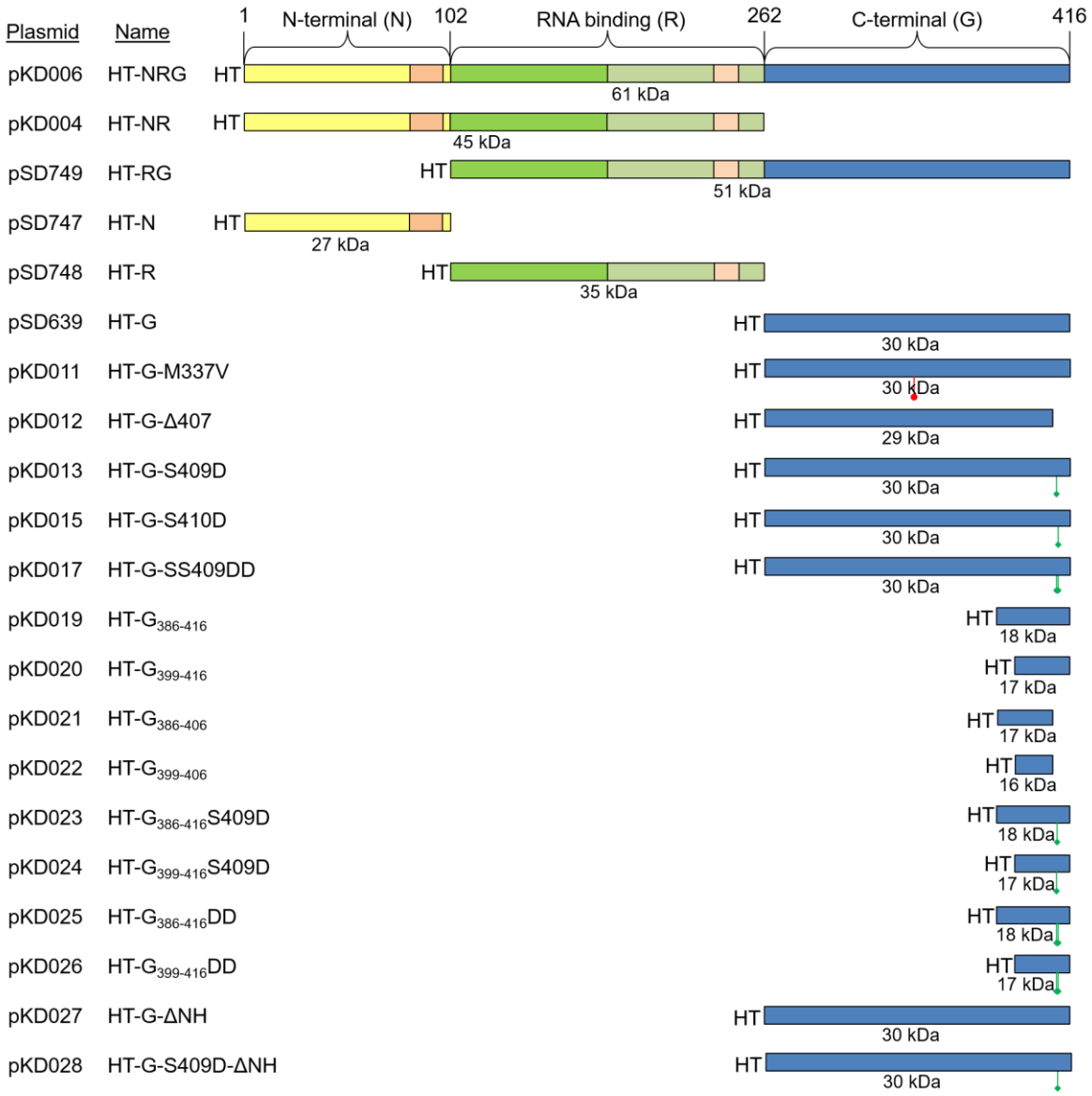


Figure 11. Schematic diagram of the TDP-43 variants used in this work. The calculated molecular weight for each expressed protein is below each diagram. The red marker shows the location of an ALS associated mutation. The green marker shows the location of a phosphomimetic mutation. Unless indicated by subscript position numbers, HT-G constructs contain the full C-terminal domain, modified with the further descriptors.

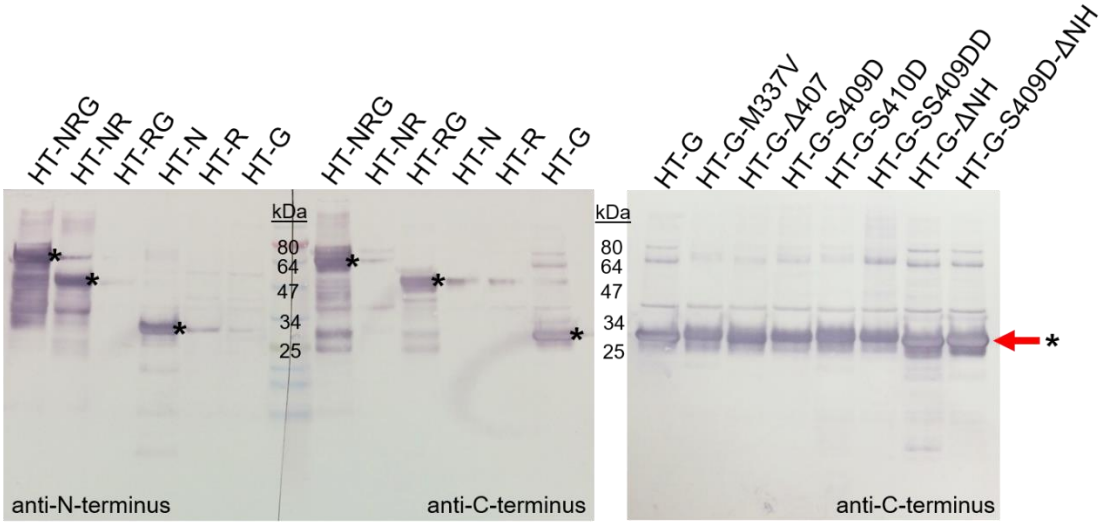


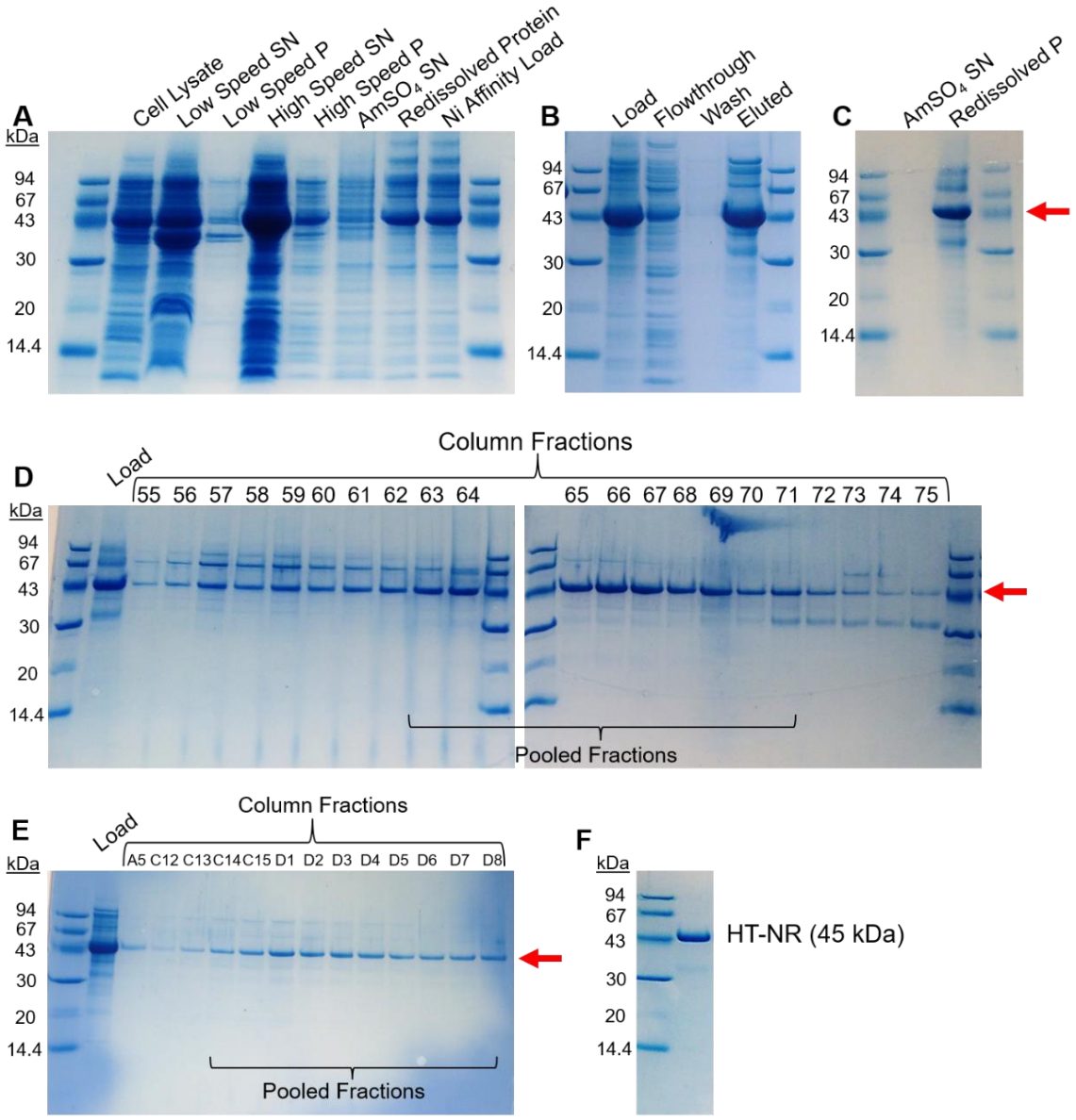
Figure 12. TDP-43 constructs detected by Western blot. SDS cell extracts were prepared, blotted and probed with antibodies for the N-terminus and C-terminus of TDP-43. The blotted constructs are labelled above. The expected bands are starred (*****). Any other present bands are likely due to non-specific binding by the antibodies used, oligomeric species of the constructs or degradation species of the constructs.

3.2 Polypeptide Purification

Pure proteins were required for some of the experiments in this work and were purified as outlined in Chapter 2.

The first, HT-NR, was purified through a series of precipitations and columns, as demonstrated in Figure 13. After lysis and centrifugation, the majority of the expressed HT-NR was soluble (Figure 13A). It was precipitated by 40% saturated NH_4SO_4 and redissolved in 50 mM Tris-HCl, 100 mM NaCl and 10 mM imidazole, pH 8.0, buffer for loading on the Ni^{2+} affinity column (Figure 13B). At this point, there was a possibility that RNA or DNA might still be bound to the RNA-binding domain. Therefore, the bound HT-NR was washed with 1 M NaCl and 200 mM $NaNO_3$, a chaotropic agent, in 50 mM Tris-HCl, 100 mM NaCl and 10 mM imidazole, pH 8.0, to disrupt the nucleotide interactions. RNase and DNase were digests were attempted, however there was some interaction with the column and elution of the nucleases overlapped with the eluted HT-NR (data not shown). The bound protein was eluted with 50 mM Tris-HCl, 100 mM NaCl and 500 mM imidazole, pH 8.0, and the fractions with HT-NR were pooled and precipitated by 40% saturated NH_4SO_4 for transition into 750 mM ArgHCl, 50 mM Tris-HCl, 1 mM DTT, pH 8.0 (Figure 13C). ArgHCl was used to solubilize the more concentrated HT-NR without structure denaturation and to assist in breaking up any protein-protein interactions that survived the previous column. The HT-NR was purified further by running through a gel filtration column (S200), equilibrated with 750 mM ArgHCl, 50 mM Tris-HCl, 1 mM DTT, pH 8.0 (Figure 13D). In consecutive fractions, two peaks of HT-NR were observed: the first, centred around fraction 59,

Figure 13. Purification of the HT-NR polypeptide. The red arrow indicates the band of interest, HT-NR. A) SDS samples of the cell lysate, centrifugation supernatants (SN) and pellets (P) and initial 40% saturated NH_4SO_4 precipitation. The low speed and high speed pellets were resuspended in an equal volume to the supernatant before sampling. B) Sample of HT-NR loaded onto the Ni2+ affinity column and samples from the FT, wash and eluted protein. C) SDS samples of the 40% saturated NH_4SO_4 SN and redissolved P of the pooled HT-NR eluted from the Ni2+ affinity column. D) SDS samples of the loaded HT-NR and eluted fractions from the S200 gel filtration column. The pooled fractions are indicated by the bracket. E) SDS samples of the eluted HT-NR from the IX column. The pooled fractions are indicated by the bracket. F) An example of the final purified HT-NR for use in further study.


A
kDa
94
67
43
30
20
14.4
Cell Lysate
Low Speed SN
Low Speed P
High Speed SN
High Speed P
AmSO4 SN
Redissolved Protein
Ni Affinity Load
B
Load
Flowthrough
Wash
Eluted
C
AmSO4 SN
Redissolved P
D
Load
Column Fractions
55 56 57 58 59 60 61 62 63 64
65 66 67 68 69 70 71 72 73 74 75
Pooled Fractions
E
Load
Column Fractions
A5 C12 C13 C14 C15 D1 D2 D3 D4 D5 D6 D7 D8
Pooled Fractions
F
HT-NR (45 kDa)

coeluted with an unknown *E. coli* protein and the second, centred around fraction 66, contained predominantly HT-NR. This second peak was pooled, precipitated by 40% saturated NH_4SO_4, then redissolved and dialyzed in 50 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8.0, for final purification by IX (Figure 13E). HT-NR eluted between 250-350 mM NaCl and the most concentrated samples were pooled and stored for later use. Figure 13F shows the final purified HT-NR protein.

As previously mentioned, the HT-G purification protocol was established by Dr. Stan Dunn, Christina Chung, Thamiya Vasanthakumar and Jethro Kwong. Many different salts and buffering conditions were unsuccessful at preventing precipitation of HT-G. The established procedure that successfully achieved soluble and purified HT-G was in a buffer containing 750 mM ArgHCl. This construct contains the aggregation prone C-terminal domain of TDP-43 and ArgHCl was used to solubilize HT-G (102). The high concentration of ArgHCl was not an ideal buffer condition for interaction studies, so it was necessary to find another buffer condition. TFA is a known additive for dissolving and solubilizing aggregation prone proteins, including β-amyloid (103). HT-G remained soluble after room temperature dialysis into 1 mM TFA, then 0.1 mm TFA in ddH_2O (Figure 14A). To increase the pH, the TFA solubilized HT-G was buffered with Tris-HCl, pH 8.0, at various concentrations to determine solubility (Figure 14B). Adding 25 mM Tris-HCl, pH 8.0, maintained maximum solubility of HT-G and 1 mM of neutralized TCEP was added (Figure 14C). For some experiments, a buffer with a higher ionic strength was suitable since many proteins require such conditions to adopt proper structures and interactions. Tricine is similar to Tris in structure

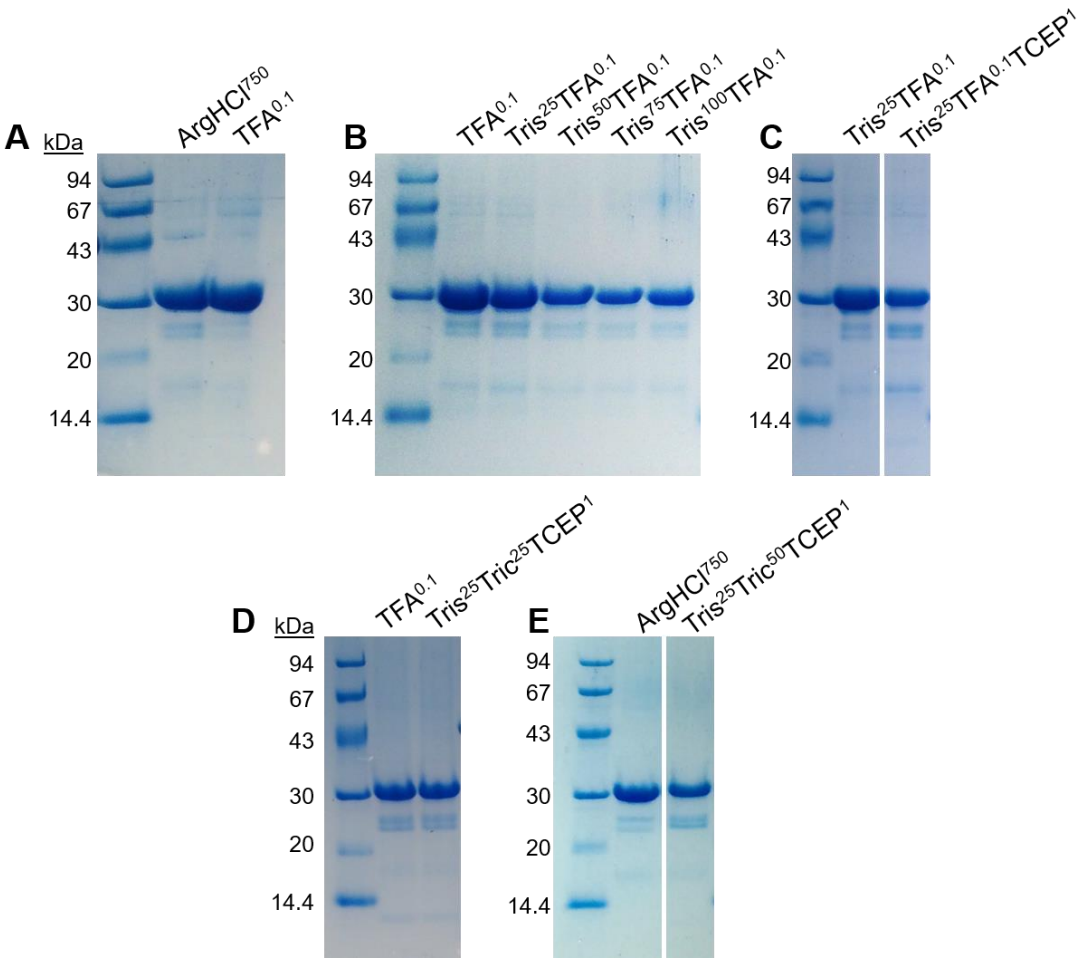


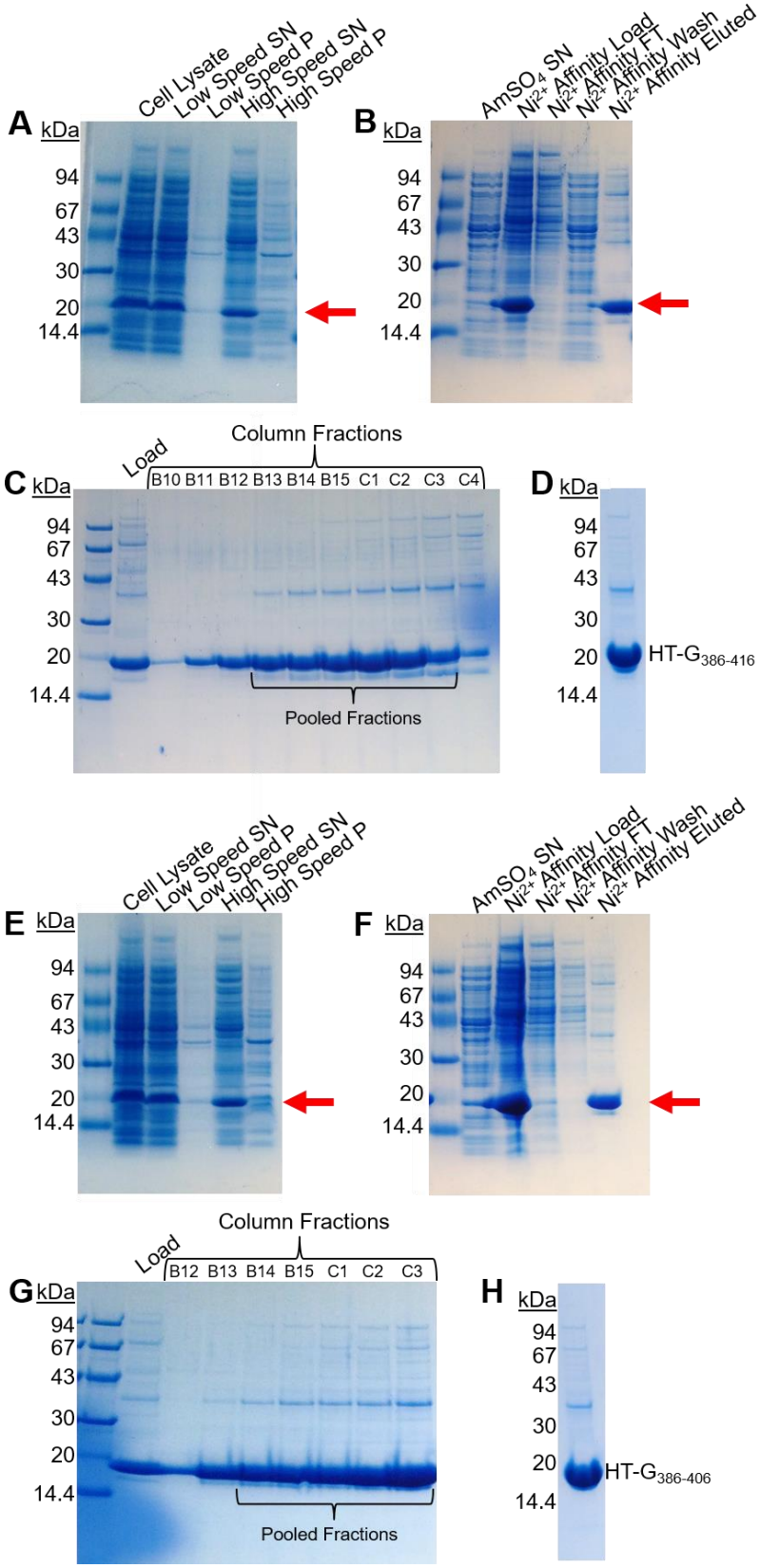
Figure 14. Solubility trials of HT-G. A) SDS-samples demonstrating HT-G in solution before (Arg^{750}) and after dialysis into 0.1 mM TFA and centrifugation at 13,000 rpm for 15 min ($TFA^{0.1}$). B) SDS samples of the tests for HT-G solubility after dialysis into 25, 50, 75 and 100 mM Tris-HCl, pH 8.0 ($Tris^{25\text{-}100}$). Each solution was centrifuged at 13,000 rpm for 15 min after dialysis. C) SDS samples demonstrating the solubility of HT-G after addition of 1 mM neutralized TCEP ($Tris^{25}TCEP^{1}$). D) SDS samples demonstrating that HT-G in 0.1 mM TFA remained soluble after dialysis against 25 mM Tris-HCl, 25 mM tricine, 1 mM TCEP, pH 7.3, and centrifugation at 13,000 rpm for 15 min ($Tris^{25}Tric^{25}TCEP^{1}$). E) SDS samples showing that HT-G in 750 mM ArgHCl was soluble after dialysis against 25 mM Tris base, 50 mM tricine, 1 mM TCEP, pH 7.5, and centrifugation at 13,000 rpm for 15 min ($Tris^{25}Tric^{50}TCEP^{1}$).

but it is an acid rather than a base and has a more spread out charge distribution. HT-G in 0.1 mM TFA was dialyzed overnight into 25 mM Tris-HCl, 25 mM tricine, 1 mM TCEP, pH 7.3, and remained soluble (Figure 14D). TFA was not included in this dialysis because it is believed to bind directly to the protein and is not easily removed. In an additional sample preparation, ArgHCl solubilized HT-G was dialyzed overnight against 25 mM Tris base, 50 mM tricine, 1 mM TCEP, pH 7.5, without any intermediate TFA dialysis. HT-G was soluble under this condition and used in further studies (Figure 14E).

The two polypeptides HT-$G_{386\text{-}416}$ and H-$G_{386\text{-}406}$ were purified through a series of precipitations and columns, as demonstrated in Figure 15A-D and Figure 15E-H, respectively. Both were purified by the same method. After lysis and centrifugation, the majority of the expressed polypeptides were soluble (Figure 15A and E). The high speed supernatants were precipitated by 40% saturated NH_4SO_4 and redissolved in 50 mM Tris-HCl, 100 mM NaCl and 10 mM imidazole, pH 8.0, buffer for loading on the Ni^{2+} affinity column (Figure 15B and F). Samples of the FT, wash and eluted protein are shown in Figure 15B and F. The fractions that contained the highest amount of eluted protein were dialyzed against 50 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8.0, for final purification by IX (Figure 15C and G). Both HT-G mutants eluted between 230-360 mM NaCl and the most concentrated samples were pooled and stored for later use. Figure 15D and H show the final purified HT-$G_{386\text{-}416}$ and HT-$G_{386\text{-}406}$, respectively.

Finally, a purification of HT was performed to have a negative control for any influence of the tag during interaction experiments. The sample acquired from

Figure 15. Purification of HT-$G_{386-416}$ and HT-$G_{386-406}$. A-D are for HT-$G_{386-416}$ and E-H are for HT-$G_{386-406}$. The red arrow indicates the band of interest. A and D) Cell lysate, centrifugation supernatants (SN) and pellets (P). The low speed and high speed pellets were resuspended in an equal volume to the supernatant before sampling. B and F) SDS samples of the NH_4SO_4 SN and pellet (Ni^{2+} affinity load) and samples from the FT, wash and eluted protein from the Ni^{2+} affinity column. C and G) SDS samples of the eluted HT-$G_{386-416}$ and HT-$G_{386-406}$ from the IX column. The pooled fractions are indicated by the bracket. D and H) An example of the final purified HT-$G_{386-416}$ and HT-$G_{386-406}$ for use in further study.


A
kDa
94
67
43
30
20
14.4
Cell Lysate
Low Speed SN
Low Speed P
High Speed SN
High Speed P
B
kDa
94
67
43
30
20
14.4
AmSO4 SN
Ni2+ Affinity Load
Ni2+ Affinity FT
Ni2+ Affinity Wash
Ni2+ Affinity Eluted
C
kDa
94
67
43
30
20
14.4
Load
Column Fractions
B10 B11 B12 B13 B14 B15 C1 C2 C3 C4
Pooled Fractions
D
kDa
94
67
43
30
20
14.4
HT-G386-416
E
kDa
94
67
43
30
20
14.4
Cell Lysate
Low Speed SN
Low Speed P
High Speed SN
High Speed P
F
kDa
94
67
43
30
20
14.4
AmSO4 SN
Ni2+ Affinity Load
Ni2+ Affinity FT
Ni2+ Affinity Wash
Ni2+ Affinity Eluted
G
kDa
94
67
43
30
20
14.4
Load
Column Fractions
B12 B13 B14 B15 C1 C2 C3
Pooled Fractions
H
kDa
94
67
43
30
20
14.4
HT-G386-406

Jethro Kwong was dialyzed against 50 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8.0, and loaded onto the IX column (Figure 16A). HT mutants eluted between 240-300 mM NaCl and the most concentrated samples were pooled, 70% saturated NH_4SO_4 precipitated and dialyzed against 25 mM Tris base, 50 mM tricine, 0.5 mM TCEP, pH 7.5. Figure 16B shows the final purified HT.

3.3 Far Western Blot

3.3.1 Interactions with TDP-43 domains—The Far Western blot technique has been widely used to visualize protein-protein interactions (104). To visualize the interactions, SDS-extracts of *E. coli* with the induced TDP-43 variants were separated by SDS-PAGE, blotted and incubated with a radio-labelled probe that was a suspected interactor. Purified HT-NR and HT-N were both labelled with ^{125}I, referred to as ^{125}I-HT-NR and ^{125}I-HT-N. HT-N was purified by Jethro Kwong. A *Rhodobacter capsulatus* F_1-ATPase delta subunit, also fused to HT (HT-δ), is included in this experiment as a negative control. The background *E. coli* proteins in the extracts acted as further controls for off-target binding by the probe. The first blots were to test the interaction capability of ^{125}I-HT-NR with the panel of wildtype TDP-43 domain variants. Figure 17B shows the exposed blot and the corresponding Coomassie stained gel. ^{125}I-HT-NR was observed to interact strongly with HT-NRG, HT-RG, HT-R and HT-G. A weaker interaction was observed with HT-NR and almost no interaction was observed with HT-N. There was no interaction with HT-δ. Following this result, the interactions of ^{125}I-HT-NR and ^{125}I-HT-N were compared (Figure 17C). The interactions observed with the

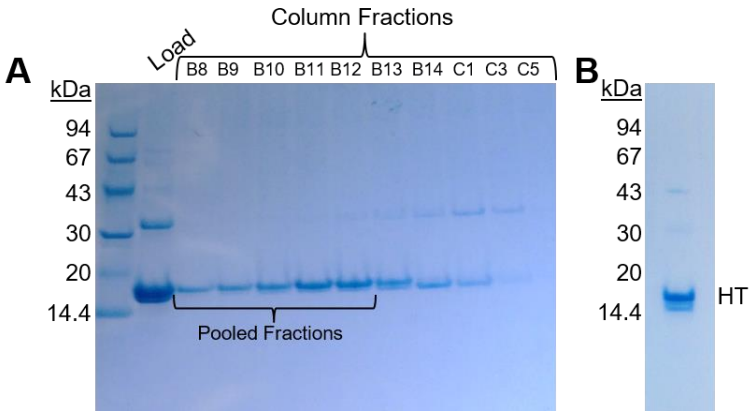


Figure 16. Purification of the HT polypeptide. A) SDS samples of the eluted HT fractions from the IX column. The pooled fractions are indicated by the bracket. B) Final sample of purified HT.

^{125}I-HT-NR probe were similar to those in Figure 17B. However, much weaker ^{125}I-HT-N interactions were observed, with only faint signals for HT-NRG, HT-RG, HT-R and HT-G. These results indicated an interaction by either the N-terminal or RNA-binding domain with the C-terminal domain of TDP-43. The weaker results with ^{125}I-HT-N suggested a stronger influence by the RNA-binding domain for binding with the C-terminal domain.

3.3.2 Interactions with C-terminal domain mutants—The C-terminal domain mutants were blotted and incubated with the probes to see how mutations or phosphomimetics affected interaction. Figure 18A shows the results for interaction with ^{125}I-HT-NR. Strong interactions were observed with HT-G (Figure 17) and with HT-G-M337V. The truncation in HT-G-Δ407 greatly reduced the interaction. A reduction in interaction was also observed with the phosphomimetic variants. HT-G-S409D had a weaker interaction signal than HT-G-S410D but similar to that of the double phosphomimetic. These results indicate that the extreme C-terminus of the C-terminal domain is critical for interaction. Two sets of repetitive blots were performed to verify these results (Figure 18B and C).

3.3.3 Interactions with C-terminal domain deletions—To further analyze the importance of the extreme C-terminal domain to binding, the constructed deletion mutants were analyzed by Far Western blotting (Figure 19). The HT-G positive control showed the expected binding with ^{125}I-HT-NR. There was no observable binding with any of the deletion mutants included. This was a surprising result

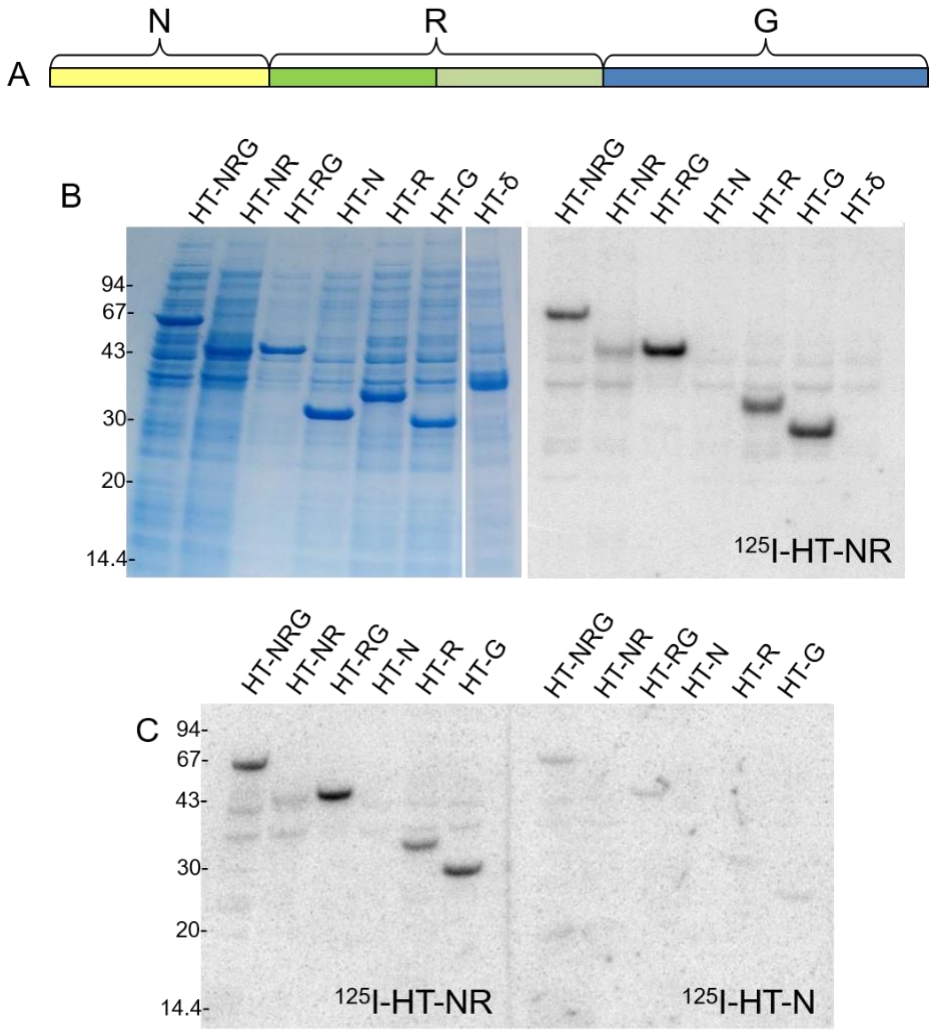


Figure 17. Far Western blot results for the panel of wild-type domains of TDP-43. A) Schematic diagram of the labelling of the TDP-43 domains. B) Shows the Coomassie blue stained gel and exposed blot. Ten microliters of the same samples were used for both the stained and transferred gels. The TDP-43 variants were blotted and incubated overnight with the probe, ^{125}I-HT-NR. The radioactive signal was observed using a phosphorimager. C) The two sets of TDP-43 variants were blotted on the same PVDF membrane, then the membrane was cut in half for separate incubation with either ^{125}I-HT-NR or ^{125}I-HT-N. HT-δ is a negative control. The signals were detected using a phosphorimager.

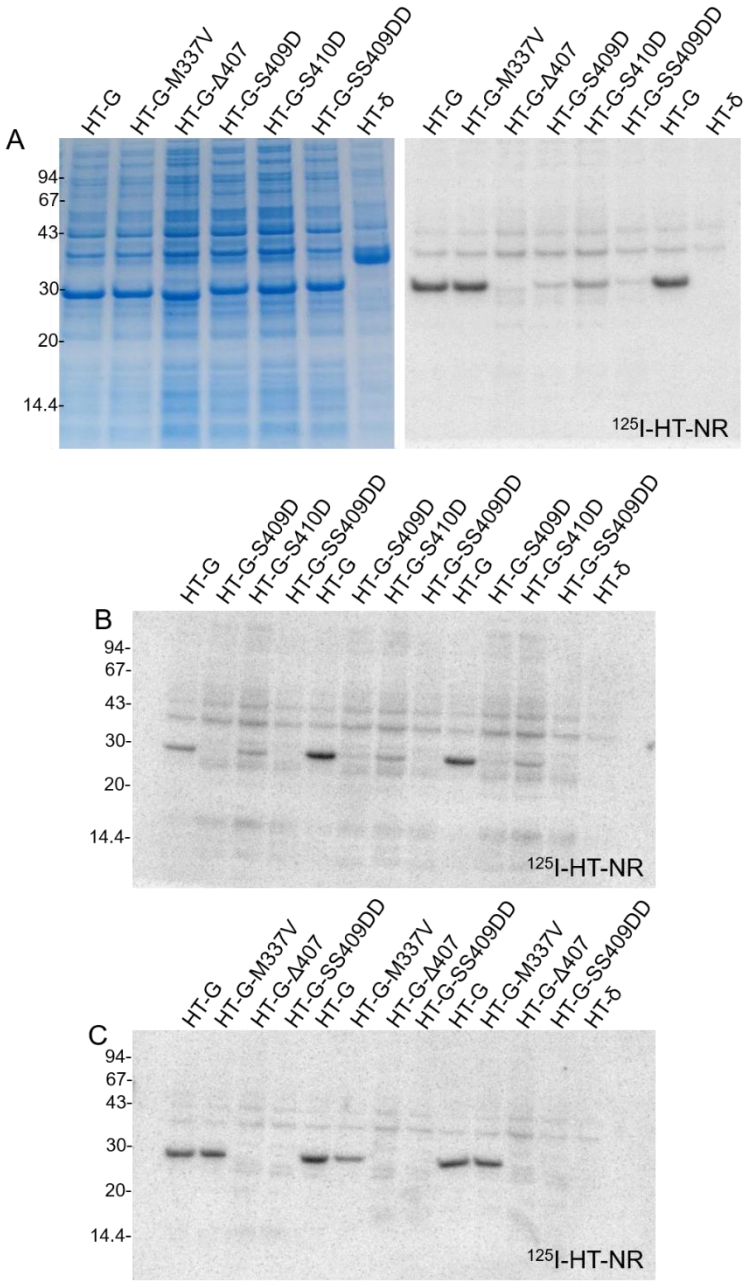


Figure 18. Far Western blot results for the C-terminal domain variants. A) Shown are the Coomassie blue stained gel and exposed blot. The TDP-43 variants were blotted and incubated overnight with the probe, ^{125}I-HT-NR. B) and C) show duplicated sets of the C-terminal domain variants for monitoring reproducibility of the binding level. The radioactive signal was observed using a phosphorimager.

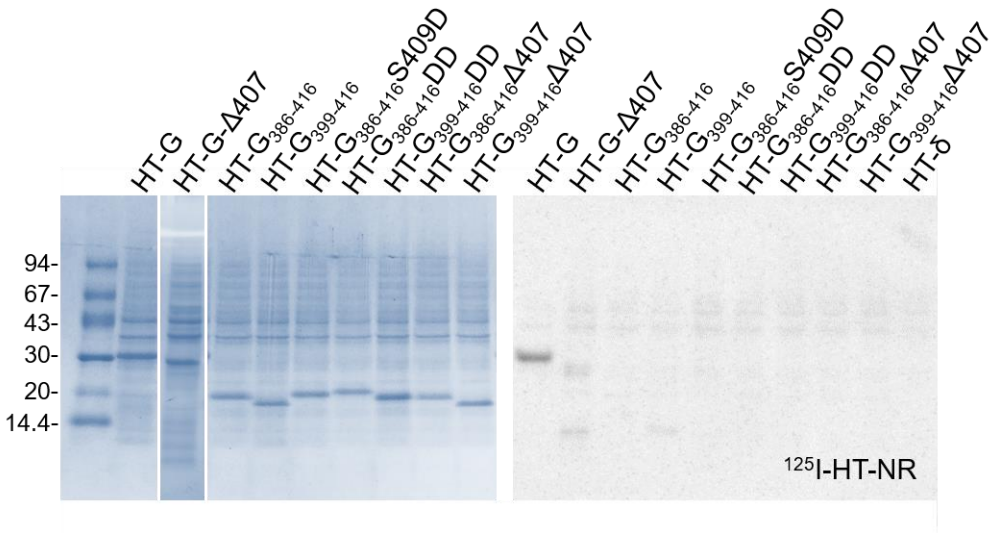


Figure 19. Far Western blot results for the C-terminal domain deletion variants. Shown are the Coomassie blue stained gel and exposed blot. The TDP-43 variants were blotted and incubated overnight with the probe, ^{125}I-HT-NR. HT-G was included for a positive control and HT-G-Δ407 and HT-δ were included for negative controls.

for the constructs HT-$G_{386-416}$ and HT-$G_{399-416}$ especially, as they were expected to maintain some capacity for interaction. It is possible that the binding region includes additional residues within the C-terminal domain and those remaining in these constructs were not sufficient to maintain an interaction to be detected by this method.

As mentioned in the introductions, TDP-43 is typically reported as a 414 amino acid protein but this work has used a 416 amino acid isoform. It was of interest to create the C-terminal domain variant with the last two residues deleted (HT-G-ΔNH, comparable to the 414 amino acid TDP-43) and compare its binding with ^{125}I-HT-NR to the larger isoform. In Figure 20A, three dilution levels of the C-terminal domain variants were blotted. As previously observed, HT-G-S409D had reduced binding compared to HT-G. HT-G-ΔNH appeared to have increased binding compared to HT-G that was reduced in the phosphomimetic, HT-G-S409D-ΔNH. Visually, HT-G-ΔNH appeared to have 2x stronger binding with the probe than HT-G. This was verified through repeated sets of HT-G and HT-G-ΔNH in Figure 20B.

3.4 Analytical Ultracentrifugation

Analytical ultracentrifugation (AUC) is a useful tool for analyzing the solution behaviour of proteins including shape and interactions. Both sedimentation equilibrium (SE) and sedimentation velocity (SV) were used in this work to further characterize the TDP-43 variants and their interactions.

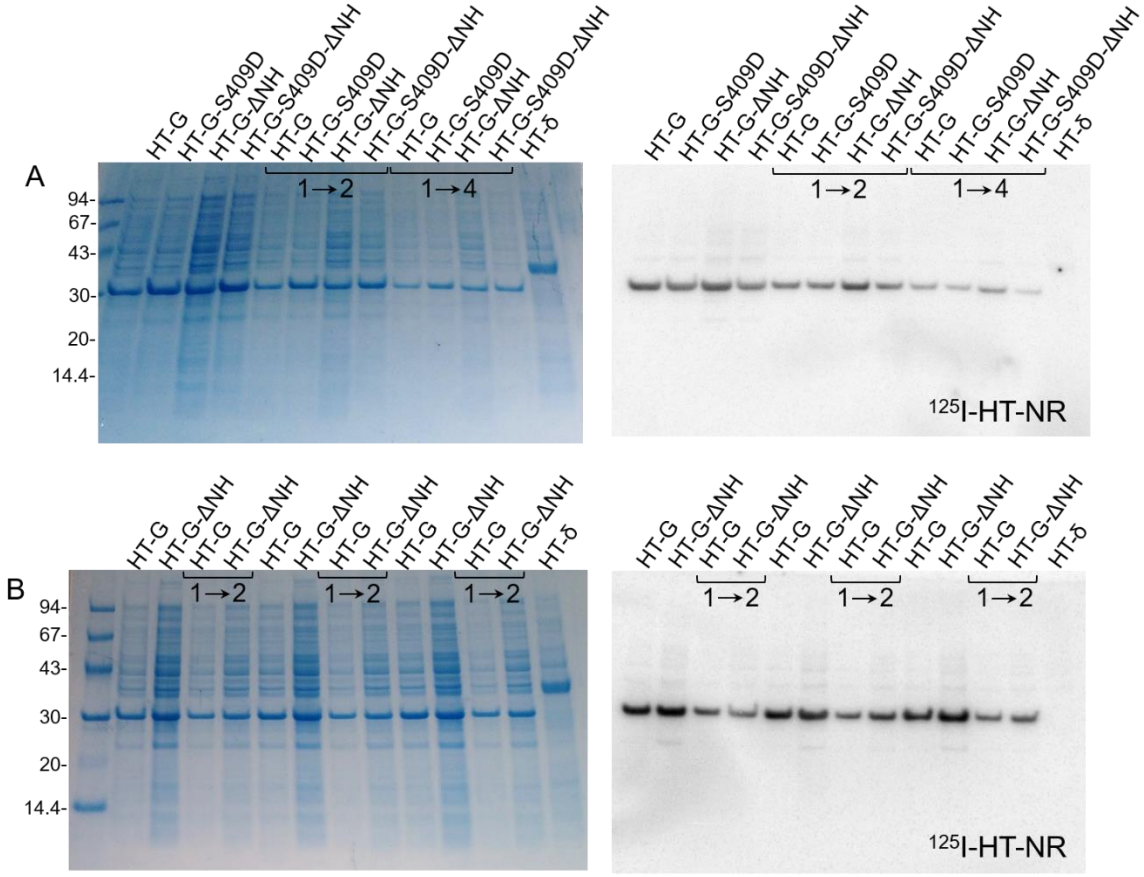


Figure 20. Far Western blot results for the C-terminal domain variant, HT-G-ΔNH. Shown are the Coomassie blue stained gel and exposed blot. The TDP-43 variants were blotted and incubated overnight with the probe, ^{125}I-HT-NR. HT-δ was included as a negative control. A) The four blotted variants were loaded at three dilution levels: 1 to 1, 1 to 2 and 1 to 4. B) The 1 to 1 and 1 to 2 samples of HT-G and HT-G-ΔNH were blotted in triplicate to verify the reproducibility of the detected interaction level.

3.4.1 Sedimentation velocity of His_6-Thioredoxin—To verify that the HT tag would not form homodimers, SV experiments were done to observe any concentration dependant dimer formation (Figure 21). At increasing concentrations of purified HT (1.7 mg/mL, 2.8 mg/mL and 5.6 mg/mL), there was no change in the S_{obs} for the major sedimenting species. Therefore, there was no dimerization of HT and it is a reasonable tag for further studies.

3.4.2 Classification of individual TDP-43 variants—Purified HT-NR and HT-G were first analyzed separately by SE and SV to establish the behaviour of each.

Purified HT-NR was dialyzed into 50 mM Tris-HCl, 0.1 mM EDTA, 1 mM TCEP, pH 8.0 and diluted to A_{280}=0.25. The protein was loaded into the centrifuge, as described in Experimental Procedures, for both SE and SV. HT-NR sedimented as a single, monomeric species in both SE and SV experiments. From SE, the experimental molecular weight (MW_{exp}) was calculated to be 45.6 kDa by fitting the data to a global single ideal species model (Figure 22A). HT-NR also sedimented largely as a single species with an observed sedimentation coefficient (S_{obs}) =2.08 S and with a fitted MW_{exp}=42.6 kDa (Figure 22B). The smaller peak, corresponding to a faster sedimenting species, was likely a small amount of HT-NR homodimer. This is consistent with the Far Western blots that had a low level of 125I-HT-NR interaction with HT-NR (Figure 17).

Purified HT-G was dialyzed into 25 mM Tris-HCl, 0.1 mM TFA, 1 mM TCEP and diluted to A280=0.25. The protein was loaded into the centrifuge, as described in Experimental Procedures for both SE and SV. HT-G sedimented as a single,

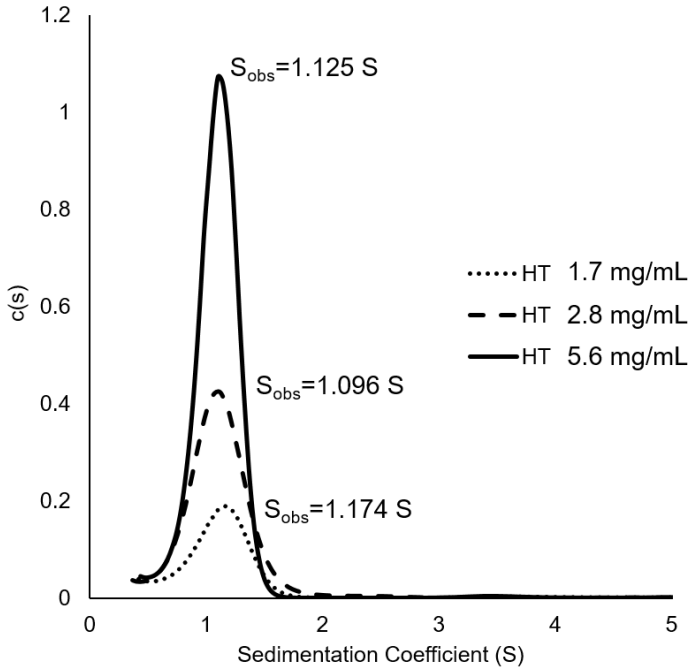
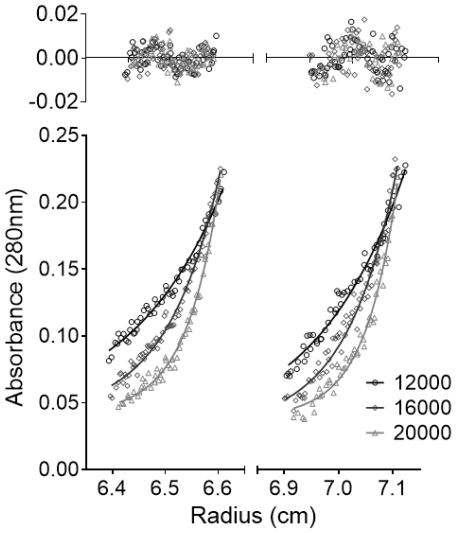


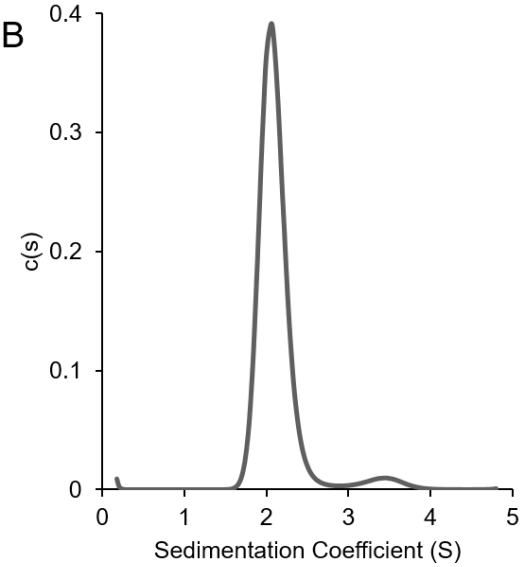
Figure 21. Sedimentation velocity experiments for various concentrations of HT. The cells were loaded with 1.7 mg/mL, 2.8 mg/mL or 5.6 mg/mL of purified HT in into 50 mM Tris-HCl, 0.1 mM EDTA, 1 mM TCEP, pH 8.0. HT $\bar{\upsilon}$=0.7383 mL/g, ρ=0.9984 g/mL and η=1.009 cP. All samples were centrifuged at 45,000 rpm, 4°C and sedimentation was monitored by A_{280} every 10 min for 30 scans. The data were analyzed by Sedfit, using a continuous c(s) distribution model.

Figure 22. Equilibrium and velocity sedimentation results for HT-NR. A) For SE, the cell was loaded with 0.22 mg/mL HT-NR in SV to establish the behaviour of each. Purified HT-NR was dialyzed into 50 mM Tris-HCl, 0.1 mM EDTA, 1 mM TCEP, pH 8.0and sedimentation was observed by A_{280} after equilibration at 12,000, 16,000 and 20,000 rpm at 4°C. The data were fit to a global single ideal species model with a calculated MW_{exp}= 45.6 kDa. The calculated residuals are shown above the plot. B) For SV, the cell was loaded with HT-NR at 0.23 mg/mL, centrifuged at 45,000 rpm at 4°C and sedimentation was monitored by A_{280} every 10 min for 30 scans. The data were analyzed by Sedfit, using a continuous c(s) distribution model. A single species with S_{obs}=2.08 S and with a predicted MW_{exp}=42.6 kDa. Both SE and SV were analyzed with $\bar{\upsilon}$=0.7349 mL/g, ρ=1.00145 g/mL and η=1.017 cP. C) Summary table of the information from the SE and SV experiments. The polypeptide MW was calculated from the amino acid sequence and assumed removal of the N-terminal Met.

A



B



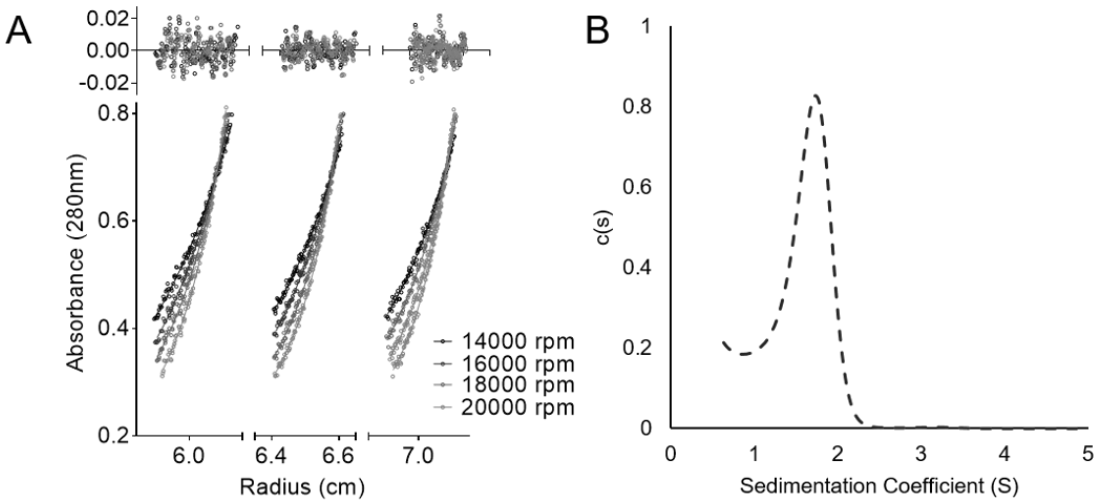
C

	HT-NR
Polypeptide MW	45,400 Da
Equilibrium Sedimentation	
MW_{exp}	45.6 kDa
Velocity Sedimentation	
S_{obs}	2.08 S
$S_{(20,w)}$	3.33 S
Frictional Ratio	1.342
MW_{exp}	42.6 kDa

monomeric species in both SE and SV experiments. From SE, the MWexp was calculated to be 22.2 kDa by fitting the data to a global single ideal species model (Figure 23A). In SV, HT-G sedimented as a single species with a sedimentation coefficient=1.60 S and with a fitted MWexp=20.8 kDa (Figure 23B). It is of note that the fitted molecular weights for both SE and SV were much lower than the calculated molecular weight.

3.4.3 Sedimentation velocity of combined TDP-43 domains—SV experiments were then performed to see if the two isolated fragments of TDP-43 interact when mixed in solution. In the first experiment (Figure 24A), HT-NR and HT-G were dialyzed into 25 mM Tris-HCl, 0.1 mM TFA, 1 mM TCEP, pH 7.3 and the individual polypeptides and the mixture were compared. To detect transient or weak interactions, the slower sedimenting HT-G was added in excess to HT-NR to maximize the effect of weak binding by HT-NR as it sediments through HT-G. The peaks corresponding to the HT-G boundary overlap and have similar S_{obs} (1.583 S or 1.603 S). In the combined experiment, the peak for HT-NR (2.267 S) was shifted to a larger S_{obs} compared to HT-NR only (2.174 S). This suggested that there was residual binding with HT-NR and HT-G which increased the sedimentation rate of HT-NR. HT-G was not similarly affected because it sedimented largely without HT-NR influence. There is again a small peak in HT-NR alone corresponding to a faster sedimenting species that is likely a small amount of HT-NR homodimer (S=3.472 S). In the combined run, there is a similar faster sedimenting species that is shifted (S=3.560 S), likely due to interaction with HT-G. The 25 mM Tris-HCl, 0.1 mM TFA, 1 mM TCEP, pH 7.5, buffer was very

Figure 23. Equilibrium and velocity sedimentation results for HT-G. A) For SE, the cell was loaded with 0.19 mg/mL HT-G in 25 mM Tris-HCl, 0.1 mM TFA, 1 mM TCEP and sedimentation was observed by A_{280} after equilibration at 14,000, 16,000, 18,000 and 20,000 rpm at 4°C. The data were fit to a global single ideal species model with a calculated MW_{exp}=22.2 kDa. The calculated residuals are shown above the plot. B) For SV, the cell was loaded with HT-G at 0.19 mg/mL, centrifuged at 45,000 rpm at 4°C and sedimentation was monitored by A_{280} every 10 min for 30 scans. The data were analyzed by Sedfit, using a continuous c(s) distribution model. A single species with S_{obs}=1.60 S was observed, with a predicted MW_{exp}=20.8 kDa. Both SE and SV were analyzed with $\bar{\upsilon}$=0.7142 mL/g, ρ=0.9991 g/mL and η=1.009 cP. C) Summary table of the information from the SE and SV experiments. The polypeptide MW was calculated from the amino acid sequence and assumed removal of the N-terminal Met.



C

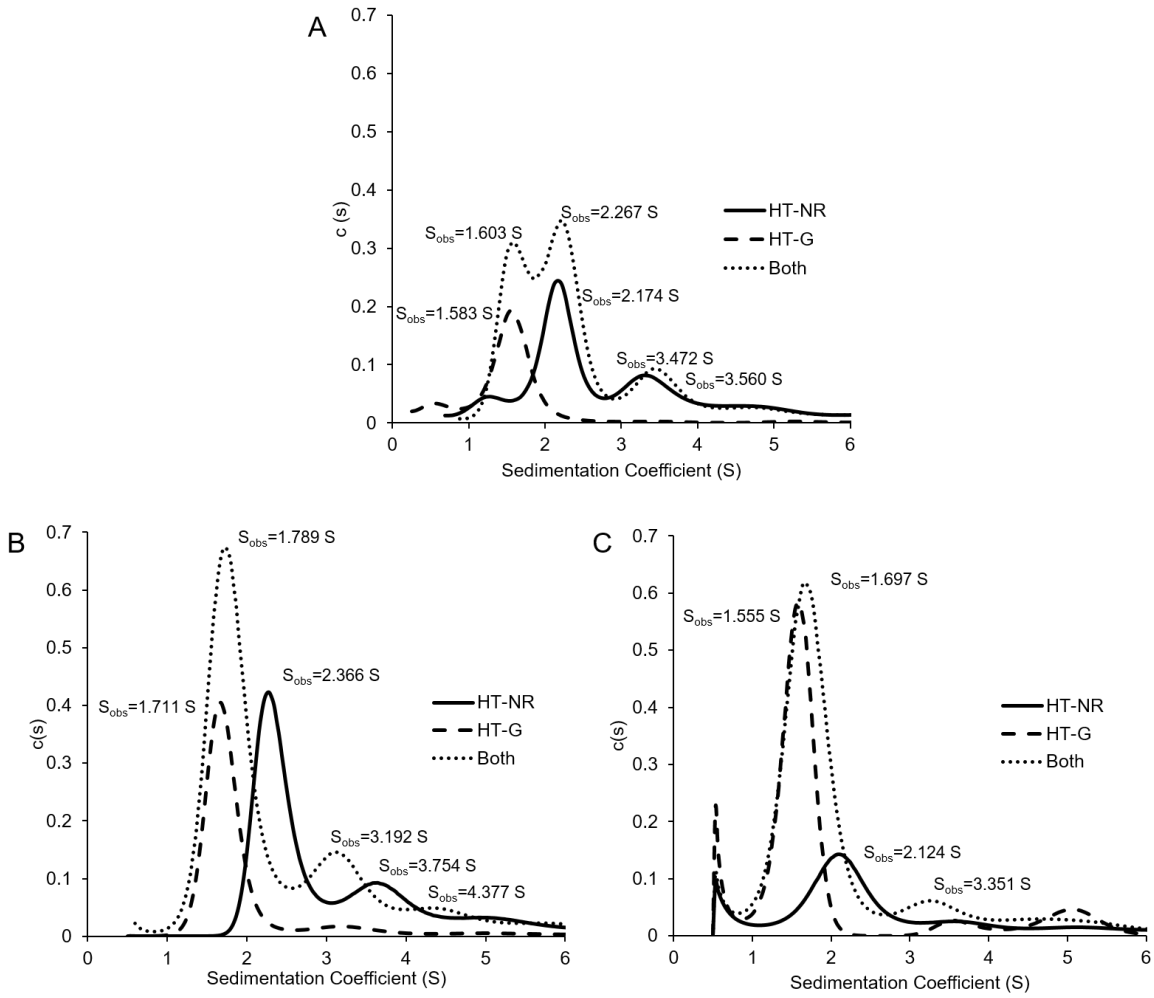
	HT-G
Polypeptide MW	30,580 Da
Equilibrium Sedimentation	
MW_{exp}	22.2 kDa
Velocity Sedimentation	
S_{obs}	1.73 S
$S_{(20,w)}$	1.79 S
Frictional Ratio	1.50
MW_{exp}	19.4 kDa

low in ionic strength which may not promote protein interactions. Furthermore, while the concentration of TFA was low, it might also affect such interactions. Therefore, it was of interest to change to a different buffer.

The second experiment (Figure 24B) used 25 mM Tris-HCl, 25 mM tricine, 1 mM TCEP, pH 7.3, as the sample buffer. HT-NR was added in excess to HT-G. The sample of HT-G in this experiment was solubilized first in ArgHCl, then in TFA. TFA was left out of the final buffer for SV to reduce the effects of TFA on the structure and interactions made in solution. The residual TFA after dialysis into the 25 mM Tris-HCl, 25 mM tricine, 1 mM TCEP, pH 7.3, should not have an effect on the results. In the combined sample, there was a single sedimenting species detected with S_{obs}=1.789 S. This boundary was sedimenting faster than HT-G alone (S_{obs}=1.711 S) and slower than HT-NR alone (S_{obs}=2.366 S). There was also an even faster sedimenting species at S_{obs}=3.192 S in the combined sample.

In the third experiment, the polypeptides were dialyzed into 25 mM Tris base, 50 mM tricine, 1 mM TCEP, pH 7.5. This sample of HT-G was dialyzed directly from ArgHCl and therefore was not exposed to any TFA to prevent any structural influence due to TFA binding. HT-G was also added in excess to HT-NR to provide the same opportunity for observing weak binding as in the initial experiment. Figure 24C shows the results for this SV run. Similar to the second experiment, when HT-NR and HT-G were mixed, a single sedimenting species was observed with a S_{obs}=1.697 S, which was again faster than HT-G alone (S_{obs}=1.555 S) and slower than HT-NR alone (S_{obs}=2.124 S). A faster sedimenting species (S_{obs}=3.351 S) was again observed in the combined sample.

Figure 24. Sedimentation velocity results of the HT-NR (solid), HT-G (dashed) and combined (dotted) in various buffers. A) The cells were loaded with HT-NR (0.12 mg/mL), HT-G (0.25 mg/mL) or Both (0.37 mg/mL), all in 25 mM Tris-HCl, 0.1 mM TFA, 1 mM TCEP, pH 7.3. The S_{obs} are shown on the graph. B) The cells were loaded with HT-NR (0.61 mg/mL), HT-G (0.29 mg/mL) or Both (0.90 mg/mL), all in used 25 mM Tris-HCl, 25 mM tricine, 1 mM TCEP, pH 7.3. The S_{obs} are shown on the graph. C) The cells were loaded with HT-NR (0.18 mg/mL), HT-G (0.18 mg/mL) or Both (0.75 mg/mL), all in used 25 mM Tris base, 50 mM tricine, 1 mM TCEP, pH 7.5. The S_{obs} are shown on the graph. SV data were analyzed with HT-NR $\bar{\upsilon}$=0.7349 mL/g, HT-G $\bar{\upsilon}$ = 0.7142 mL/g, Both $\bar{\upsilon}$=0.7246 mL/g, ρ = 1.0004 g/mL and η = 1.009 cP. All samples were centrifuged at 45,000 rpm, 4°C and sedimentation was monitored by A_{280} every 10 min for 30 scans. The data were analyzed by Sedfit, using a continuous c(s) distribution model.


A
c (s)
Sedimentation Coefficient (S)
S_{obs}=1.603 S
S_{obs}=2.267 S
S_{obs}=1.583 S
S_{obs}=2.174 S
S_{obs}=3.472 S
S_{obs}=3.560 S
HT-NR
HT-G
Both
B
c(s)
Sedimentation Coefficient (S)
S_{obs}=1.789 S
S_{obs}=2.366 S
S_{obs}=1.711 S
S_{obs}=3.192 S
S_{obs}=3.754 S
S_{obs}=4.377 S
HT-NR
HT-G
Both
C
c(s)
Sedimentation Coefficient (S)
S_{obs}=1.697 S
S_{obs}=1.555 S
S_{obs}=2.124 S
S_{obs}=3.351 S
HT-NR
HT-G
Both

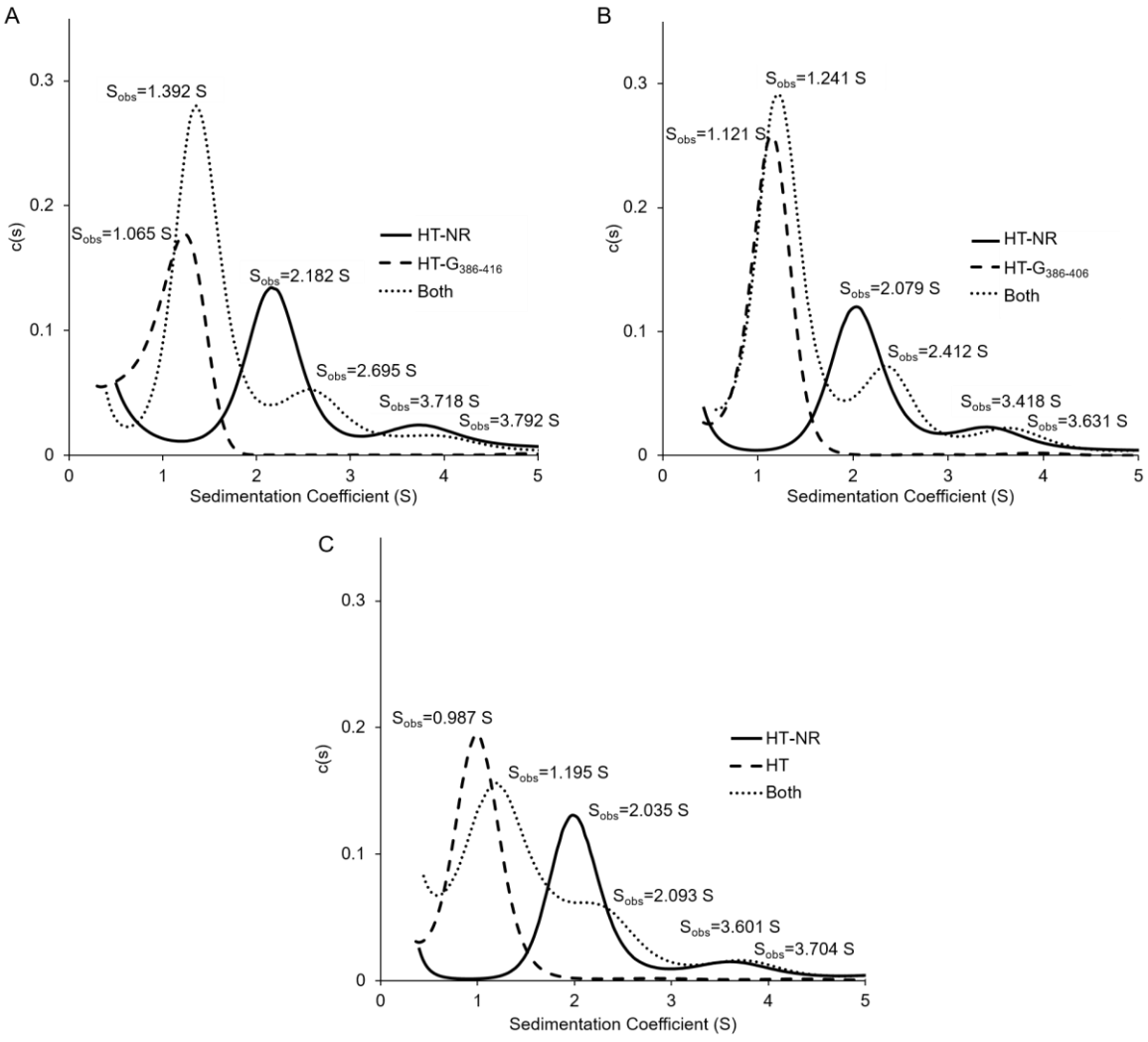
Again in these experiments, small peaks corresponding to a faster sedimenting HT-NR homodimer were observed in the HT-NR alone samples.

To test the theory that the extreme C-terminus of TDP-43 is responsible for binding, two of the deletion mutants were purified for analysis by SV. HT-NR, HT-$G_{386-416}$ and HT-$G_{386-406}$ were dialyzed against 25 mM Tris base, 50 mM tricine, 0.5 mM TCEP, pH 7.5 and were loaded into the centrifuge. The combination samples were loaded with a molar excess of the HT-G variants. As a negative control, a sample of the HT tag was also purified and analyzed by SV. In all experiments, the HT-G variant or HT was in excess to HT-NR.

Figure 25A shows the results for the SV with HT-$G_{386-416}$. It was predicted that this shorter region of the C-terminal domain would form a heterodimer with HT-NR that would sediment faster than HT-NR alone and would not form the extended structure as suggested in the previous experiment. In the combined sample, the strongest boundary sedimenting at S_{obs}=1.392 S was faster than HT-$G_{386-416}$ alone (S_{obs}=1.065 S) and slower than HT-NR alone (S_{obs}=2.182 S). A faster sedimenting peak at S_{obs}=2.695 S was also present in the combined sample.

Figure 25B shows the results for the SV with HT-$G_{386-406}$. It was hypothesized that this even shorter region of the C-terminal domain would not participate in any interactions, as observed in the Far Western blot experiments. Again in the combined sample, a boundary was observed at S_{obs}=1.241 S that was sedimenting slightly faster than HT-G alone, S_{obs}=1.121 S, and slower than HT-NR alone, S_{obs}=2.079 S. There was again a faster sedimenting species at S_{obs}=2.412 S.

Figure 25. Sedimentation velocity results for the truncated HT-G constructs. A) The cells were loaded with HT-NR (0.21 mg/mL), HT-$G_{386\text{-}416}$ (0.18 mg/mL) or Both (0.39 mg/mL), all in 25 mM Tris base, 50 mM tricine, 0.5 mM TCEP, pH 7.5. The S_{obs} are shown on the graph. B) The cells were loaded with HT-NR (0.21 mg/mL), HT-$G_{386\text{-}406}$ (0.18 mg/mL) or Both (0.39 mg/mL), all in 25 mM Tris base, 50 mM tricine, 0.5 mM TCEP, pH 7.5. C) The cells were loaded with HT-NR (0.21 mg/mL), HT (0.17 mg/mL) or Both (0.38 mg/mL), all in 25 mM Tris base, 50 mM tricine, 0.5 mM TCEP, pH 7.5. The S_{obs} are shown on the graphs. SV data were analyzed with HT-NR $\bar{\upsilon}$=0.7349 mL/g, HT-$G_{386\text{-}416}$ $\bar{\upsilon}$ = 0.7282 mL/g, HT-$G_{386\text{-}406}$ $\bar{\upsilon}$ = 0.7304 mL/g, HT $\bar{\upsilon}$=0.7378 mL/g, $Both_A$ $\bar{\upsilon}$=0.7316 mL/g, $Both_B$ $\bar{\upsilon}$=0.7327 mL/g, $Both_C$ $\bar{\upsilon}$= 0.7364 mL/g, ρ = 1.0004 g/mL and η = 1.009 cP. All samples were centrifuged at 45,000 rpm, 4°C and sedimentation was monitored by A_{280} every 10 min for 30 scans. The data were analyzed by Sedfit, using a continuous c(s) distribution model.


A
0.3
0.2
0.1
0
c(s)
S_{obs}=1.392 S
S_{obs}=1.065 S
S_{obs}=2.182 S
S_{obs}=2.695 S
S_{obs}=3.718 S
S_{obs}=3.792 S
HT-NR
HT-$G_{386-416}$
Both
0 1 2 3 4 5
Sedimentation Coefficient (S)
B
0.3
0.2
0.1
0
c(s)
S_{obs}=1.241 S
S_{obs}=1.121 S
S_{obs}=2.079 S
S_{obs}=2.412 S
S_{obs}=3.418 S
S_{obs}=3.631 S
HT-NR
HT-$G_{386-406}$
Both
0 1 2 3 4 5
Sedimentation Coefficient (S)
C
0.3
0.2
0.1
0
c(s)
S_{obs}=0.987 S
S_{obs}=1.195 S
S_{obs}=2.035 S
S_{obs}=2.093 S
S_{obs}=3.601 S
S_{obs}=3.704 S
HT-NR
HT
Both
0 1 2 3 4 5
Sedimentation Coefficient (S)

The final graph, Figure 25C, shows the results of the potential influence of pure HT on the sedimentation of HT-NR. The two peaks in the combined sample were not well resolved in the analysis and the apparent shift in the slower sedimenting species, S_{obs}=1.195 S, may be influenced by the overlapping peaks. The peaks at S_{obs}=2.035 S and S_{obs}=2.093 S are likely both HT-NR monomers, the latter shifted due to the overlapping peaks or analysis variation.

In all HT-NR experiments, a faster sedimenting species corresponding to an HT-NR homodimer was observed.

3.5 *In vivo* Colocalization

HEK293T cells were cotransfected with fluorescently labelled NR and G, then visualized with confocal microscopy to observe colocalization by these polypeptides *in vivo*. The cells expressed either GFP-NR/mCherry-G or mCherry-NR/GFP-G; colocalization is indicated by yellow fluorescence due to overlapping GFP (green) and mCherry (red). Figure 26 shows that in both combinations, there was colocalization of the two fragments of TDP-43 indicated by yellow fluorescence in the merged image. It was also interesting to note that the NR fragment contains both the NLS and NES but both fragments appeared to localize to the nucleus. This is demonstrated by the overlap of the blue Hoechst fluorescence with the red and green from the expressed polypeptides.

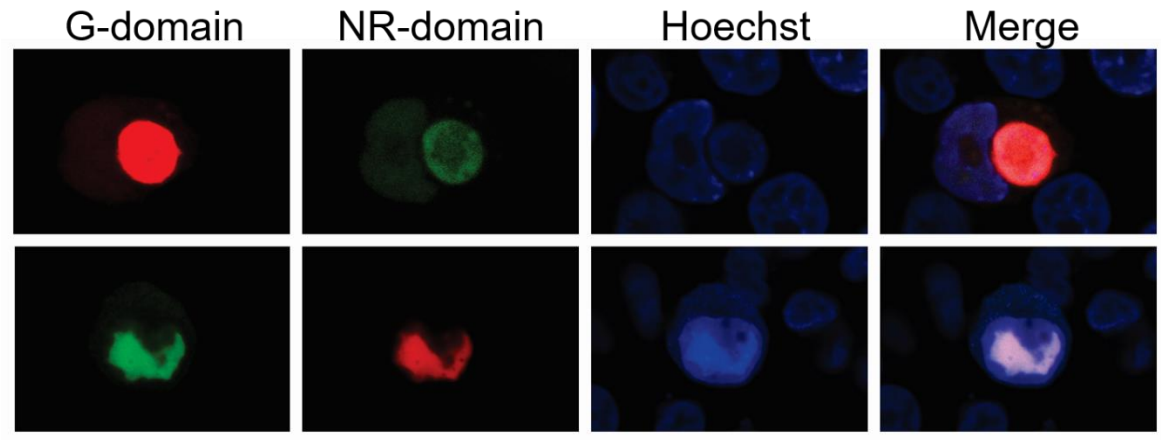


Figure 26. Confocal microscopy of fluorescently labelled TDP-43 domains in HEK293T cells. The top row shows cells cotransfected with mCherry-G/GFP-NR and Hoechst stained. The bottom row shows cells cotransfected with GFP-G/mCherry-NR and Hoechst stained. Both show colocalization of the two fragments of TDP-43. The TDP-43 fragments also appeared to localize to the nucleus as indicated by the overlap with the blue fluorescence. The images were taken at 600x magnification.

Chapter 4: Discussion, Conclusion and Future Direction

4.1 Detecting TDP-43 Domain Interactions

Two biophysical techniques, Far Western blotting and AUC, were used to investigate the existence of an interaction between the C-terminal domain and either the N-terminal domain or the RNA-binding domain of TDP-43. Recombinant TDP-43 constructs containing various combinations of the three domains of TDP-43 were used in these techniques and analyzed for intermolecular interactions. TDP-43 is a highly aggregation prone protein and it was difficult to isolate because it aggregated in inclusion bodies after overexpression in *E. coli*. The recombinant polypeptides included a His_6-thioredoxin (HT) tag that was utilized in affinity purification and in solubility. *E. coli* thioredoxin is a well-studied molecule that increases solubility of proteins that are prone to form inclusion bodies, such as TDP-43 (104). For this reason, the polypeptides were not cleaved from the HT-tag for this work.

In a Far Western blot, the blotted polypeptides were probed with a radioactively labelled protein probe that was a suspected binding partner. This technique requires a strong binding interaction that is detectable at very low probe concentrations compared to the blotted polypeptides during incubation and physical separation after significant washing to dissociate non-specific binding and excess radiation. This is in contrast to AUC, where the two polypeptides were present in much higher concentrations and the faster sedimenting polypeptide is always in contact with its potential binding partner, maximizing the effects of an

interaction on the rate of sedimentation. This situation is demonstrated in Figure 27. For this reason, much weaker interactions are detectable by AUC that may have presented as negative in a Far Western blot. These two techniques complemented each other to establish both strength and minimal residue requirements of the domains of TDP-43 for interaction.

4.1.1 The effects of buffers in AUC—The influence of various salts on protein stability in solution has been well studied and those that more closely mimic physiological levels are ideal for studying native interactions (105). TDP-43 is an aggregation-prone protein and many of the typical buffers for AUC, particularly high salt buffers, were not suitable for studying interactions because they caused rapid precipitation of HT-G. Some work was done to introduce low levels of NaCl into the buffer but HT-G rapidly precipitated and was not usable (data not shown). Instead, the purified domains were typically solubilized in buffers with components that prevent protein interactions/aggregation including ArgHCl (not used in AUC) and TFA, with low levels of Tris base and tricine. TFA is a known additive for dissolving and solubilizing aggregation prone proteins, including β-amyloid (103). Adding TFA was useful for the initial AUC interactions studies with HT-G and there appeared to be some shift in the sedimentation of HT-NR. However, the mechanism by which TFA solubilizes proteins is not well understood and it was not determined if TFA was removed with dialysis. Therefore, it was of interest to find different conditions to further investigate binding to eliminate any possibility of TFA affecting interactions. The solubility of HT-G in low levels of Tris-HCl, Tris base and tricine was critical to the interaction studies.

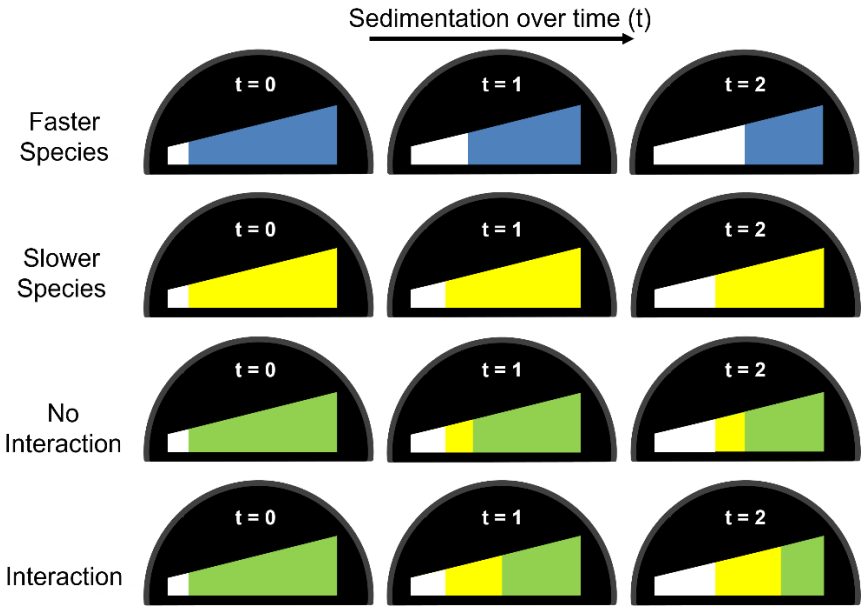


Figure 27. Schematic representation of the sedimentation of two species, with and without interaction. The black figures represent the centrepiece where the sample is loaded for analysis. White space is buffer. Blue is the faster sedimenting species that over time, sediments to the back of the centrepiece. Yellow is the slower species that sediments at a slower rate compared to blue. Green represents the space where both species are sedimenting. In a non-interacting system, the two species sediment at the same rates as observed individually. In an interacting system, the blue species sediments at a faster rate than when alone because it is always in the presence of yellow (shown in green). Yellow sediments at the same rate because it has lost contact with blue.

By dialyzing HT-G directly into a buffer with 25 mM Tris base and 50 mM tricine, the effects of bound TFA were eliminated and more reliable results achieved from the SV studies.

4.1.2 The RNA-binding and C-terminal domains interact in Far Western blots—To investigate the ability of TDP-43 domains to interact, Far Western blots were used to visualize domain interactions by probing immobilized TDP-43 with purified ^{125}I-HT-NR or ^{125}I-HT-N (Figure 17). Strong interactions were observed between ^{125}I-HT-NR and HT-NRG, HT-RG, HT-R and HT-G while ^{125}I-HT-NR/HT-NR had a weak signal. No interaction was observed between ^{125}I-HT-NR and HT-N. These results suggested that the RNA-binding domain and C-terminal domain likely contained the major sites of interaction. Therefore, it is possible that in the intact protein the intrinsically disordered C-terminal domain folds back to interact with RNA-binding domain to modulate some function of TDP-43 similar to that in other IDPRs (92, 93). In TDP-43, this could involve RNA-binding or cellular localization by blocking the NES. The C-terminal domain has been observed to bind other proteins (33, 42). These interactions may relieve intramolecular inhibition by the C-terminal domain or may be blocked by the interaction. In either case, cellular functions of TDP-43 would be modulated by a C-terminal/RNA-binding domain interaction.

The weaker interactions with ^{125}I-HT-NR/HT-NR and ^{125}I-HT-N with any of the blotted TDP-43 constructs suggest that TDP-43 may participate in additional interactions but these may only assist the stronger, RNA-binding/C-terminal

domain interaction or may require additional molecules such as RNA or proteins. Dimerization by the N-terminal domain has been observed through crosslinking techniques (48) and it is possible that Far Western blotting is not a sufficient technique for detecting this interaction.

One common ALS mutation in the C-terminal domain, Met337Val, was also tested for interaction with ^{125}I-HT-NR (Figure 18). The level of interaction was similar to that of the wildtype HT-G. It would be interesting in further work to compare this with other ALS mutations to see if there is a different result for the other known mutations. It is not clear what the impact of the various mutations is on the fate of TDP-43 but comparing the interaction abilities of the mutants may provide some information.

4.1.3 The extreme C-terminus is critical for interactions—In addition to binding with nucleotides or proteins, inhibition by IDPRs may also be relieved through post-translational modifications such as phosphorylation (92, 93). With some confirmed sites of phosphorylation within the C-terminal domain at Ser409/Ser410 (79), mutants of HT-G were created that contained phosphomimetic residues to observe the effects on the interactions detected by Far Western blotting (Figure 18). In all three phosphomimetic constructs, binding with ^{125}I-HT-NR was reduced. Of the single phosphomimetics, HT-G-S409D showed a greater reduction in binding with ^{125}I-HT-NR than HT-G-S410D. HT-G-S409D also had similar reduction to that of the double phosphomimetic. A virtually complete reduction was observed in the deletion construct where residues 407-416 were deleted. These results together

indicate that the extreme C-terminus of TDP-43 is important for the binding observed between the C-terminal domain and HT-NR. Addition of a negative charge by the phosphate group to this region may interrupt the binding interface through steric hindrance or electrostatic repulsion.

4.1.4 HT-NR and HT-G interaction increases sedimentation in solution—The interpretation of the data collected from the sedimentation velocity experiments was not entirely clear but some general trends were evident that do support the presence of an interaction. Intramolecular interactions, such as the one hypothesized in this work, are likely weak in binding energy but are strengthened by proximity when both regions are within the same polypeptide (92). By separating the interacting domains, the weaker interaction was more likely to be observed through a shift in the sedimentation boundary of monomers rather than a single heterodimer boundary. In theory, the effect on sedimentation would be observable because the 45-kDa HT-NR polypeptide was both larger and less extended than the disordered, 30-kDa HT-G polypeptide and the HT-NR would therefore sediment more rapidly than HT-G. If weak interactions occur, the heterodimer would sediment faster, shifting the apparent sedimentation of HT-NR to a larger S_{obs}. HT-NR passing through a “sea” of HT-G would also increase that chances of multiple binding and release events, making the overall evidence of interaction more apparent. The sedimentation of HT-G would not be as significantly influenced because that boundary would be separated from HT-NR during the experiment. (106)

The shifts in the S_{obs} for the peaks in the first three SV experiments (Figure 24) are summarized in Table 3. In each experiment, the boundary assigned to HT-NR sedimented more rapidly in the combined sample than the HT-NR alone. The shift also increased as the buffer conditions became more favourable for protein interactions. This result is as expected from the above explanation of sedimentation and interactions with HT-G and both the monomeric and heterodimeric forms of HT-NR were observed.

4.1.5 The extreme C-terminus is not sufficient for interactions—The results discussed in the previous section supported an interaction by the C-terminal domain and implied that the extreme C-terminus was critical for the interaction. All of the Far Western studies between HT-NR and wild-type HT-G showed a positive interaction. When modifications were made to the extreme C-terminal region of HT-G through phosphomimetics or deletion, the interaction was reduced or nearly abolished. The next step was to determine if this region was sufficient for the interaction.

To further investigate the binding region of the C-terminal domain, mutants were constructed that deleted residues 262-385, leaving the extreme C-terminus attached to HT for interaction studies. There was no observable interaction in the Far Western blots with either of the tested polypeptides (Figure 19). This could be due to two situations: either how these proteins bound to the Immobilon membrane or else that more of the C-terminal domain contributes to binding. In the first situation, accessibility of the probe to the blotted protein domain is dependent on

Buffer Condition		Observed Sedimentation Coefficient (S_{obs}) Shift	
		HT-NR → Combined (Difference)	HT-G → Combined (Difference)
$Tris^{25}TFA^{0.1}TCEP^{1}$	Peak 1	2.174 S → 2.267 S (0.0930 S)	1.583 S → 1.603 S (0.020 S)
	Peak 2	3.472 S → 3.560 S (0.088 S)	
$Tris^{25}Tris^{25}TCEP^{1}$	Peak 1	2.366 S → 3.192 S (0.826 S)	1.711 S → 1.789 S (0.078 S)
	Peak 2	3.754 S → 4.377 S (0.623 S)	
$Tris^{25}Tris^{50}TCEP^{1}$	Peak 1	2.124 S → 3.351 S (1.227 S)	1.555 S → 1.697 S (0.142 S)

Table 3. Observed sedimentation coefficient shifts in SV experiments with HT-NR and HT-G. Three different buffer conditions were used to study the same interaction. Shown are the S_{obs} for the individual samples shifted (→) to the corresponding peak in the mixed sample. The difference is provided in brackets. Peak 1 refers to the monomeric HT-NR boundary. Peak 2 is the potentially homodimeric HT-NR boundary, if resolved by the analysis.

the orientation of the transferred protein on the Immobilon membrane and whether the binding site is accessible to the probe in solution. In these small constructs, the binding region is so close to the His_6-thioredoxin tag, there is a chance that the remaining residues of the C-terminal domain were inaccessible to the probe and therefore no binding could occur. In the second case, there may be necessary residues for binding elsewhere in the C-terminal domain that were removed with both deletion mutants (HT-G-Δ407 and HT-$G_{386-416}$) and deletion of either reduced the binding energy such that it was too low for the Far Western blot technique. Since the C-terminal domain is disordered, it is possible to have multiple sites of contact spread through the domain that individually have low binding energies but collectively create a strong interaction. Through studying the binding kinetics of IDPRs, many of these domains adapt an induced fit structure that depends on an initial binding step that triggers a conformational change for full binding (reviewed in 107). The final bound state would involve multiple sites through the domain and therefore weak binding may be detectable with many regions of the domain.

In the second set of SV experiments (Figure 25), a series of truncations were used to try to elucidate a more defined region of binding by the C-terminal domain. Even though there was no detection of an interaction in the Far Western blot experiments with the HT-$G_{386-416}$ construct, SV would present a more favourable environment for binding because the technique is more sensitive to weak binding due to the higher concentrations employed. Table 4 summarizes the S_{obs} shifts for these experiments. As anticipated, the shift in HT-NR sedimentation observed in the presence of HT-$G_{386-416}$ was greater than that with HT-$G_{386-406}$ or

HT. This agrees with our prediction that there are residues within the extreme C-terminus of TDP-43 that form an interaction with the RNA-binding domain.

HT-$G_{386-406}$ was expected to have no interaction with HT-NR based on the observation in the Far Western blot experiments with HT-G-Δ407. However, the same sensitivity argument mentioned above for using SV applies here. The residues in HT-G-Δ407 may not have had high enough binding energy to be detectable by a Far Western blot but the few remaining residues within HT-$G_{386-406}$ had a strong enough binding to increase the sedimentation of HT-NR detected by SV. Overall, the extreme C-terminus of TDP-43 is critical to the interaction with HT-NR but it is not sufficient for high affinity binding. The results suggested some binding energy is provided by other residues within the C-terminal domain.

Finally, there was very little change in the sedimentation of HT-NR in the presence of the purified HT tag alone. This confirmed that it is the presence of the C-terminal domain of TDP-43 that binds HT-NR and not the tag.

4.1.6 Variation in S_{obs} due to Sedfit analysis and user error—As alluded to previously, the apparent shift in the HT boundary when in the presence of HT-NR may be attributed to the poor resolution of the two peaks during analysis. There is also clear variation in the Sobs for identical polypeptides run in different experiments. The software used for this analysis, Sedfit, required significant user input through placement of limits and boundaries during the analysis. For example, if there is poor data collection due to a small leak during centrifugation, there may be fewer usable data points in one cell compared to another.

	Observed Sedimentation Coefficient (S_{obs}) Shift		
		HT-NR → Combined (Difference)	Other → Combined (Difference)
HT-$G_{386-416}$	Peak 1	2.182 S → 2.695 S (0.513 S)	1.065 S → 1.392 S (0.327 S)
	Peak 2	3.718 S → 3.792 S (0.074 S)	
HT-$G_{386-406}$	Peak 1	2.079 S → 2.412 S (0.333 S)	1.121 S → 1.241 S (0.120 S)
	Peak 2	3.418 S → 3.631 S (0.213 S)	
HT	Peak 1	2.035 S → 2.093 S (0.058 S)	0.987 S → 1.195 S (0.208 S)
	Peak 2	3.601 S → 3.704 S (0.103 S)	

Table 4. Observed sedimentation coefficient shifts in SV experiments with HT-NR and HT-$G_{386-416}$, HT-$G_{386-406}$ and HT. Shown are the S_{obs} for the individual samples shifted (→) to the corresponding peak in the mixed sample. The difference is provided in brackets. Peak 1 refers to the monomeric HT-NR boundary. Peak 2 is the potentially homodimeric boundary.

This results in a smaller radial range of analyzable points, which has some impact on the observed sedimentation coefficients reported. The radial range of data analyzed within an experiment (three SV data sets) was kept relatively constant but between experiments, the range was not as closely monitored. This explains the variability in S_{obs} for HT-NR across the experiments but does not negate the larger shifts due to interaction. Other sources of variability include cell alignment, window imperfections and temperature variability, to name a few. In a study on the impact of such variables, Arthur *et al.* (2008) found that S_{obs} was not significantly affected by these variables (108). Therefore, the above conclusions based on gross changes in sedimentation are more likely due to real interactions and not simply artifacts or experimental error.

4.2 Implications for intramolecular interactions in TDP-43 function

The functions of TDP-43 within a cell involve RNA-binding, protein-protein interactions and transport between the nucleus and cytoplasm (19). These functions of TDP-43 must be tightly regulated by the cell to ensure that miRNA is correctly processed, or mRNA is spliced, transported or even degraded. The presence of an IDPR within TDP-43, the C-terminal domain, suggests the possibility of this domain as a regulatory module that forms an intramolecular interaction to moderate TDP-43, for example, binding with RNA, binding with other proteins or even cellular localization.

The most obvious role of the C-terminal domain would be in regulating or responding to RNA binding. If the binding interface for the C-terminal domain

covers the RNA-binding surface, this intramolecular interaction could act by inhibiting RNA binding. Relieving inhibition could involve post-translational modifications to the C-terminal domain or proximity to an intermolecular binding partner of TDP-43. Alternatively, proximity to an ideal binding sequence within an RNA molecule could dissociate the C-terminal domain, making it available for protein interactions.

Similar to that of RNA-binding, an intramolecular interaction by the C-terminal domain could moderate access to the binding region for interactions with other protein binding partners of TDP-43 within a cell. As discussed in Section 1.2.4, the C-terminal domain participates in a number of protein-protein interactions and through an intramolecular interaction involving the C-terminal domain, these interactions could be inhibited. Upon release, again by RNA-binding or post-translational modification, the region for binding would be exposed to nearby proteins. The C-terminal domain could then bind with the appropriate proteins for its current role, perhaps to be transported out of the nucleus. Upon release of the protein partner after TDP-43 has completed its role, the C-terminal domain could rebind intramolecularly and once more shield interactions.

TDP-43 contains both a NLS and a NES (31). The C-terminal domain could act as a regulator of access to the NLS or NES. The NES is located within the RNA-binding domain (residues 239-250). The experiments in this work have suggested that the C-terminal domain forms an interaction with the RNA-binding domain. In the bound state, the C-terminal domain may prevent recognition of the NES by transport machinery. Masking of localization signals has been observed

in many other proteins (reviewed in 109). Upon release of the C-terminal domain interaction by RNA-binding or post-translational modification, for example, the NES would be exposed and export of TDP-43 could occur. TDP-43 would then be returned to the nucleus after rebinding of the C-terminal domain and subsequent masking of the NES. TDP-43 aggregates are found localized to the cytoplasm of ALS-affected neurons and it is the C-terminal domain that is implicated in nucleating aggregation (1, 3, 31, 64, 110). Loss of intramolecular binding by the C-terminal domain, possibly through excessive interactions leading to aggregation, may strand TDP-43 in the cytoplasm. The *in vivo* confocal microscopy studies demonstrated that GFP-NR and mCherry-G expression were colocalized to the nucleus in HEK293T cells while little (if any) expression was seen in the cytosol. Considering the above discussion about the impact of the NES, there is also the consideration of the NLS, located within the N-terminal domain at residues 82-98. The structure formed by the N-terminal domain is just starting to be studied and it is possible that masking of the NLS also occurs within this domain. At this time, the relative strength between the two signals is not known, but it would be interesting to study further.

4.3 Implications for intramolecular interactions in TDP-43 dependent pathogenesis

4.3.1 ΔNH increased the interaction—The 416 amino acid isoform of TDP-43 used throughout this study was isolated from neuropathologically normal human brain

(18) but there is a 414 amino acid isoform that is more commonly cited in the literature. The 414 amino acid isoform was originally isolated from HeLa cells (7) and is also detectable in HEK293T cells (18). The longer isoform has two amino acid residues added at the C-terminus and its presence in the nervous system has interesting implications on the importance of the extreme C-terminus in TDP-43 regulation. While the expression of this isoform was not determined throughout human tissues, it would be interesting to determine if this isoform is neuronal tissue specific. Deleting the last two residues of the 416 amino acid HT-G, making the HT-G-ΔNH C-terminal domain, appeared to increase binding with ^{125}I-HT-NR about two-fold. If intramolecular binding by TDP-43 is critical to stability and aggregation prevention, expression of the longer isoform in only certain tissues would explain the more aggregation-prone state of TDP-43. The shorter isoform may form a more stable intramolecular interaction and therefore be less susceptible to the triggers for aggregation. TDP-43 is a ubiquitously expressed protein, present in many different tissues (8), but its aggregation has thus far largely been observed in the nervous system. There is some evidence for TDP-43 involvement in inclusion body myopathy, a disease of the muscle, but the isoform of TDP-43 present in muscle was not determined (111). Exploring this interaction further will be interesting for understanding the susceptibility of neurons to aggregation of TDP-43 and if it is related to the weaker binding isoform that is expressed.

4.3.2 Phosphorylation of C-terminal serines in TDP-43—As discussed in Chapter 1, the aggregates found in neurons often contain TDP-43 phosphorylated at Ser409/Ser410 (82) and it is unclear whether phosphorylation or aggregation occurs first. Considering the current results, arguments for both exist (Figure 28). If aberrant phosphorylation of the C-terminal domain occurred, a potentially critical step in moderating TDP-43 function would be disrupted, thus allowing for excessive interactions to occur and aggregate formation. Alternatively, if aggregates of TDP-43 form, phosphorylation and cleavage may be preventative steps by a cell to reduce further aggregation with full length TDP-43 from occurring. In either case, disruption of this interaction could have an impact on the fate of TDP-43 in a neuron. There has been little exploration into any *in vivo* pathways for dephosphorylation of TDP-43.

4.4 Conclusion and Future Directions

My hypothesis was that the C-terminal domain of TDP-43 is involved in intra- or intermolecular interactions. The experimental results presented in this work support my hypothesis and demonstrate that the C-terminal domain is capable of interaction with the RNA-binding domain of TDP-43. Disruption of the interaction due to phosphomimetic mutations, similar to pathological phosphorylation, was also demonstrated. I further hypothesize that this interaction could be critical to TDP-43 function within a cell and disruption is one critical step in disease pathogenesis. Future work should be focused both *in vitro* and *in vivo* to further investigate the functional implications of an intramolecular interaction.

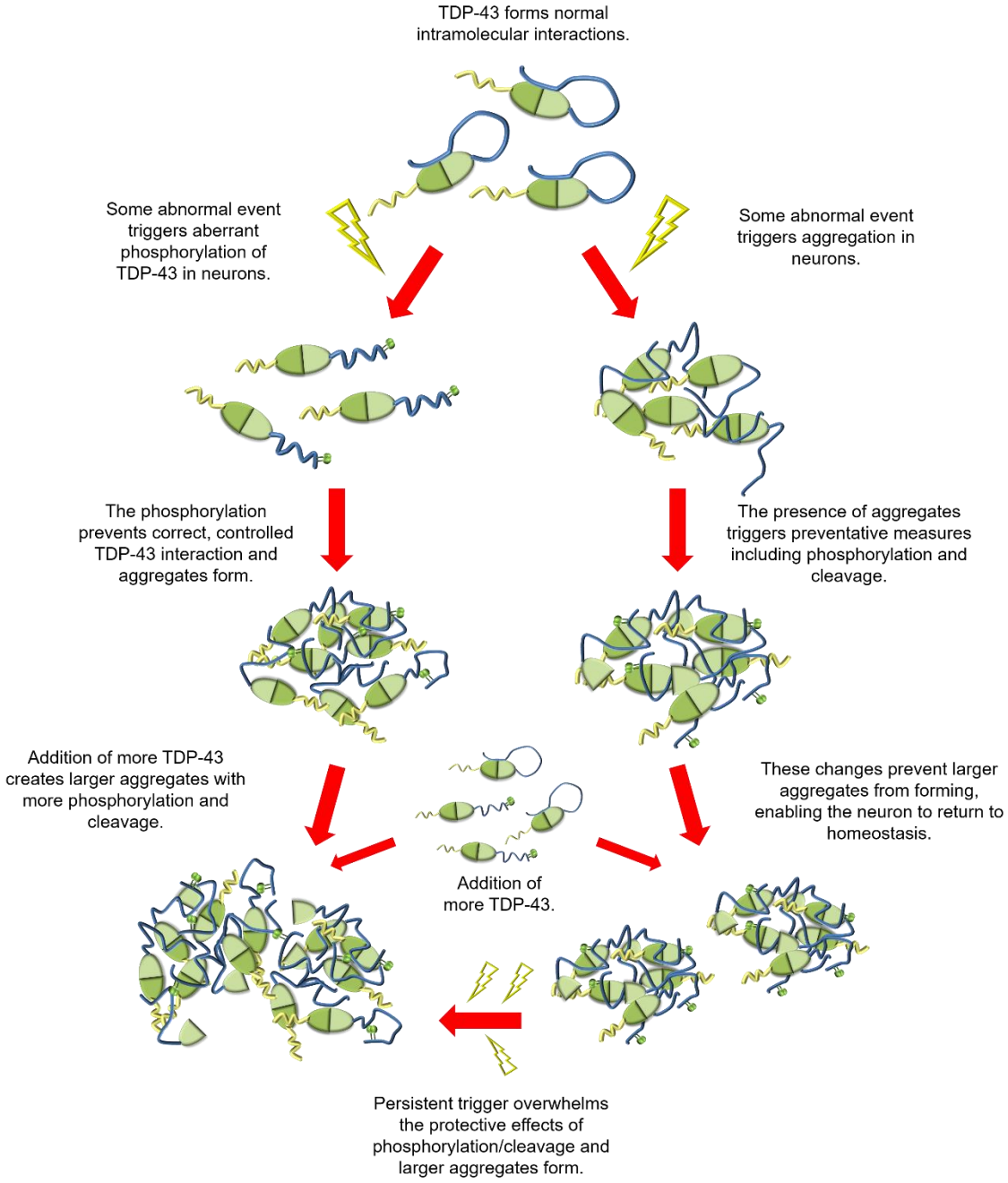


Figure 28. Schematic diagram illustrating two possible functions of phosphorylation in aggregation. On the left side, TDP-43 is aberrantly phosphorylated in the disease state which leads to aggregate formation. On the right side, there is some trigger that causes TDP-43 to aggregate and phosphorylation of TDP-43 is a mechanism to prevent further aggregation.

One important *in vitro* step should be purification of the full length TDP-43 protein with greater yields. This will provide an important piece of information for comparison with the current results. More precise mapping of the binding interface within the RNA-binding domain would be important. This could be achieved by introducing a crosslinking system where a photo-crosslinkable agent would be added in the C-terminal domain and upon binding and activation of crosslinking, the C-terminal domain would be chemically bound at the interface. Then trypsin digestion and a peptide identification technique such as mass spectrometry could follow to identify the bound region in the RNA-binding domain. Once a more precise region of interaction is known for both domains, site directed mutagenesis could be used to mutate certain residues that might be critical for an interaction and then perform the same interaction studies to determine the impact of the mutations.

Investigation of the impact of oligonucleotides on the interaction would provide important information of what functions of TDP-43 might be regulated. Using a technique such as an electrophoretic mobility shift assay, a radio-labelled oligonucleotide could be added with the HT-NR and HT-G to observe how the movement of HT-NR and HT-G are altered by the presence of RNA. The same techniques described in this work and future work should also be used to identify any differences in C-terminal domain binding due to other known mutations in disease. This would complement purification of the mutant C-terminal domains and analysis by AUC. Co-immunoprecipitation could also be used to detect an

interaction between co-expressed TDP-43 polypeptides from a eukaryotic cell lysate.

For *in vivo* work, more colocalization studies would be important for deciphering the biological relevance of the interaction. This could be achieved through studies with lower construct expression and through introducing some of the other mutations or truncations used in this work. Performing an RNA stability assay, such as a luciferase assay, would provide a more direct functional relevance of an interaction through observing the difference between full-length TDP-43 and the co-expression of NR and G polypeptides.

Overall, the work presented here is the initial step to defining an intramolecular interaction in TDP-43. This is important to the field because it provides insight into the possible mechanisms of TDP-43 that may be altered in disease. An interaction such as this has significant implications on how TDP-43 should be studied both in the RNA-binding and neurodegenerative disease fields. The presence of the C-terminal domain in studies is variable depending on the desired results of the work, whether it be observing aggregation with an excess of this domain or determining structural behavior with the C-terminal domain cleaved. This type of interaction is critical to the normal behaviour of TDP-43 and therefore the C-terminal domain should not be cleaved for ease of purification or minimizing *in vivo* aggregation to properly study the natural behaviour of TDP-43. Since TDP-43 is a highly conserved protein, all domains are important to its function. Through interdomain interactions, its variable roles within a cell could be much more tightly regulated and integral to cell survival than initially presumed. Further efforts to

establish the molecular and biological basis for this interaction should provide better insight into the role of TDP-43 in normal cellular homeostasis and in disease pathogenesis.

References

1. Arai, T., Hasegawa, M., Akiyama, H., Ikeda, K., Nonaka, T., Mori, H., Mann, D., Tsuchiya, K., Yoshida, M., Hashizume, Y., and Oda, T. (2006) TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem. Biophys. Res. Commun.* **351**, 602–611
2. Neumann, M., Sampathu, D. M., Kwong, L. K., Truax, A. C., Micsenyi, M. C., Chou, T. T., Bruce, J., Schuck, T., Grossman, M., Clark, C. M., McCluskey, L. F., Miller, B. L., Masliah, E., Mackenzie, I. R., Feldman, H., Feiden, W., Kretzschmar, H. A., Trojanowski, J. Q., and Lee, V. M.-Y. (2006) Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Science*. **314**, 130–133
3. Arai, T. (2014) Significance and limitation of the pathological classification of TDP-43 proteinopathy. *Neuropathology*. 10.1111/neup.12138
4. Van Deerlin, V. M., Leverenz, J. B., Bekris, L. M., Bird, T. D., Yuan, W., Elman, L. B., Clay, D., Wood, E. M., Chen-Plotkin, A. S., Martinez-Lage, M., Steinbart, E., McCluskey, L., Grossman, M., Neumann, M., Wu, I.-L., Yang, W.-S., Kalb, R., Galasko, D. R., Montine, T. J., Trojanowski, J. Q., Lee, V. M.-Y., Schellenberg, G. D., and Yu, C.-E. (2008) TARDBP mutations in amyotrophic lateral sclerosis with TDP-43 neuropathology: a genetic and histopathological analysis. *Lancet Neurol.* **7**, 409–416
5. Yokoseki, A., Shiga, A., Tan, C.-F., Tagawa, A., Kaneko, H., Koyama, A., Eguchi, H., Tsujino, A., Ikeuchi, T., Kakita, A., Okamoto, K., Nishizawa, M., Takahashi, H., and Onodera, O. (2008) TDP-43 mutation in familial amyotrophic lateral sclerosis. *Ann. Neurol.* **63**, 538–542
6. Al-Chalabi, A., Jones, A., Troakes, C., King, A., Al-Sarraj, S., and van den Berg, L. H. (2012) The genetics and neuropathology of amyotrophic lateral sclerosis. *Acta Neuropathol. (Berl.)*. **124**, 339–352
7. Ou, S. H., Wu, F., Harrich, D., García-Martínez, L. F., and 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. *J. Virol.* **69**, 3584–3596

8. Buratti, E., Dörk, T., Zuccato, E., Pagani, F., Romano, M., and Baralle, F. E. (2001) Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. *EMBO J.* **20**, 1774–1784
9. Buratti, E., Brindisi, A., Pagani, F., and Baralle, F. E. (2004) Nuclear Factor TDP-43 Binds to the Polymorphic TG Repeats in CFTR Intron 8 and Causes Skipping of Exon 9: A Functional Link with Disease Penetrance. *Am. J. Hum. Genet.* **74**, 1322–1325
10. Nakashima-Yasuda, H., Uryu, K., Robinson, J., Xie, S. X., Hurtig, H., Duda, J. E., Arnold, S. E., Siderowf, A., Grossman, M., Leverenz, J. B., Woltjer, R., Lopez, O. L., Hamilton, R., Tsuang, D. W., Galasko, D., Masliah, E., Kaye, J., Clark, C. M., Montine, T. J., Lee, V. M.-Y., and Trojanowski, J. Q. (2007) Co-morbidity of TDP-43 proteinopathy in Lewy body related diseases. *Acta Neuropathol. (Berl.)*. **114**, 221–229
11. Arai, T., Mackenzie, I. R. A., Hasegawa, M., Nonoka, T., Niizato, K., Tsuchiya, K., Iritani, S., Onaya, M., and Akiyama, H. (2009) Phosphorylated TDP-43 in Alzheimer's disease and dementia with Lewy bodies. *Acta Neuropathol. (Berl.)*. **117**, 125–136
12. Walker, A. K., Daniels, C. M. L., Goldman, J. E., Trojanowski, J. Q., Lee, V. M.-Y., and Messing, A. (2014) Astrocytic TDP-43 Pathology in Alexander Disease. *J. Neurosci.* **34**, 6448–6458
13. Wang, H.-Y., Wang, I.-F., Bose, J., and Shen, C.-K. J. (2004) Structural diversity and functional implications of the eukaryotic TDP gene family. *Genomics*. **83**, 130–139
14. He, Y., and Smith, R. (2008) Nuclear functions of heterogeneous nuclear ribonucleoproteins A/B. *Cell. Mol. Life Sci.* **66**, 1239–1256
15. Jean-Philippe, J., Paz, S., and Caputi, M. (2013) hnRNP A1: The Swiss Army Knife of Gene Expression. *Int. J. Mol. Sci.* **14**, 18999–19024
16. Ayala, Y. M., Pantano, S., D'Ambrogio, A., Buratti, E., Brindisi, A., Marchetti, C., Romano, M., and Baralle, F. E. (2005) Human, Drosophila, and C.elegans

TDP43: nucleic acid binding properties and splicing regulatory function. *J. Mol. Biol.* **348**, 575–588

17. Consortium, T. U. (2015) UniProt: a hub for protein information. *Nucleic Acids Res.* **43**, D204–D212
18. Strong, M. J., Volkening, K., Hammond, R., Yang, W., Strong, W., Leystra-Lantz, C., and Shoesmith, C. (2007) TDP43 is a human low molecular weight neurofilament (hNFL) mRNA-binding protein. *Mol. Cell. Neurosci.* **35**, 320–327
19. Buratti, E., and Baralle, F. E. (2012) TDP-43: gumming up neurons through protein-protein and protein-RNA interactions. *Trends Biochem. Sci.* **37**, 237–247
20. Buratti, E., and Baralle, F. E. (2001) Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. *J. Biol. Chem.* **276**, 36337–36343
21. Neumann, M., Kwong, L. K., Lee, E. B., Kremmer, E., Flatley, A., Xu, Y., Forman, M. S., Troost, D., Kretzschmar, H. A., Trojanowski, J. Q., and Lee, V. M.-Y. (2009) Phosphorylation of S409/410 of TDP-43 is a consistent feature in all sporadic and familial forms of TDP-43 proteinopathies. *Acta Neuropathol. (Berl.)*. **117**, 137–149
22. Igaz, L. M., Kwong, L. K., Chen-Plotkin, A., Winton, M. J., Unger, T. L., Xu, Y., Neumann, M., Trojanowski, J. Q., and Lee, V. M.-Y. (2009) Expression of TDP-43 C-terminal Fragments in Vitro Recapitulates Pathological Features of TDP-43 Proteinopathies. *J. Biol. Chem.* **284**, 8516–8524
23. Arai, T., Hasegawa, M., Nonoka, T., Kametani, F., Yamashita, M., Hosokawa, M., Niizato, K., Tsuchiya, K., Kobayashi, Z., Ikeda, K., Yoshida, M., Onaya, M., Fujishiro, H., and Akiyama, H. (2010) Phosphorylated and cleaved TDP-43 in ALS, FTLD and other neurodegenerative disorders and in cellular models of TDP-43 proteinopathy. *Neuropathology*. **30**, 170–181
24. Fuentealba, R. A., Udan, M., Bell, S., Wegorzewska, I., Shao, J., Diamond, M. I., Weihl, C. C., and Baloh, R. H. (2010) Interaction with Polyglutamine

Aggregates Reveals a Q/N-rich Domain in TDP-43. *J. Biol. Chem.* **285**, 26304–26314

25. Liachko, N. F., Guthrie, C. R., and Kraemer, B. C. (2010) Phosphorylation promotes neurotoxicity in a C. elegans model of TDP-43 proteinopathy. *J. Neurosci. Off. J. Soc. Neurosci.* **30**, 16208–16219
26. Che, M.-X., Jiang, Y.-J., Xie, Y.-Y., Jiang, L.-L., and Hu, H.-Y. (2011) Aggregation of the 35-kDa fragment of TDP-43 causes formation of cytoplasmic inclusions and alteration of RNA processing. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **25**, 2344–2353
27. Miguel, L., Frébourg, T., Campion, D., and Lecourtois, M. (2011) Both cytoplasmic and nuclear accumulations of the protein are neurotoxic in Drosophila models of TDP-43 proteinopathies. *Neurobiol. Dis.* **41**, 398–406
28. Yang, C., Wang, H., Qiao, T., Yang, B., Aliaga, L., Qiu, L., Tan, W., Salameh, J., McKenna-Yasek, D. M., Smith, T., Peng, L., Moore, M. J., Brown, R. H., Cai, H., and Xu, Z. (2014) Partial loss of TDP-43 function causes phenotypes of amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E1121–1129
29. Buratti, E., and Baralle, F. E. (2008) Multiple roles of TDP-43 in gene expression, splicing regulation, and human disease. *Front. Biosci. J. Virtual Libr.* **13**, 867–878
30. Freibaum, B. D., Chitta, R., High, A. A., and Taylor, J. P. (2010) Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery. *J. Proteome Res.* **9**, 1104–1120
31. Winton, M. J., Igaz, L. M., Wong, M. M., Kwong, L. K., Trojanowski, J. Q., and Lee, V. M.-Y. (2008) Disturbance of Nuclear and Cytoplasmic TAR DNA-binding Protein (TDP-43) Induces Disease-like Redistribution, Sequestration, and Aggregate Formation. *J. Biol. Chem.* **283**, 13302–13309
32. Passoni, M., De Conti, L., Baralle, M., and Buratti, E. (2012) UG Repeats/TDP-43 Interactions near 5′ Splice Sites Exert Unpredictable Effects on Splicing Modulation. *J. Mol. Biol.* **415**, 46–60

33. Buratti, E., Brindisi, A., Giombi, M., Tisminetzky, S., Ayala, Y. M., and Baralle, F. E. (2005) TDP-43 binds heterogeneous nuclear ribonucleoprotein A/B through its C-terminal tail: an important region for the inhibition of cystic fibrosis transmembrane conductance regulator exon 9 splicing. *J. Biol. Chem.* **280**, 37572–37584
34. Romano, M., Buratti, E., Romano, G., Klima, R., Del Bel Belluz, L., Stuani, C., Baralle, F., and Feiguin, F. (2014) Evolutionarily conserved heterogeneous nuclear ribonucleoprotein (hnRNP) A/B proteins functionally interact with human and Drosophila TAR DNA-binding protein 43 (TDP-43). *J. Biol. Chem.* **289**, 7121–7130
35. Budini, M., Romano, V., Quadri, Z., Buratti, E., and Baralle, F. E. (2014) TDP-43 loss of cellular function through aggregation requires additional structural determinants beyond its C-terminal Q/N prion-like domain. *Hum. Mol. Genet.* 10.1093/hmg/ddu415
36. Volkening, K., Leystra-Lantz, C., Yang, W., Jaffee, H., and Strong, M. J. (2009) Tar DNA binding protein of 43 kDa (TDP-43), 14-3-3 proteins and copper/zinc superoxide dismutase (SOD1) interact to modulate NFL mRNA stability. Implications for altered RNA processing in amyotrophic lateral sclerosis (ALS). *Brain Res.* **1305**, 168–182
37. Fiesel, F. C., Voigt, A., Weber, S. S., Van den Haute, C., Waldenmaier, A., Görner, K., Walter, M., Anderson, M. L., Kern, J. V., Rasse, T. M., Schmidt, T., Springer, W., Kirchner, R., Bonin, M., Neumann, M., Baekelandt, V., Alunni-Fabbroni, M., Schulz, J. B., and Kahle, P. J. (2010) Knockdown of transactive response DNA-binding protein (TDP-43) downregulates histone deacetylase 6. *EMBO J.* **29**, 209–221
38. Bose, J. K., Huang, C.-C., and Shen, C.-K. J. (2011) Regulation of autophagy by neuropathological protein TDP-43. *J. Biol. Chem.* **286**, 44441–44448
39. Ayala, Y. M., Misteli, T., and Baralle, F. E. (2008) TDP-43 regulates retinoblastoma protein phosphorylation through the repression of cyclin-dependent kinase 6 expression. *Proc. Natl. Acad. Sci.* **105**, 3785–3789

40. Ayala, Y. M., De Conti, L., Avendaño-Vázquez, S. E., Dhir, A., Romano, M., D'Ambrogio, A., Tollervey, J., Ule, J., Baralle, M., Buratti, E., and Baralle, F. E. (2011) TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J.* **30**, 277–288
41. Buratti, E., De Conti, L., Stuani, C., Romano, M., Baralle, M., and Baralle, F. (2010) Nuclear factor TDP-43 can affect selected microRNA levels. *FEBS J.* **277**, 2268–2281
42. Kawahara, Y., and Mieda-Sato, A. (2012) TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 3347–3352
43. Dewey, C. M., Cenik, B., Sephton, C. F., Dries, D. R., Mayer, P., Good, S. K., Johnson, B. A., Herz, J., and Yu, G. (2011) TDP-43 is directed to stress granules by sorbitol, a novel physiological osmotic and oxidative stressor. *Mol. Cell. Biol.* **31**, 1098–1108
44. McDonald, K. K., Aulas, A., Destroismaisons, L., Pickles, S., Beleac, E., Camu, W., Rouleau, G. A., and Velde, C. V. (2011) TAR DNA-Binding Protein 43 (TDP-43) Regulates Stress Granule Dynamics via Differential Regulation of G3BP and TIA-1. *Hum. Mol. Genet.* 10.1093/hmg/ddr021
45. Zhang, T., Baldie, G., Periz, G., and Wang, J. (2014) RNA-Processing Protein TDP-43 Regulates FOXO-Dependent Protein Quality Control in Stress Response. *PLoS Genet*. **10**, e1004693
46. Stoica, R., De Vos, K. J., Paillusson, S., Mueller, S., Sancho, R. M., Lau, K.-F., Vizcay-Barrena, G., Lin, W.-L., Xu, Y.-F., Lewis, J., Dickson, D. W., Petrucelli, L., Mitchell, J. C., Shaw, C. E., and Miller, C. C. J. (2014) ER–mitochondria associations are regulated by the VAPB–PTPIP51 interaction and are disrupted by ALS/FTD-associated TDP-43. *Nat. Commun.* 10.1038/ncomms4996
47. Ayala, Y. M., Zago, P., D'Ambrogio, A., Xu, Y.-F., Petrucelli, L., Buratti, E., and Baralle, F. E. (2008) Structural determinants of the cellular localization and shuttling of TDP-43. *J. Cell Sci.* **121**, 3778–3785

48. Zhang, Y.-J., Caulfield, T., Xu, Y.-F., Gendron, T. F., Hubbard, J., Stetler, C., Sasaguri, H., Whitelaw, E. C., Cai, S., Lee, W. C., and Petrucelli, L. (2013) The dual functions of the extreme N-terminus of TDP-43 in regulating its biological activity and inclusion formation. *Hum. Mol. Genet.* **22**, 3112–3122
49. Yachdav, G., Kloppmann, E., Kajan, L., Hecht, M., Goldberg, T., Hamp, T., Hönigschmid, P., Schafferhans, A., Roos, M., Bernhofer, M., Richter, L., Ashkenazy, H., Punta, M., Schlessinger, A., Bromberg, Y., Schneider, R., Vriend, G., Sander, C., Ben-Tal, N., and Rost, B. (2014) PredictProtein--an open resource for online prediction of protein structural and functional features. *Nucleic Acids Res.* **42**, W337–343
50. Chang, C., Wu, T.-H., Wu, C.-Y., Chiang, M., Toh, E. K.-W., Hsu, Y.-C., Lin, K.-F., Liao, Y., Huang, T., and Huang, J. J.-T. (2012) The N-terminus of TDP-43 promotes its oligomerization and enhances DNA binding affinity. *Biochem. Biophys. Res. Commun.* **425**, 219–224
51. Qin, H., Lim, L.-Z., Wei, Y., and Song, J. (2014) TDP-43 N terminus encodes a novel ubiquitin-like fold and its unfolded form in equilibrium that can be shifted by binding to ssDNA. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 18619–18624
52. Shiina, Y., Arima, K., Tabunoki, H., and Satoh, J. (2010) TDP-43 Dimerizes in Human Cells in Culture. *Cell. Mol. Neurobiol.* **30**, 641–652
53. Zhang, Y.-J., Xu, Y.-F., Cook, C., Gendron, T. F., Roettges, P., Link, C. D., Lin, W.-L., Tong, J., Castanedes-Casey, M., Ash, P., Gass, J., Rangachari, V., Buratti, E., Baralle, F., Golde, T. E., Dickson, D. W., and Petrucelli, L. (2009) Aberrant cleavage of TDP-43 enhances aggregation and cellular toxicity. *Proc. Natl. Acad. Sci.* **106**, 7607–7612
54. Maris, C., Dominguez, C., and Allain, F. H.-T. (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.* **272**, 2118–2131
55. Mackness, B. C., Tran, M. T., McClain, S. P., Matthews, C. R., and Zitzewitz, J. A. (2014) Folding of the RNA Recognition Motif (RRM) Domains of the Amyotrophic Lateral Sclerosis (ALS)-linked Protein TDP-43 Reveals an Intermediate State. *J. Biol. Chem.* **289**, 8264–8276

56. Kuo, P.-H., Doudeva, L. G., Wang, Y.-T., Shen, C.-K. J., and Yuan, H. S. (2009) Structural insights into TDP-43 in nucleic-acid binding and domain interactions. *Nucleic Acids Res.* **37**, 1799–1808
57. Lukavsky, P. J., Daujotyte, D., Tollervey, J. R., Ule, J., Stuani, C., Buratti, E., Baralle, F. E., Damberger, F. F., and Allain, F. H.-T. (2013) Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat. Struct. Mol. Biol.* **20**, 1443–1449
58. Adam, S. A., Nakagawa, T., Swanson, M. S., Woodruff, T. K., and Dreyfuss, G. (1986) mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. *Mol. Cell. Biol.* **6**, 2932–2943
59. Kuo, P.-H., Chiang, C.-H., Wang, Y.-T., Doudeva, L. G., and 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 Res.* **42**, 4712–4722
60. Bhardwaj, A., Myers, M. P., Buratti, E., and Baralle, F. E. (2013) Characterizing TDP-43 interaction with its RNA targets. *Nucleic Acids Res.* **41**, 5062–5074
61. Ward, J. J., McGuffin, L. J., Bryson, K., Buxton, B. F., and Jones, D. T. (2004) The DISOPRED server for the prediction of protein disorder. *Bioinformatics*. **20**, 2138–2139
62. Xue, B., Dunbrack, R. L., Williams, R. W., Dunker, A. K., and Uversky, V. N. (2010) PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta*. **1804**, 996–1010
63. Jiang, L.-L., Che, M.-X., Zhao, J., Zhou, C.-J., Xie, M.-Y., Li, H.-Y., He, J.-H., and Hu, H.-Y. (2013) Structural Transformation of the Amyloidogenic Core Region of TDP-43 Protein Initiates Its Aggregation and Cytoplasmic Inclusion. *J. Biol. Chem.* **288**, 19614–19624
64. Budini, M., Buratti, E., Stuani, C., Guarnaccia, C., Romano, V., De Conti, L., and Baralle, F. E. (2012) Cellular model of TAR DNA-binding protein 43 (TDP-

43) aggregation based on its C-terminal Gln/Asn-rich region. *J. Biol. Chem.* **287**, 7512–7525

65. D'Ambrogio, A., Buratti, E., Stuani, C., Guarnaccia, C., Romano, M., Ayala, Y. M., and Baralle, F. E. (2009) Functional mapping of the interaction between TDP-43 and hnRNP A2 in vivo. *Nucleic Acids Res.* **37**, 4116–4126
66. Taylor, J. P., Tanaka, F., Robitschek, J., Sandoval, C. M., Taye, A., Markovic-Plese, S., and Fischbeck, K. H. (2003) Aggresomes protect cells by enhancing the degradation of toxic polyglutamine-containing protein. *Hum. Mol. Genet.* **12**, 749–757
67. Chiti, F., and Dobson, C. M. (2006) Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **75**, 333–366
68. Mompeán, M., Buratti, E., Guarnaccia, C., Brito, R. M. M., Chakrabartty, A., Baralle, F. E., and Laurents, D. V. (2014) Structural characterization of the minimal segment of TDP-43 competent for aggregation. *Arch. Biochem. Biophys.* **545**, 53–62
69. Wang, I.-F., Chang, H.-Y., Hou, S.-C., Liou, G.-G., Way, T.-D., and James Shen, C.-K. (2012) The self-interaction of native TDP-43 C terminus inhibits its degradation and contributes to early proteinopathies. *Nat. Commun.* **3**, 766
70. Da Cruz, S., and Cleveland, D. W. (2011) Understanding the role of TDP-43 and FUS/TLS in ALS and beyond. *Curr. Opin. Neurobiol.* **21**, 904–919
71. Nehls, J., Koppensteiner, H., Brack-Werner, R., Floss, T., and Schindler, M. (2014) HIV-1 replication in human immune cells is independent of TAR DNA binding protein 43 (TDP-43) expression. *PloS One*. **9**, e105478
72. Ayala, Y. M., Pagani, F., and Baralle, F. E. (2006) TDP43 depletion rescues aberrant CFTR exon 9 skipping. *FEBS Lett.* **580**, 1339–1344
73. Zhang, Y.-J., Xu, Y., Dickey, C. A., Buratti, E., Baralle, F., Bailey, R., Pickering-Brown, S., Dickson, D., and Petrucelli, L. (2007) Progranulin mediates caspase-dependent cleavage of TAR DNA binding protein-43. *J. Neurosci. Off. J. Soc. Neurosci.* **27**, 10530–10534
74. De Marco, G., Lomartire, A., Mandili, G., Lupino, E., Buccinnà, B., Ramondetti, C., Moglia, C., Novelli, F., Piccinini, M., Mostert, M., Rinaudo, M.

T., Chiò, A., and Calvo, A. (2014) Reduced cellular Ca2+ availability enhances TDP-43 cleavage by apoptotic caspases. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1843**, 725–734

75. Josephs, K. A., Whitwell, J. L., Weigand, S. D., Murray, M. E., Tosakulwong, N., Liesinger, A. M., Petrucelli, L., Senjem, M. L., Knopman, D. S., Boeve, B. F., Ivnik, R. J., Smith, G. E., Jr, C. R. J., Parisi, J. E., Petersen, R. C., and Dickson, D. W. (2014) TDP-43 is a key player in the clinical features associated with Alzheimer's disease. *Acta Neuropathol. (Berl.)*. **127**, 811–824
76. Kabashi, E., Valdmanis, P. N., Dion, P., Spiegelman, D., McConkey, B. J., Vande Velde, C., Bouchard, J.-P., Lacomblez, L., Pochigaeva, K., Salachas, F., Pradat, P.-F., Camu, W., Meininger, V., Dupre, N., and Rouleau, G. A. (2008) TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nat. Genet.* **40**, 572–574
77. Sreedharan, J., Blair, I. P., Tripathi, V. B., Hu, X., Vance, C., Rogelj, B., Ackerley, S., Durnall, J. C., Williams, K. L., Buratti, E., Baralle, F., de Belleroche, J., Mitchell, J. D., Leigh, P. N., Al-Chalabi, A., Miller, C. C., Nicholson, G., and Shaw, C. E. (2008) TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science*. **319**, 1668–1672
78. Corrado, L., Ratti, A., Gellera, C., Buratti, E., Castellotti, B., Carlomagno, Y., Ticozzi, N., Mazzini, L., Testa, L., Taroni, F., Baralle, F. E., Silani, V., and D'Alfonso, S. (2009) High frequency of TARDBP gene mutations in Italian patients with amyotrophic lateral sclerosis. *Hum. Mutat.* **30**, 688–694
79. Hasegawa, M., Arai, T., Nonaka, T., Kametani, F., Yoshida, M., Hashizume, Y., Beach, T. G., Buratti, E., Baralle, F., Morita, M., Nakano, I., Oda, T., Tsuchiya, K., and Akiyama, H. (2008) Phosphorylated TDP-43 in frontotemporal lobar degeneration and ALS. *Ann. Neurol.* **64**, 60–70
80. Kametani, F., Nonaka, T., Suzuki, T., Arai, T., Dohmae, N., Akiyama, H., and Hasegawa, M. (2009) Identification of casein kinase-1 phosphorylation sites on TDP-43. *Biochem. Biophys. Res. Commun.* **382**, 405–409

81. Liachko, N. F., McMillan, P. J., Guthrie, C. R., Bird, T. D., Leverenz, J. B., and Kraemer, B. C. (2013) CDC7 inhibition blocks pathological TDP-43 phosphorylation and neurodegeneration. *Ann. Neurol.* **74**, 39–52
82. Inukai, Y., Nonaka, T., Arai, T., Yoshida, M., Hashizume, Y., Beach, T. G., Buratti, E., Baralle, F. E., Akiyama, H., Hisanaga, S., and Hasegawa, M. (2008) Abnormal phosphorylation of Ser409/410 of TDP-43 in FTLD-U and ALS. *FEBS Lett.* **582**, 2899–2904
83. Kim, K. Y., Lee, H.-W., Shim, Y., Mook-Jung, I., Jeon, G. S., and Sung, J.-J. (2015) A phosphomimetic mutant TDP-43 (S409/410E) induces Drosha instability and cytotoxicity in Neuro 2A cells. *Biochem. Biophys. Res. Commun.* **464**, 236–243
84. Zhang, Y.-J., Gendron, T. F., Xu, Y.-F., Ko, L.-W., Yen, S.-H., and Petrucelli, L. (2010) Phosphorylation regulates proteasomal-mediated degradation and solubility of TAR DNA binding protein-43 C-terminal fragments. *Mol. Neurodegener.* **5**, 33
85. Brady, O. A., Meng, P., Zheng, Y., Mao, Y., and Hu, F. (2011) Regulation of TDP-43 aggregation by phosphorylation and p62/SQSTM1. *J. Neurochem.* **116**, 248–259
86. Li, H.-Y., Yeh, P.-A., Chiu, H.-C., Tang, C.-Y., and Tu, B. P. (2011) Hyperphosphorylation as a Defense Mechanism to Reduce TDP-43 Aggregation. *PLoS ONE*. 10.1371/journal.pone.0023075
87. Dormann, D., Capell, A., Carlson, A. M., Shankaran, S. S., Rodde, R., Neumann, M., Kremmer, E., Matsuwaki, T., Yamanouchi, K., Nishihara, M., and Haass, C. (2009) Proteolytic processing of TAR DNA binding protein-43 by caspases produces C-terminal fragments with disease defining properties independent of progranulin. *J. Neurochem.* **110**, 1082–1094
88. Liu, Y., Duan, W., Guo, Y., Li, Z., Han, H., Zhang, S., Yuan, P., and Li, C. (2014) A new cellular model of pathological TDP-43: The neurotoxicity of stably expressed CTF25 of TDP-43 depends on the proteasome. *Neuroscience*. **281C**, 88–98

89. Nonaka, T., Arai, T., Buratti, E., Baralle, F. E., Akiyama, H., and Hasegawa, M. (2009) Phosphorylated and ubiquitinated TDP-43 pathological inclusions in ALS and FTLD-U are recapitulated in SH-SY5Y cells. *FEBS Lett.* **583**, 394–400
90. Tompa, P. (2002) Intrinsically unstructured proteins. *Trends Biochem. Sci.* **27**, 527–533
91. Uversky, V. N. (2013) The most important thing is the tail: Multitudinous functionalities of intrinsically disordered protein termini. *FEBS Lett.* **587**, 1891–1901
92. Pufall, M. A., and Graves, B. J. (2002) Autoinhibitory domains: modular effectors of cellular regulation. *Annu. Rev. Cell Dev. Biol.* **18**, 421–462
93. Trudeau, T., Nassar, R., Cumberworth, A., Wong, E. T. C., Woollard, G., and Gsponer, J. (2013) Structure and Intrinsic Disorder in Protein Autoinhibition. *Structure*. **21**, 332–341
94. Jayaraman, B., and Nicholson, L. K. (2007) Thermodynamic Dissection of the Ezrin FERM/CERMAD Interface†. *Biochemistry (Mosc.)*. **46**, 12174–12189
95. Christie, D. A., Lemke, C. D., Elias, I. M., Chau, L. A., Kirchhof, M. G., Li, B., Ball, E. H., Dunn, S. D., Hatch, G. M., and Madrenas, J. (2011) Stomatin-Like Protein 2 Binds Cardiolipin and Regulates Mitochondrial Biogenesis and Function. *Mol. Cell. Biol.* **31**, 3845–3856
96. Schägger, H. (2006) Tricine-SDS-PAGE : Nature Protocols. *Nat Protoc.* **1**, 16–22
97. Chung, C. (2014) Structure and Interactions of the C-terminal Domain of TDP-43. *Honors Biochem. Thesis*
98. Vasanthakumar, T. (2014) Characterization of the C-terminal domain of TDP-43. *Honors Biochem. Thesis*
99. Dunn, S. D. (1986) Effects of the modification of transfer buffer composition and the renaturation of proteins in gels on the recognition of proteins on western blots by monoclonal antibodies. *Anal. Biochem.* **157**, 144–153
100. Walker, J. M. (ed.) (1996) The Iodogen Method for Radiolabeling Protein - Springer, Humana Press, [online]

http://link.springer.com/protocol/10.1007%2F978-1-60327-259-9_115 (Accessed August 20, 2015)

101. Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J., and Schubert, D. (2002) Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. *Biophys. J.* **82**, 1096–1111

102. Baynes, B. M., Wang, D. I. C., and Trout, B. L. (2005) Role of Arginine in the Stabilization of Proteins against Aggregation†. *Biochemistry (Mosc.)*. **44**, 4919–4925

103. Jao, S.-C., Ma, K., Talafous, J., Orlando, R., and Zagorski, M. G. (1997) Trifluoroacetic acid pretreatment reproducibly disaggregates the amyloid β-peptide. *Amyloid*. **4**, 240–252

104. Jadwin, J. A., Mayer, B. J., and Machida, K. (2015) Detection and quantification of protein-protein interactions by far-Western blotting. *Methods Mol. Biol. Clifton NJ*. **1312**, 379–398

105. Bondos, S. E., and Bicknell, A. (2003) Detection and prevention of protein aggregation before, during, and after purification. *Anal. Biochem.* **316**, 223–231

106. Dam, J., and Schuck, P. (2005) Sedimentation Velocity Analysis of Heterogeneous Protein-Protein Interactions: Sedimentation Coefficient Distributions c(s) and Asymptotic Boundary Profiles from Gilbert-Jenkins Theory. *Biophys. J.* **89**, 651–666

107. Dogan, J., Gianni, S., and Jemth, P. (2014) The binding mechanisms of intrinsically disordered proteins. *Phys. Chem. Chem. Phys.* **16**, 6323–6331

108. Arthur, K. K., Gabrielson, J. P., Kendrick, B. S., and Stoner, M. R. (2009) Detection of protein aggregates by sedimentation velocity analytical ultracentrifugation (SV-AUC): Sources of variability and their relative importance. *J. Pharm. Sci.* **98**, 3522–3539

109. Boulikas, T. (1992) Nuclear localization signals (NLS). *Crit. Rev. Eukaryot. Gene Expr.* **3**, 193–227

110. Fuentealba, R. A., Udan, M., Bell, S., Wegorzewska, I., Shao, J., Diamond, M. I., Weihl, C. C., and Baloh, R. H. (2010) Interaction with Polyglutamine Aggregates Reveals a Q/N-rich Domain in TDP-43. *J. Biol. Chem.* **285**, 26304–26314

111. Weihl, C. C., Temiz, P., Miller, S. E., Watts, G., Smith, C., Forman, M., Hanson, P. I., Kimonis, V., and Pestronk, A. (2008) TDP-43 accumulation in inclusion body myopathy muscle suggests a common pathogenic mechanism with frontotemporal dementia. *J. Neurol. Neurosurg. Psychiatry*. **79**, 1186–1189

Curriculum Vitae

Post-Secondary Education

2013-2015 **Master of Science Candidate, Biochemistry**
The University of Western Ontario, London, Ontario, Canada
Title: “Interdomain interactions of the transactive response DNA binding protein 43 kDa (TDP-43)”
Co-supervisors: Dr. Stanley D. Dunn and Dr. Michael J. Strong

2008-2013 **Honors Specialization in Clinical Biochemistry**
Bachelor of Medical Science
The University of Western Ontario, London, Ontario, Canada

Scholarships and Awards

2013-2015 **Western Graduate Research Scholarship**
- Awarded to eligible full-time graduate students

2014 **Third Place Poster Presenter**, ALS Society of Canada 10th Annual Research Forum, Toronto, Ontario
- Poster entitled: “Inter-domain interactions of the RNA-binding protein, TDP-43”
- Awarded based on a judging panel

2013 **First Place Poster Presenter**, Fourth International Research Workshop on Frontotemporal Dementia in ALS
- Poster entitled: “Inter-domain interactions of the RNA-binding protein, TDP-43”
- Awarded based on a judging panel

2008 **Western Scholarship of Excellence**
- Awarded for achieving an above 90% average in final year of secondary school

Research Experience

2013-2015 **Research-based Master's Thesis**, Department of Biochemistry
The University of Western Ontario, London, ON, Canada
Co-supervisors: Dr. Stanley D. Dunn and Dr. Michael J. Strong

- Identified an interaction between two domains of the protein TDP-43
- Studied protein interactions through biophysical techniques including analytical ultracentrifugation and Far Western blotting
- Helped with planning colocalization studies
- Purified recombinant proteins via column chromatography
- Planned and cloned various TDP-43 expression plasmids through molecular biology techniques
- Prepared presentations/posters of my research for meetings and conferences

2012-2013 **Honors Undergraduate Thesis**, Department of Biochemistry
The University of Western Ontario, London, ON, Canada
Supervisor: Dr. Stanley D. Dunn

- Found preliminary evidence for an interaction between two domains of TDP-43
- Planned and cloned various TDP-43 expression plasmids through molecular biology techniques
- Purified recombinant proteins via column chromatography
- Wrote a final thesis report
- Prepared presentations and a poster for symposia within the department

2011-2012 **Science Internship Program**, LANXESS Inc., London ON, Canada

- Worked in Product Research Group, on TPV Team
- Planned and conducted lab experiments modifying butyl rubber
- Analyzed results using IR, NMR, titration and other rubber industry equipment
- Wrote multiple technical reports detailing work performed, result analysis and future recommendations
- Took part in weekly team meetings, preparing presentations detailing resent results and paths forward

Presentations and Abstracts

2015 **Poster and Podium Presentation**, Fifth International Research Workshop on Frontotemporal Dementia in ALS, London, Ontario, Canada

- Title: Inter-domain interactions of the RNA-binding protein, TDP-43
- Authors: Karen Dunkerley[1], Yumin Bi[1], Kathryn Volkening[2,3], Michael J. Strong[2,3] and Stanley D. Dunn[1]

2015 **Poster Presentation and Abstract Submission**, London Health Research Day 2015, London, Ontario, Canada

- Title: Inter-domain interactions of the RNA-binding protein, TDP-43
- Authors: Karen Dunkerley, Yumin Bi, Kathryn Volkening, Michael J. Strong and Stanley D. Dunn

2015 **Presentation**, Graduate Student Seminars, Department of Biochemistry

- Title: Characterization of interdomain interactions in the RNA-binding protein, TDP-43
- Author: Karen Dunkerley

2014 **Poster Presentation**, ALS Society of Canada 10th Annual Research Forum, Toronto, Ontario

- Title: Inter-domain interactions of the RNA-binding protein, TDP-43
- Authors: Karen Dunkerley, Yumin Bi, Kathryn Volkening, Michael J. Strong and Stanley D. Dunn

2014 **Poster Presentation and Abstract Submission**, London Health Research Day 2014, London, Ontario, Canada

- Title: Inter-domain interactions of the RNA-binding protein, TDP-43
- Authors: Karen Dunkerley, Yumin Bi, Kathryn Volkening, Michael J. Strong and Stanley D. Dunn

2013 **Poster and Podium Presentation**, Fourth International Research Workshop on Frontotemporal Dementia in ALS, London, Ontario, Canada

- Title: Characterization of TDP-43 Domain Interactions
- Authors: Karen Dunkerley, Yumin Bi, Kathryn Volkening, Michael J. Strong and Stanley D. Dunn

2013 **Poster Presentation**, Harold B. Stewart Memorial Lecture and Research Showcase, London, Ontario, Canada

- Title: Characterization of TDP-43 Domain Interactions

- Authors: Karen Dunkerley, Yumin Bi and Stanley D. Dunn

Teaching Experience

2013-2014 **Teaching Assistant**, Department of Biochemistry, The University of Western Ontario, London, Ontario, Canada
Courses: Biochemistry 3381A, 4420A

- Planned and conducted tutorials
- Held office hours
- Graded student assignments and exams
- Proctored exams
- Reported student progress/concerns to course instructors