Matrin3 Misfolding in Amyotrophic Lateral Sclerosis

Ahmed Salem, The University of Western Ontario

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Anatomy and Cell Biology

© Ahmed Salem 2021

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Recommended Citation
https://ir.lib.uwo.ca/etd/7884

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.
Abstract

Amyotrophic Lateral Sclerosis (ALS) is a neurodegenerative disorder characterized by degeneration of upper and lower motor neurons in the brain and the spinal cord, respectively. ALS is associated with protein misfolding and inclusion formation of several RNA binding proteins, such as TAR DNA binding protein (TDP-43) and Fused in Sarcoma (FUS). Matrin3 is a nuclear DNA and RNA binding protein and mutations in the gene encoding Matrin3 have been identified as a cause of familial ALS (fALS). Matrin3 is an intrinsically disordered RNA binding protein with numerous phosphorylation sites. This study attempts to understand the role of the intrinsically disordered regions and protein phosphorylation on Matrin3 misfolding and mis-localization using a novel yeast model, mammalian neuronal cells, and post-mortem human neuronal tissue from the spinal cords of ALS patients. We propose that the intrinsically amino terminal disordered region and protein phosphorylation drive Matrin3 misfolding in ALS.

Keywords

Neurodegenerative diseases, Amyotrophic Lateral Sclerosis (ALS), protein quality control, protein misfolding, Matrin3, protein phosphorylation, intrinsically disordered domains
Summary for lay audience

Amyotrophic lateral sclerosis (ALS) is a fatal disease that affects a specific group of nerve cells inside our brains and spinal cords called motor neurons. Motor neurons are cells that control the movement of our muscles. Scientists have discovered a link between ALS and defects in proteins. Proteins are the molecules that carry out most of the functions inside our cells. For proteins to be able to perform their function, they have to attain a specific shape through a well-regulated process called protein folding. In ALS, some proteins lose their proper shape and acquire a different one which is a process called protein misfolding. These misfolded proteins can accumulate inside the neurons, become toxic and prevent the neurons from functioning properly. There are several factors that can regulate the process of folding, such as protein phosphorylation and intrinsically disordered regions. Protein phosphorylation is the addition of phosphate groups at specific sites in the protein to help its proper folding. There is evidence that there are changes in protein phosphorylation in ALS patients. Intrinsically disordered regions are parts in some proteins that do not have a specific shape due to their high flexibility. This flexibility helps the protein perform its function and bind to many other proteins. The downside to this flexibility is that it renders the protein prone to misfolding. Scientists have found that many proteins misfolding in ALS have intrinsically disordered regions.

In this thesis we explore the misfolding of Matrin3, a nuclear protein that becomes abnormally altered in ALS individuals. Matrin3 has many phosphorylation sites and two large intrinsically disordered regions, so we decided to study how these factors contribute to Matrin3 misfolding in ALS. We began our studies in a yeast model, which allows us to isolate Matrin3 and test its toxicity and localization. Interestingly, we discovered that protein phosphorylation and one of the intrinsically disordered regions plays a major role in Matrin3 toxicity. Thus, we speculate that altered protein phosphorylation and intrinsically disordered regions drive Matrin3 misfolding in ALS patients leading to motor neuron dysfunction.
Abstract ........................................................................................................................................... 1
Summary for lay audience ................................................................................................................... II
Co-Authorship Statement .................................................................................................................. VII
Acknowledgments ............................................................................................................................. VIII
Table of Contents ............................................................................................................................... IX
List of Tables ........................................................................................................................................ IX
List of Figures ........................................................................................................................................ IX
List of Appendices .............................................................................................................................. X
List of Abbreviations .......................................................................................................................... XI

Chapter 1 ........................................................................................................................................... 1

1 Introduction................................................................................................................................. 1
   1.1 Neurodegenerative diseases ................................................................................................. 1
   1.1.1 Amyotrophic Lateral Sclerosis ................................................................................. 1
         1.1.1.1 Amyotrophic Lateral Sclerosis (ALS)-Associated protein misfolding .......... 4
              1.1.1.1.1 Transactive Response DNA Binding Protein 43 ... 6
              1.1.1.2 Matrin3 ......................................................................................................... 10
   1.2 Protein folding and misfolding inside the cell ................................................................. 15
   1.3 Mechanisms of protein misfolding in Amyotrophic Lateral Sclerosis ................. 17
         1.3.1 Liquid-liquid phase separation ............................................................................. 17
              1.3.1.1 Intrinsically disordered proteins .............................................................. 19
              1.3.1.2 Prion like domains .................................................................................. 20
         1.3.2 Posttranslational modification .............................................................................. 21
              1.3.2.1 Protein phosphorylation ........................................................................ 21
              1.3.2.2 Matrin3 phosphorylation ...................................................................... 22
   1.4 Yeast as a Model Organism to study protein misfolding .................................. 23
   1.5 Mammalian neuronal cell model: SN56 cells ...................................................... 24
   1.6 Rationale ......................................................................................................................... 24
Chapter 4 ......................................................................................................................... 76

4 Discussion .................................................................................................................... 76

  4.1 Aberrant Matrin3 localization in motor neurons of the anterior horn in the spinal cords of ALS patients ............................................................ 76
  4.2 WT Matrin3 yeast model ................................................................................. 78
  4.3 Matrin3 ALS associated variants ................................................................. 79
  4.4 Matrin3 phosphorylation ............................................................................. 80
  4.5 The intrinsically disordered domains of Matrin3 .................................... 83
  4.6 Overall conclusions and significance ....................................................... 85

References ............................................................................................................... 88

Appendices ............................................................................................................. 100

Curriculum Vitae .................................................................................................... 106
Co-Authorship Statement

Figure 11: Sonja Di Gregorio helped establishing Matrin3 immunohistochemistry staining protocol.

Figure 12: Sali Farhan helped finding all Matrin3 ALS associated mutations in databases.

Figure 15: Sonja Di Gregorio helped establishing the Matrin3 yeast model.

Figure 20: Sonja Di Gregorio provided Matrin3 F115C fluorescence microscopy images.

Figure 14: Carter Wilson provided Matrin3 phosphorylation in-silico analysis.

All other experiments and data analysis were conducted by myself.
Acknowledgements

First, I would like to give a special thanks to my supervisor, Dr. Martin Duennwald. Thank you for your continued support, direction and understanding. Without your guidance, I wouldn’t be the scientist I am today! You created a positive collaborative atmosphere and provided me with such an amazing opportunity. You have been such a great mentor and have taught me more than I can give you credit for here. I would also like to thank my advisory committee, Dr. Shawn Whitehead, Dr. Patrick Lajoie, and Dr. Qi Zhang, for your encouragement and direction.

To the past and present members of the Duennwald laboratory, Dr. Mohammad Esmaeili, Dr. Sonja DiGregorio, Dr. Vy Ngo, Mihai Dumbrava, Andrey Petropavllovskiy, Meaghan Stoltz, Raga Alshwari, Donovan McDonald, Brendan Charles, Christian Lotz, and Ayushi Bhatt, your friendship and technical support were imperative to this project. You are the best research team any student could dream of having.

Sonja Di Gregorio, thank you for taking me under your wing from day one. You are a such a great scientist and mentor. Mohammad, thank you for providing me with your experienced direction and putting up with my questions all the time. Vy, thank you for your support, guidance and advice. Thank you, Meaghan. You are a great friend, scientist and the best desk mate I could ever ask for. Donovan and Brendan, thank you for the laughs, and continuous support. Christian, we had so much fun playing squash. I hope we can play again one day.

Thank you to the members of the Lajoie laboratory, Dr. Patrick Lajoie, Dr. Sarah Chadwick Maram Albakri, Marina Sertsis, Afrin Bari and Malissa Fernando for always providing valuable advice. Without your help this project wouldn’t be as successful.
List of Tables

Table 1: Yeast growth media
Table 2: E. coli growth media
Table 3: Mammalian cell growth media
Table 4: Plasmids
Table 5: Antibodies
Table 6: Additional key reagents and chemicals
Table 7: Human spinal cord tissue samples
Table 8: Summary of the immunohistochemical staining results

List of Figures:

Figure 1: Amyotrophic lateral sclerosis (ALS) affects upper and lower motor neurons
Figure 2: ALS associated genes
Figure 3: A schematic diagram of the domain structure of TDP-43 and its mis-localization and aggregation in ALS.
Figure 4: A schematic diagram of the domain structure of Matrin3
Figure 5: The interaction network of Matrin3
Figure 6: ALS-causing mutations in MATR3
Figure 7: Cellular protein quality control and its role in protein folding
Figure 8: Aberrant liquid-liquid phase separation can lead to the formation of protein aggregates
Figure 9: Localization of Matrin3 in the anterior horn of the human spinal cord of unaffected controls, sALS and fALS patients.
Figure 10: MATR3 fALS mutations
Figure 11: Matrin-3 post-translational modifications
Figure 12: In-silico analyses predict Matrin3 to contain intrinsically disordered domains
Figure 13: Expression of human full-length WT Matrin3 causes a growth defect in the yeast model
Figure 14: High and low expression of WT Matrin3 in the yeast model
Figure 15: Matrin3 forms nuclear and cytosolic inclusions in the yeast model
Figure 16: Expressing ALS-associated Matrin3 variants in the yeast model
Figure 17: Matrin3 ALS-associated mutations do not change Matrin3 protein levels in the yeast model.
Figure 18: ALS-associated mutations does not affect Matrin3 localization in the yeast model.
Figure 19: Matrin3 S188A and S610D mutations significantly reduced Matrin3 growth defect in the yeast model.
Figure 20: S188A mutation leads to Matrin3 degradation in the yeast model.
Figure 21: S188A mutation leads to low steady-state levels of Matrin3 and S610D inhibits Matrin3 inclusion formation in the yeast model.
Figure 22: Matrin3 phospho-variants localization in mammalian cells
Figure 23: S610D, T622A and T622D variants reduces SN56 cell viability.
Figure 24: Matrin3 amino-terminal truncated fragments caused a significant growth defect in the yeast model.
Figure 25: Matrin3 amino-terminal 399 aa truncated fragment shows a significant increase in the steady-state protein levels compared to WT Matrin3.
Figure 26: F115C mutation changes the localization Matrin3 amino-terminal 399 aa truncated fragment in yeast.
Figure 27: amino-terminal 399 aa truncated fragment leads to the formation of nuclear and cytoplasmic inclusions.
Figure 28: Overexpression of the amino-terminal 399 aa truncated fragment significantly reduces SN56 cells viability.
Figure 29: Proposed mechanism of Matrin3 Misfolding in ALS.

List of Appendices

Appendix A: Supplementary figures
List of Abbreviations

2xYT 2x yeast extract tryptone
AD Alzheimer’s disease
ADP Adenosine diphosphate
ALS Amyotrophic lateral sclerosis
Amp Ampicillin
ATP Adenosine triphosphate
BCA Bicinchoninic acid assay
BSA Bovine serum albumin
C9orf72 Chromosome 9 open reading frame 72
CNS Central nervous system
DAB Diaminobenzidine
DMEM Dulbecco’s modified eagle medium
DMSO Dimethyl Sulfoxide
ECL Enhanced chemiluminescence
fALS Familial amyotrophic lateral sclerosis
FBS Fetal bovine serum
FTLD Frontotemporal lobar dementia
FUS Fused in sarcoma
g/L grams per litre
GFP Green fluorescent protein
H2O2 Hydrogen peroxide
HD Huntington’s disease
Htt Huntingtin protein
IF Immunofluorescence
IHC Immunohistochemistry
Kan Kanamycin
kDa Kilodalton
KLD reaction Kinase, Ligase, and Dpn1 reaction
LB Lysogeny Broth
Li Lithium
MATR3 Matrin3 gene
mRNA Messenger ribonucleic acid
NaOH Sodium Hydroxide
NES Nuclear export sequence
ng Nanogram
NLS Nuclear localization sequence
PBST Phosphate-buffered saline/tween
PD Parkinson’s disease
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEG</td>
<td>Polyethylene glycol</td>
</tr>
<tr>
<td>PGK-1</td>
<td>Phosphoglycerate kinase-1</td>
</tr>
<tr>
<td>polyQ</td>
<td>polyglutamine</td>
</tr>
<tr>
<td>PrLD</td>
<td>Prion like domain</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene fluoride or polyvinylidene difluoride</td>
</tr>
<tr>
<td>RRM</td>
<td>RNA-binding domain</td>
</tr>
<tr>
<td>sALS</td>
<td>sporadic amyotrophic lateral sclerosis</td>
</tr>
<tr>
<td>SD</td>
<td>Selective dextrose</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SGD</td>
<td>Saccharomyces genome database</td>
</tr>
<tr>
<td>SOD1</td>
<td>copper-zinc superoxide dismutase-1</td>
</tr>
<tr>
<td>TDP-43</td>
<td>Transactive response DNA binding protein 43</td>
</tr>
<tr>
<td>TE</td>
<td>Tris-EDTA</td>
</tr>
<tr>
<td>v/v</td>
<td>volume/volume</td>
</tr>
<tr>
<td>WT</td>
<td>Wild type</td>
</tr>
<tr>
<td>YFP</td>
<td>Yellow fluorescent protein</td>
</tr>
<tr>
<td>YPD</td>
<td>Yeast extract peptone dextrose</td>
</tr>
<tr>
<td>ZnF</td>
<td>Zinc-finger domain</td>
</tr>
<tr>
<td>µg</td>
<td>microgram</td>
</tr>
<tr>
<td>µL</td>
<td>microlitre</td>
</tr>
<tr>
<td>µm</td>
<td>micromolar</td>
</tr>
</tbody>
</table>
1. **Chapter 1**

1.1. **Neurodegenerative diseases**

Neurodegeneration is a broad term encompassing multiple disorders, such as Alzheimer’s disease (AD), Parkinson’s disease (PD), Amyotrophic Lateral Sclerosis (ALS), and Frontotemporal dementia (FTD). These disorders manifest in an age-dependent manner and are characterized by progressive loss of neurons. A major hallmark of neurodegenerative disorders is protein misfolding and inclusion formation, exemplified by Aβ plaques in AD, Lewy bodies in PD and TDP-43 aggregates in ALS.\(^1\)\(^2\) My thesis focuses on describing specific aspects of ALS pathology and explores underlying mechanisms.

1.1.1. **Amyotrophic lateral sclerosis**

Amyotrophic lateral sclerosis (ALS), also known as Lou Gehrig’s disease, belongs to a specific category of neurodegenerative diseases called motor neuron diseases, which are characterized by the loss of upper motor neurons (Betz cells) in the primary motor cortex and lower motor neurons in the anterior horn cells of the spinal cord (Figure 1).\(^3\)\(^4\) The onset of ALS is age-dependent with the majority of patients being diagnosed in the sixth decade of life.\(^5\) ALS affects approximately two in 100,000 individuals with higher prominence in males compared to females and an increased risk for military veterans and athletes.\(^6\)

ALS is characterized clinically by progressive muscle weakness, paralysis, and respiratory failure, which is the major cause of death in ALS patients.\(^5\)\(^6\) ALS can be classified based on its heritability into sporadic ALS (sALS), which means there is no family history of the disease. sALS represents approximately 90% of the ALS cases. Familial ALS (fALS), which means there is a family history of the disease, accounts for the remaining 10% of patients.\(^6\)\(^7\)\(^8\) The onset of ALS varies between patients; some patients present with limb onset, which means symptoms affect the upper and lower limbs first, whereas other patients present with bulbar onset, which means patients first notice difficulty in speaking and swallowing. The prognosis of the disease also differs between patients depending on multiple factors, such as age of diagnosis, gender, onset pattern (limb or bulbar),
type of ALS (sporadic or familial), the mutated gene, and even different mutations in the same gene.\textsuperscript{9}
Currently, there is no effective treatment for ALS. Riluzole and Radicava are the only two FDA approved medications for ALS, however studies showed that they can only delay the onset of ventilator-dependence and increase the survival of patients by two to three months.\textsuperscript{10} Similar to many other neurodegenerative disorders, such as Alzheimer’s disease and Parkinson’s disease, ALS is characterized by a defect in protein processing resulting in protein misfolding, mis-localization and inclusion formation in motor neurons.\textsuperscript{11}
Figure 1: Amyotrophic lateral sclerosis (ALS) affects upper and lower motor neurons

A Diagram showing upper motor neurons (UMNs) in the primary motor cortex and lower motor neurons (LMNs) in the anterior horn cells of spinal cord. B Diagram showing progressive muscle atrophy due to degeneration of motor neurons in ALS.\textsuperscript{11,12}
1.1.1.1. **ALS-associated protein misfolding**

Protein misfolding and mis-localization in the form of protein inclusions in motor neurons is a major pathologic hallmark of ALS and many different proteins have been identified to misfold in motor neurons of ALS patients. The first gene mutation identified as a cause of fALS affects the gene encoding the protein Superoxide dismutase one (SOD1) (Figure 2). SOD1 mutations represent 20% of all fALS cases and 1% of all sALS cases.\(^{10,13}\) SOD1 ALS-associated variants destabilize native SOD1 and lead to the formation of misfolded aggregates and degradation of SOD1.\(^{14}\) Furthermore, SOD1 variants cause the misfolding of wild-type (WT) SOD1, thus inducing the aggregation of WT SOD1 and the propagation of SOD1 misfolding in a prion-like manner.\(^{15,16}\)

Subsequently, several gene mutations have been identified as a cause of ALS, such as in the gene encoding the TAR DNA-binding protein 43 (TDP43), C9orf72, which is the most frequent ALS causing gene, Fused in Sarcoma (FUS), and Matrin3 (Figure 2). All these proteins show misfolding and mis-localization not only in fALS patients, but they are also characterised in the majority of sALS cases and fALS cases caused by mutations in different genes.\(^{16,17,18}\)
Figure 2: ALS associated genes.

Diagram showing the timeline of the discovery of several ALS causing genes. The size of the circle represents ALS gene frequency. Matrin3 which is the focus of my project in placed in the red box. Modified after Kirby et al. 19
1.1.1.1. TAR DNA binding protein (TDP-43)

Many RNA binding proteins associated with ALS contain prion like domains (PrLDs), including FUS, TDP-43, and TAF15. PrLDs plays a critical role in the normal function of these proteins as they mediate protein–protein interactions and the formation of stress granules.20,21 Transactive response DNA binding protein 43 (TDP-43) is a 414 amino acid, 43 kDa DNA and RNA binding protein.22 TDP-43 is a ubiquitously expressed nuclear protein that consists of two RNA recognition motifs (RRM1 and RRM2), a nuclear localization signal (NLS), nuclear export signal (NES), and a carboxy-terminal PrLD, which mediates multiple protein-protein interactions (Figure 3, A).22,23

TDP-43 binds DNA and RNA and is involved in RNA regulation and protein synthesis in the cell, such as binding, stabilizing, and alternative splicing of mRNA.24 TDP-43 plays a role in other cellular processes, such as the transport of RNA granules between the cell body and dendrites of neurons and the formation of cytoplasmic stress granules, indicating its role under cellular stress conditions, such as oxidative stress.25,26

TDP-43 proteinopathy represented by protein misfolding and inclusion formation of TDP-43 is considered a major pathologic hallmark not only in ALS patients, but also in patients suffering from other neurodegenerative diseases, such as AD and PD.27 TDP-43 misfolds and forms aggregates in 95% of all ALS cases.27 These aggregates are composed of ubiquitinated, hyperphosphorylated and truncated fragments of TDP-43 containing the carboxy-terminal of the protein (Figure 3, C).7,12,28 Even though protein aggregates are major hallmarks for neurodegenerative diseases, their role in the pathogenesis of these disorders is controversial. It has long been thought that these aggregates cause neurotoxicity, however, in the recent few years, some evidence has emerged indicating that these aggregates are beneficial to neurons by sequestering highly toxic oligomers.29

Over 50 mutations in the gene encoding TDP-43 have been identified as a cause of fALS most of which are located in the carboxy-terminal PrLD, indicating a possible role of the PrLD on TDP-43 misfolding (Figure 3, B).30 This has led to extensive studies aimed at understanding the role of the PrLD in TDP-43 proteinopathy. The N-terminal domain and RRM1 of TDP-43 do not form
aggregates, while fragments containing RRM2 and the C-terminal region aggregate with similar efficiency as full-length TDP-43, further supporting the theory that the PrLD is essential for TDP-43 aggregation.\textsuperscript{31,32}
Figure 3: Schematic diagram of TDP-43 and its mis-localization and aggregation in ALS.

A

NH2

82 98 106 176 191 262 274 414

COOH

B

NH2

82 98 106 176 191 262 274 414

COOH

C

Healthy motor neuron

Degenerated motor neuron

TDP-43 inclusion
A TDP-43 contains two RNA recognition motifs (RRM), a nuclear localization signal (NLS), a nuclear export signal (NES) and a carboxy-terminal prion like domain. B The majority of the ALS causing mutations in TDP-43 cause amino acid exchanges in the carboxy-terminal PrLD. The black lines represent fALS causing amino acid exchanges caused by missense mutations. C TDP-43 shows nuclear localization in healthy motor neurons. In ALS, TDP-43 mis-localizes and forms aggregates in the cytosol. TDP-43 aggregates contain a truncated and hyperphosphorylated form of TDP-43.
1.1.1.2. Matrin3

Matrin3 is 125kDa DNA and RNA-binding nuclear matrix protein. Matrin3 is encoded by the MATR3 gene which locates on chromosome 5 in humans. Matrin3 is one of twelve major inner nuclear matrix structural protein that binds to the nuclear lamina. It is ubiquitously expressed in almost all body tissues with highest expression in the brain. The primary structure of Matrin3 includes: two RNA recognition motifs (RRM1 and RRM2), that recognize and bind RNA. Matrin3 interacts in an RNA-dependent manner with several proteins with established roles in RNA processing e.g., TDP-43 and FUS and maintains its interaction with RNA via its RRM2 domain; two Zinc finger domains (ZnF1 and ZnF2), that bind DNA and are also involved in Matrin3 protein interactions; a carboxy-terminal nuclear localization signal (NLS); and amino-terminal nuclear export signal (NES) (Figure 4).
Matrin3 contains two RNA recognition motifs (RRM), a carboxy-terminal nuclear localization signal (NLS), an amino-terminal nuclear export signal (NES) and two Zinc finger domains (ZnF).
Although the exact function that Matrin3 performs in the cells is still unclear, previous studies suggest that Matrin3 is involved in different cellular processes, such as binding and stabilizing mRNA, nuclear retention of hyper edited RNA, regulation of alternative splicing, regulation of chromosomal distribution, and DNA damage repair.\textsuperscript{39,40,41,42}

Directed studies and unbiased proteome- and transcriptome-wide analyses revealed interactions of Matrin3 with different macromolecules, including DNA, RNA and many different proteins. In fact, most cellular Matrin3 functions seem to be executed and defined through its interactions with other proteins. For example, Matrin3 performs its role in DNA damage response through its interaction with NONO/SFPQ. Matrin3 was also found to interact with many well-established ALS-associated RNA binding proteins, such as TDP-43, FUS, and HNRNPA1 (Figure 5).\textsuperscript{43,44}
Matrin3 interacts with different proteins that help Matrin3 perform its multiple functions. Matrin3 also interacts with several ALS-associated RNA binding proteins. The thickness of the line between the nodes indicates the confidence level of the interaction. This interaction network was created using data from the STRING consortium data base. Matrin3 was chosen as the network central node and all interactions above the confidence level (0.9) are shown.
The first evidence that identified mutations in MATR3 as a cause of fALS was in published by Johnson et al. in 2014. Upon performing genome sequencing, mutations leading to the amino acid substitution Phenylalanine 115 to Cysteine (F115C) in Matrin3 was identified as the cause of fALS (Figure 6). After the discovery of MATR3 as an ALS causing genes, further investigations were performed on MATR3 identifying additional mutations as a cause of ALS (Figure 6). Most of the ALS causing mutations are scattered across the entire genes resulting in amino acid changes in different regions and are not concentrated in specific domains, as observed for TDP-43 and FUS where most ALS-associated mutations localize to their PrLD.30,45 In fact, none of the ALS-causing mutations in MATR3 identified so far are located within the RNA binding domains (RRM1 and 2) and the Zinc finger encoding domains.

Figure 6: ALS-causing variants of Matrin3.

Diagram of the domain structure of Matrin3 including the locations of the ALS causing amino acid changes caused by missense mutations.
1.2. Protein folding and misfolding inside the cell

Proteins are complex macromolecules that play essential roles inside our cells, such as maintaining cellular structural integrity, intracellular transport, and regulation of gene expression. Therefore, maintaining the integrity of the proteome, i.e., the sum of all proteins in a cell or organisms, is essential for maintaining viability. For proteins to perform their function inside the cells, they must attain their distinct native three-dimensional structure. Thus, protein function does not only depend on the primary amino acid sequence of the protein, but also on the folding of the protein into its distinct, thermodynamically stable conformation.

Protein quality control and proper protein folding are facilitated by molecular chaperones. Molecular chaperones are a group of proteins, which bind to the nascent peptide during translation and fold into its proper three-dimensional structure by providing a favourable environment for folding of either the entire protein or a specific region of a protein.

Protein misfolding is a process by which proteins do not reach or lose their most stable conformation and often displays their hydrophobic amino acid residues on the surface of the protein, which is rarely found in properly folded proteins. Protein misfolding can occur due to stochastic fluctuations, destabilizing mutations, cellular stress conditions, such as a heat shock or oxidative stress, or defects in the cellular protein quality control as observed during ageing (Figure 7).

Under normal conditions, cells can prevent the toxic consequences of misfolded proteins through multiple pathways facilitated by molecular chaperones and others branches of cellular protein quality control, such as protein degradation by the ubiquitin proteasome system (UPS) and autophagy. Chaperones can bind to misfolded proteins and re-fold them into their proper conformation. Chaperones can also target misfolded proteins for degradation by both the UPS and autophagy. Finally, if the cellular protein quality control system fails to refold or eliminate misfolded protein, they accumulate inside the cell and form inclusions, which are major hallmarks of neurodegenerative disorders, including ALS (Figure 7).
Figure 7: Cellular protein quality control regulates protein folding and misfolding.

Diagram showing proper protein folding and the fate of misfolded proteins. Misfolded proteins can be either refolded, sequestered, or targeted for degradation.
1.3. **Mechanisms of protein misfolding in ALS**

1.3.1. **Liquid-liquid phase separation**

Liquid-liquid phase separation (LLPS) emerges as a new paradigm underlying the intracellular assembly of proteins and other macromolecules, such as RNA. LLPS is a reversible process of a homogenous fluid de-mixing into two distinct phases, a condensed phase, and a dilute phase.\(^{54,55,56}\) LLPS has been identified as a critical physical principle underlying the formation of numerous membrane-less cellular organelles, such as p-bodies, stress granules and nucleoli.\(^{57,58}\) In addition to membrane-less organelles, other structures are also formed through LLPS, such as transport channels in the nuclear pore complex, and membrane receptor clusters at the cell membrane.\(^{59,60}\)

Defects in LLPS have been speculated to drive protein misfolding and aggregation in different human diseases, such as cataract and neurodegenerative disorders, including ALS.\(^{61,62}\) For example, proteins inside stress granules can undergo liquid-solid phase transition or can nucleate to form protein aggregates. This process can be accelerated by disease causing mutations in ALS-associated genes. For example, ALS-causing mutations in the gene encoding FUS causes drastic changes in the biophysical properties of the liquid droplets and induce rapid formation of protein fibrils, which are speculated to be cytotoxic (Figure 8).\(^{63}\)

A major category of proteins that can undergo LLPS are intrinsically disordered proteins (IDPs). Specifically, proteins with intrinsically disordered domains and RNA binding domains, such as Matrin3, can facilitate the formation of multivalent interactions allowing them to undergo LLPS.\(^{64}\)
Figure 8: Aberrant liquid-liquid phase separation (LLPS) can lead to the formation of protein inclusions.

Diagram illustrating LLPS of RNA binding proteins (Blue) and RNA molecules (Red) and the formation of membrane-less organelles. Defects in LLPS leads to protein misfolding.\textsuperscript{65}
1.3.1.1. **Intrinsically disordered proteins**

Proteins with large intrinsically disordered domains (IDDs), commonly known as intrinsically disordered proteins (IDPs), represent a large class of macromolecules in the human proteome that can undergo LLPS. Intrinsically disordered domains are regions in the protein that lack a defined secondary structures and well-structured three-dimensional conformations thus giving them a high degree of flexibility. Many of the IDD follow a coupled folding and binding process, which means that the IDD can fold into stable secondary or even tertiary structures when binding to interacting proteins.

Unlike predicting the three-dimensional structure of globular protein, which remains a mostly unsolved challenge, determining the IDD in a protein is a straightforward process since the amino acid composition of the IDD is biased toward low proportions of hydrophobic amino acids and high content of polar and charged amino acids. Since determining whether a region of a protein is disordered or not is a straightforward process, multiple computational algorithms and webservers have been developed to detect IDDs such as PONDR, FoldIndex, and DISOPRED. IDDs can be further classified according to their amino acid composition such as glutamine-rich or proline-rich regions.

IDPs are abundant in the human proteome, and they play a crucial role in various cellular processes, such as regulation of gene transcription, translation, cellular signal transduction, and, most importantly, regulation of self-assembly of large multiprotein complexes, such as ribosomes. Since IDPs play crucial roles in cells, previous studies have shown that IDD can be highly conserved in both sequence and amino acid composition.
1.3.1.2. **Prion-like domains**

Prions are infectious proteins that can misfold and adapt a specific, often insoluble conformation, which can propagate through healthy cells causing similar or the same kinds of proteins to adopt the pathological form. Prions cause many devastating neurodegenerative disorders, such as Creutzfeldt–Jakob disease, Gerstmann–Sträussler–Scheinker syndrome in humans, and chronic wasting disease, scrapie, and mad cow disease in animals.

Prion-like domains (PrLDs) are a subset of IDDs that are rich in glycine and uncharged amino acids. PrLDs were first discovered in yeast and can be identified using hidden Markov algorithms in the proteomes of many other organisms. PrLDs are found in mammalian proteins, especially RNA binding proteins, and have similar amino acid composition to that of yeast prions.

Many RNA binding proteins that have been identified to cause ALS, such as TDP-43 and FUS, contain PrLDs. PrLDs play a critical role in the normal function of these proteins; for example, the PrLD of TDP-43 mediates protein–protein interactions with splicing factors and facilitate its recruitment to stress granules. PrLDs are also crucial for the formation of TDP-43 protein aggregates in ALS. The truncation of the C-terminal PrLD eliminated TDP-43 inclusion formation and toxicity in different model systems. Furthermore, in ALS, most of the mutations in RNA binding proteins occur in the PrLDs, which indicates a major contribution of PrLDs in the pathogenesis of the disease.
1.3.1.3. Posttranslational modifications

One of the factors that regulate protein folding and misfolding is posttranslational modification (PTMs). PTMs are a group of chemical alterations of proteins at the later stages of or after protein biosynthesis. PTMs range from covalent addition of functional groups to proteolytic cleavages of whole protein subunits. PTMs, such as phosphorylation, ubiquitination or SUMOylation, can alter the conformation and biological function of proteins, can affect protein folding and aggregation, and thereby play a role in neurodegenerative diseases. For the purpose of my project, I focused on protein phosphorylation as major and ubiquitous PTM.

1.3.1.3.1. Protein phosphorylation and protein folding and misfolding

Protein phosphorylation can affect the three-dimensional structure, function, localization, and stability of proteins in different ways: it may be required for proper protein folding; it may induce conformational changes; it can function as a recognition signal for further modifications, and finally, it may modify protein-protein interactions. Changes in the protein phosphorylation pattern are associated with protein misfolding in multiple neurodegenerative disorders, including ALS. Abnormal phosphorylation takes place in deposited, aggregated TDP-43. Multiple phosphorylation sites were identified in the carboxyl-terminal prion-like domain of aggregated TDP-43. These studies suggest that phosphorylated TDP-43 is a major component of pathological inclusions, and that abnormal phosphorylation of TDP-43 may be a critical step in the pathogenesis of ALS. While it is clear is that phosphorylation is relevant in the context of protein misfolding, aggregation, mis-localization and toxicity, the mechanism by which protein phosphorylation contribute to the process of protein misfolding and then neurodegeneration remains unclear. For example, there is still no evidence suggesting that protein phosphorylation is the actual cause rather than merely a consequence of aggregation, or whether phosphorylation has a neurotoxic or neuroprotective effect. For the purpose of this study, I focused on Matrin3 phosphorylation.
1.3.1.3.2. **Matrin3 phosphorylation**

Matrin3 is a phosphoprotein with several phosphorylation sites for tyrosine and serine/threonine kinases. Phosphorylation of tyrosine residues of Matrin3 can modulate its DNA binding activity. Although multiple phosphorylation sites have been identified in Matrin3 only four different protein kinases are known to phosphorylate Matrin3. The ataxia-telangiectasia mutated (ATM) kinase phosphorylates Matrin3 at serine 208 (Ser208) in response to DNA double stranded breaks (DSBs) and thereby activates the DNA damage repair response. Previous studies confirmed Matrin3 as a novel target of Pyruvate kinase M2 (PKM2)-mediated phosphorylation. PKM2 phosphorylates Matrin3’s threonine residue 293 (T239), which inhibits Matrin3 degradation by preventing its K48-linked ubiquitylation. Matrin3 was also identified as a target for Protein kinase A (PKA)-mediated phosphorylation. In response to NMDA receptor activation, PKA-mediated Matrin3 phosphorylation leads to its degradation in the nuclei of neurons in the brain, which leads to neuronal death. Finally, Mass spectrometry results show that Protein kinase B (PKB) also known as Akt, a serine/threonine kinase, also phosphorylates Matrin3.

Although many studies have proven that Matrin3 is heavily phosphorylated by many different kinases, the contribution of Matrin3 phosphorylation in the pathogenesis of ALS is not yet fully understood. There is still no evidence whether there is a change in Matrin3 phosphorylation pattern in ALS or not. And if there are any changes, it remains unclear if phosphorylation drives Matrin3 misfolding and mis-localization or whether is it a consequence of these events.
1.4. Yeast as a model organism to study protein misfolding

Saccharomyces cerevisiae also known as baker’s yeast or budding yeast is eukaryotic single-celled organism that is frequently used to study protein misfolding associated with neurodegenerative disorders.\textsuperscript{90} There are multiple characteristics that qualifies it to be a good model system generally: its unique and vast genetic tools, rapid growth, short cell cycle, the ability to conduct high throughput studies, and the availability of genome-wide databases; in fact, yeast was the first eukaryotic organism to have its genome fully sequenced and there are deletion and overexpression libraries for every yeast gene.\textsuperscript{91,92} Nearly a third of yeast genes have a direct human orthologues and more than two thirds have significant homology with human genes.\textsuperscript{93} Yeast cells are also able to adjust their metabolism in response to the provided carbon source in the growth media. For example, in media containing fermentable carbon sources, such as glucose or galactose, yeast rely on glycolysis and aerobic fermentation as a main source of energy. However, in media containing non-fermentable carbon sources, such as glycerol, yeast cells switch their metabolism from glycolysis to oxidative phosphorylation carried out by the mitochondria as their major source of ATP, which makes yeast cells more similar to many human cells, including neurons, which also heavily rely on oxidative phosphorylation for ATP production.\textsuperscript{94} The most important characteristic that makes yeast a highly effective model to study protein misfolding and its associated toxicity is that most of the cellular pathways involved in cellular protein quality, such as molecular chaperones and most proteins involved in protein degradation are highly conserved between humans and yeast.\textsuperscript{95,96}

In sum, yeast is a very powerful tool to study protein misfolding of protein involved in neurodegenerative diseases. There are several established yeast models to study protein misfolding in various neurodegenerative disorders, such as $\alpha$-synuclein ($\alpha$-syn) and parkin (PARK2) yeast models for Parkinson’s Disease (PD), the Huntington (Htt) polyQ expansion in Huntington’s Disease (HD), and TDP-43 and FUS in ALS. Of note, many studies using yeast models have produced results that deciphered genetic and cellular mechanisms underlying protein misfolding that are highly relevant to human neurodegeneration.\textsuperscript{97} For the purpose of this study, we established a yeast model for Matrin3 protein to investigate Matrin3 misfolding and mis-localization.
1.5. **Mammalian neuronal cell model: SN56 cells**

The next step after establishing our yeast model is to evaluate the translational nature of our yeast results in mammalian neuronal cells. In this study we used SN56 cell line, which is a mouse cholinergic septal neuronal cell line that is frequently used to study different neurodegenerative disorders. For instance, SN56 cells have been used to study the misfolding of several ALS-associated proteins, such as TDP-43, FUS, and RGENF. 98,99

1.6. **Rationale**

1.6.1. **Matrin3 phosphorylation**

Protein phosphorylation can affect the three-dimensional structure, function, and fate of proteins in many different ways. Protein phosphorylation may induce conformational changes that can lead to protein misfolding and inclusion formation. 84 Many studies have explored protein phosphorylation in the context of neurodegeneration, for example, abnormal phosphorylation pattern has been observed in TDP-43 inclusions in ALS patients. These studies suggest that abnormal phosphorylation of TDP-43 is a critical step in the pathogenesis of ALS and FTD. 27 Matrin3 is heavily phosphorylated, but the contribution of Matrin3 phosphorylation to the pathogenesis of ALS is not yet fully understood. There is still no evidence whether there is a change in Matrin3 phosphorylation pattern in ALS or not. And if there are any changes, do these drive Matrin3 misfolding and mis-localization or are they just a consequence of protein misfolding? All these questions require a comprehensive investigation of Matrin3 function in ALS.
1.6.2. **Intrinsically disordered domain and Matrin3 misfolding**

LLPS is reversible process of a homogenous fluid de-mixing into two distinct phases, a condensed phase, and a dilute phase.\(^{54,55}\) There is increasing evidence that defects in phase separation process might be the driving force for protein misfolding and aggregation in neurodegenerative diseases including ALS.\(^{61,62}\) Proteins contain large intrinsically disordered regions such as prion like domains can undergo LLPS. Many of the RNA binding proteins, such as TDP-43 and FUS that are associated with ALS, contain prion-like domains and most of the fALS causing amino acid changes in these proteins occur in prion-like domains.\(^{64,80}\) This indicates that intrinsically disordered regions play a role in the process of protein misfolding and aggregate formation. Previous studies have shown that Matrin3 is an intrinsically disordered protein, and the amino terminal regions is required for Matrin3 to undergo LLPS.\(^{18,100}\) In this study, we seek to determine the role of the intrinsically disordered regions on Matrin3 misfolding and toxicity.

1.7. **Hypothesis, Objectives**

We hypothesize that Matrin3 phosphorylation and its intrinsically disordered domains drive Matrin3 aggregation and toxicity in yeast and neuronal cells models of ALS.

My specific objectives are:

1. Detect Matrin3 aggregation and mis-localization in patients with sporadic and familial ALS.
2. Determine the effect of Matrin3 phosphorylation on its toxicity and localization in the yeast model and in cultured mammalian neuronal cells.
3. Determine the role of the amino-terminal disordered domain, its toxicity, and localization in the yeast model and in cultured mammalian neuronal cells.
1.8. **Significance**

In order to fully understand the pathological mechanisms underlying protein misfolding in ALS, we must focus on fundamental mechanistic aspects, using experimentally tractable model organisms, such as yeast and cultured neuronal cells before we can start meaningful experiments in animal models. Protein aggregation is the hallmark of many neurodegenerative diseases. ¹,² Studying the mechanisms underlying Martin3 misfolding will provide insight on the role of PTMs and IDD on protein misfolding and toxicity in ALS and eventually develop a therapeutic solution for not just ALS but possibly also other neurodegenerative diseases.
2 Materials and methods

2.1 Materials

2.1.1 Yeast strains

Yeast W303 (MAT a leu2- 3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15) and W303 Δpdr5 (MAT a leu2- 3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15, Δpdr5) strains were used in this study.

2.1.2 E. coli strains

Escherichia coli Strain DH5α Genotype F–Φ80lacZΔM15 Δ(lacZYA-argF) 169 recA1 endA1 hsdR17 (rK–, mK+) phoA supE44 λ– thi-1 gyrA96 relA1 was used in this study. Subcloning efficient DH5α Competent Cells (Invitrogen, Cat. No. 18265-017) were used for the cloning work in this study.

2.1.3 Mammalian cell lines

SN56 cell line; a mouse cholinergic septal neuronal cell line was used in this study. The SN56 cells were provided by Dr. Marco Prado (Robarts Research Institute, University of Western Ontario).

Table 1: Yeast growth media

<table>
<thead>
<tr>
<th>Name</th>
<th>Composition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>YPD</td>
<td>10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose</td>
<td>Rich yeast media used for non-selective yeast growth in liquid media or on agar plates</td>
</tr>
</tbody>
</table>
**Selective Dextrose (SD)**

<table>
<thead>
<tr>
<th>Composition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2% glucose, Yeast Nitrogen Base (YNB), 6 g/L L-isoleucine, 2 g/L L-arginine, 4 g/L L-lysine, 6 g/L L-phenylalanine, 1 g/L L-threonine</td>
<td>Synthetic yeast media. Used for either non-selective yeast growth (when all selective amino acids are present) or for selective growth, when one of them is absent. Can be used as liquid media or for agar plates*</td>
</tr>
<tr>
<td>Selective amino acids: 4g/L L-tryptophan, 6 g/L L-leucine, 2 g/L L-histidine, 2g/L uracil, 1 g/L L-methionine</td>
<td></td>
</tr>
</tbody>
</table>

**Selective Galactose (SGal)**

<table>
<thead>
<tr>
<th>Composition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2% galactose, Yeast Nitrogen Base (YNB), 6 g/L L-isoleucine, 2 g/L L-arginine, 4 g/L L-lysine, 6 g/L L-phenylalanine, 1 g/L L-threonine, 1 g/L L-methionine</td>
<td>Synthetic yeast media. Used for protein expression induction Can be used as liquid media or for agar plates*</td>
</tr>
<tr>
<td>Selective amino acids: 4g/L L-tryptophan, 6 g/L L-leucine, 2 g/L L-histidine, 2g/L uracil</td>
<td></td>
</tr>
</tbody>
</table>

*23 g/L agar was used for preparing agar plates

**Table 2: E. coli growth media**

<table>
<thead>
<tr>
<th>Name</th>
<th>Composition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysogeny (LB) broth</td>
<td>5 g NaCl</td>
<td>Media for selective e. coli growth, transformations, and cloning</td>
</tr>
<tr>
<td></td>
<td>5 g yeast extract</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 g tryptone</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Selective Antibiotics:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ampicillin (1x)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kanamycin (1x)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Can be used as liquid media or for agar plates*</td>
</tr>
</tbody>
</table>

---

28
<table>
<thead>
<tr>
<th>Name</th>
<th>Composition</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dulbecco's Modified Eagle Medium (DMEM, Wisnet)</td>
<td>10% fetal bovine serum, 1x penicillin/streptomycin</td>
<td>Basal media for mammalian cell growth</td>
</tr>
<tr>
<td>Opti MEM</td>
<td></td>
<td>Reduced serum media for mammalian cell differentiation and transfection</td>
</tr>
</tbody>
</table>

### Table 3: Mammalian cell growth media

<table>
<thead>
<tr>
<th>Name</th>
<th>Purpose</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pDONR201</td>
<td>Gateway cloning vector</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pAG423GAL-Stop/YFP/HA</td>
<td>Expression of MATR3 WT, missense mutations, and truncated variants in yeast</td>
<td>Susan Lindquist. Addgene number: 14149 (stop) 14221 (YFP) 14245 (HA)</td>
</tr>
<tr>
<td>pAG413GPD-Stop/YFP</td>
<td>Expression of MATR3 WT in yeast</td>
<td>Susan Lindquist. Addgene number: 14142 (stop) 14334 (YFP)</td>
</tr>
<tr>
<td>pAG413MET-Stop</td>
<td>Expression of MATR3 WT, missense mutations and</td>
<td>Duennwald lab</td>
</tr>
</tbody>
</table>

2.1.4. DNA Plasmids

### Table 4: Plasmids

*23 g/L agar was used for preparing agar plates*
Expression of MATR3 WT, missense mutations, and truncated fragments in mammalian cells

2.1.5. Antibodies

Table 5: Antibodies

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Use</th>
<th>Concentration</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>PGK-1</td>
<td>Western Blot</td>
<td>1:1000</td>
<td>Thermofisher Scientific</td>
</tr>
<tr>
<td>αTubulin</td>
<td>Western Blot</td>
<td>1:1000</td>
<td>Abcam</td>
</tr>
<tr>
<td>Matrin3</td>
<td>Western Blot</td>
<td>1:500</td>
<td>Sigma</td>
</tr>
<tr>
<td></td>
<td>Immunohistochemistry</td>
<td>1:800/1:1600</td>
<td></td>
</tr>
<tr>
<td>Matrin3</td>
<td>Western Blot</td>
<td>1:2000</td>
<td>Thermofisher scientific</td>
</tr>
<tr>
<td></td>
<td>Immunofluorescence</td>
<td>1:1000</td>
<td></td>
</tr>
<tr>
<td>Rabbit Alexa 680</td>
<td>Western Blot Immunofluorescence</td>
<td>1:1000</td>
<td>Thermofisher Scientific</td>
</tr>
</tbody>
</table>
### 2.1.6. Additional Key Reagents

**Table 6: Additional key reagents and chemicals**

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Purpose</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>MG-132</td>
<td>Reversible inhibitor of 26s proteasome</td>
<td>SigmaAldrich</td>
</tr>
<tr>
<td>Hydrogen Peroxide</td>
<td>Oxidizing agent. Induces protein, lipid, and nucleic acid damage. Induces oxidative stress response</td>
<td>SigmaAldrich</td>
</tr>
<tr>
<td>Enhanced Chemiluminescence (ECL) Kit</td>
<td>Highly sensitive method for detecting proteins bound to western blotting membranes</td>
<td>BioRad</td>
</tr>
<tr>
<td>Gateway Cloning Kit</td>
<td>Cloning MATR3 into expression vector</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>Q5 Site-Directed Mutagenesis Kit</td>
<td>Production of MART3 mutation variants</td>
<td>New England BioLabs (NEB)</td>
</tr>
<tr>
<td>QIAprep Spin Miniprep Kit</td>
<td>Isolation of DNA from E. coli culture</td>
<td>Qiagen</td>
</tr>
<tr>
<td>QIAQuick gel extraction kit</td>
<td>Agarose gel DNA extraction</td>
<td>Qiagen</td>
</tr>
<tr>
<td>Pierce™ BCA Protein Assay Kit</td>
<td>Normalization of protein concentrations</td>
<td>ThermoFisher Scientific</td>
</tr>
<tr>
<td>Lipofectamine LTX</td>
<td>Mammalian cell transfection</td>
<td>ThermoFisher Scientific</td>
</tr>
<tr>
<td>CellTiter-Glo Luminescent Cell Viability Assay</td>
<td>Method for determining the number of viable cells in culture</td>
<td>Promega</td>
</tr>
</tbody>
</table>
Table 7: Human spinal cord tissue samples

Human spinal cord tissues samples (controls and ALS cases) were kindly provided by the team of Dr. Michael Strong (Robarts Research Institute, University of Western Ontario).

<table>
<thead>
<tr>
<th>Case Number</th>
<th>Tissue</th>
<th>Type</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>09-006 A7</td>
<td>SC</td>
<td>control</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>97-139 N8</td>
<td>SC</td>
<td>control</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>97-148 N6</td>
<td>SC</td>
<td>control</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>97-033 N5</td>
<td>SC</td>
<td>control</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>97-070 N11</td>
<td>SC</td>
<td>control</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>A01-25 14</td>
<td>SC</td>
<td>sALS</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>A01-115 N7</td>
<td>SC</td>
<td>sALS</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>A01-342 N30</td>
<td>SC</td>
<td>sALS</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>A01-363 N3</td>
<td>SC</td>
<td>sALS</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>87702 N37/33</td>
<td>SC</td>
<td>sALS</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>94165 N10</td>
<td>SC</td>
<td>sALS</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>95205 N2</td>
<td>SC</td>
<td>sALS</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>98027 N17</td>
<td>SC</td>
<td>sALS</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>98112 N10</td>
<td>SC</td>
<td>sALS</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>00417 N8</td>
<td>SC</td>
<td>sALS</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>00410 N7</td>
<td>SC</td>
<td>sALS</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>98-052 N12</td>
<td>SC</td>
<td>SOD1 (113T)</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>04-511 A24</td>
<td>SC</td>
<td>SOD1 (D76Y)</td>
<td>Strong Lab</td>
</tr>
<tr>
<td>00-394 N7</td>
<td>SC</td>
<td>SOD1 (A4T)</td>
<td>Strong Lab</td>
</tr>
</tbody>
</table>
2.2. Experimental methods

2.2.1. Matrin3 in silico analysis

A Charge hydropathy plot was produced using PONDR according to the method introduced by Uversky et al. The plotting was performed using MATLAB software. Four webservers MFDp2, PONDR, IUPred2A, and DISOPRED3 were used to predict the intrinsically disordered regions in the protein. Matrin3 post translational modification: PhosphoSitePlus data was used for the Matrin3. Low- and high-throughput numbers were summed for each site. Plotting was performed in MATLAB.

2.2.2. Cloning of MATR3 and creation of MATR3 mutant variants

2.2.2.1. Gateway cloning

2.2.2.1.1. BP reaction

The BP recombination reaction was performed by combining 2μl of destination vector plasmid DNA (about 100-200 ng of DNA) or PCR product, 1 μl of pDONR vector (150 ng/μl), and 2 μl of 5X BP Clonase (Invitrogen). The mixture is vortexed and centrifuged twice to ensure thorough mixing and then allowed to incubate overnight at 30°C. 1 μl of Proteinase K (Invitrogen) solution is added to the reaction following incubation and left to react at 37°C for 10 minutes. 2 μl of this
reaction mixture is then used to transform into Subcloning Efficiency DH5α Competent Cells (Invitrogen, Cat. No. 18265-017) following the protocol provided by the manufacturer. The cells are plated in LB kanamycin (kan) resistant agar plates and incubated overnight at 37°C. Colonies are picked from the plates and inoculated into LB kan liquid overnight at 37°C. The plasmids are then extracted from the E. coli cells by using the High-Speed Plasmid Mini Kit (Qiagen); the resulting DNA is in pDONR vector backbone.

### 2.2.2.1.2. LR reaction

The LR recombination reaction uses 1µl of the product (100-300 ng) from the BP reaction, i.e., the pDONR vector, in combination with 2µl of destination vector (150 ng/µl), 13µl of TE buffer, and 4µl of LR Clonase (Invitrogen). The destination vectors used are listed in section 2.1.4. The mixture is vortexed and centrifuged twice and allowed to incubate overnight at 37°C. The resulting procedure is the same as described for the BP recombination reaction—Proteinase K is added to the mixture, incubated, and transformed into competent cells. The destination vectors are ampicillin (amp) resistant and therefore must be plated on LB amp agar plates and then inoculated in LB amp liquid media.

### 2.2.2.2. Cloning of WT MATR3

MATR3 was PCR-amplified from human WT MATR3 template provided by Dr. Jeehye Park (Hospital of sick kids, Toronto, ON, 416-813-7670).

The PCR product was cloned into pDONR201 vector (Table 5) using Gateway® technology (Invitrogen) BP reaction and following the manufacturer’s protocols. The correct sequence of the entry clones was confirmed by Sanger sequencing (performed by Robarts London Regional Genomics Centre). MATR3 was cloned into p423GAL-Stop/YFP, p413 GPD-Stop/YFP and p413 Met-Stop (Table 5) plasmids using the LR reaction. The identity of the destination vector containing MATR3 was verified by restriction digest.
2.2.2.3. Cloning of MATR3 mutants

Wild type pDONR201 MATR3 template (10-25 ng) was PCR amplified with Q5 Hot Start High-Fidelity kit (Table 5) and non-overlapping mutagenetic primers in a 25μl reaction. The resulting PCR products were ligated using the Kinase, Ligase, and DpnI (KLD) reaction and transformed into E. coli. Several colonies were then picked and inoculated in 3mL of liquid LB Kan overnight. Cultures were then centrifuged, and DNA was extracted from the pellets using the QIAprep Spin Miniprep kit (Table 5). The presence of mutations was then verified by Sanger sequencing (performed by Robarts London Regional Genomics Centre). Mutated DNA was then cloned into pAG423 Gal-Stop/YFP, pAG413 GPD Stop/YFP destination vector (Table 5) using the LR reaction. The identity of the destination vector containing the mutated DNA was verified using restriction digest.

2.2.2.4. Cloning of MATR3 truncated fragments

Wild type pDONR201 MATR3 template (10-25 ng) was PCR amplified with Q5 High-Fidelity kit (Table 5) and non-overlapping truncation primers in a 25μl reaction The PCR product was cloned into pDONR201 vector (Table 5) using Gateway® technology (Invitrogen) BP reaction and following the manufacturer’s protocols. The correct sequence of the entry clones was confirmed by Sanger sequencing (performed by Robarts London Regional Genomics Centre). MATR3 was cloned into p423GAL-Stop/YFP, and p413 Met-Stop (Table 5) plasmids using the LR reaction. The identity of the destination vector containing MATR3 was verified by restriction digest.

2.2.2.5. Gel purification

PCR products were confirmed by running the DNA on 1% agarose gel. The PCR products were then extracted using Qiagen gel extraction kit (table 6).
2.2.2.6. DNA Isolation

After DNA cloning and E. coli transformation, DNA was isolated using Qiagen Miniprep Kit according to table 6.

2.2.3. E. coli transformation

We performed E. coli transformations to replicate and amplify plasmid DNA, resulting in abundant amounts of DNA. 100 µl aliquots of transformation competent DH5α cells were thawed on ice from storage at -80°C; 1-5 µL (0.1~0.5 µg) of plasmid DNA was added to the cells and mixed thorough by gently flicking the tubes (the competent cells should not be vortexed). The cells were allowed to recover on ice for about 30 mins and heat shocked at 40°C for 45 seconds. The cells rested on ice for 2 mins before 1 mL of 2xYT was added and the cells allowed to recover at 37°C in a shaking incubator for at least 1 hour. Following recovery, the cells were centrifuged at 10,000 xg for one minute, the supernatant was aspirated off, and the pellet was resuspended in 100 µl of 2xYT. The suspension was plated on LB agar plates with selective antibiotic depending on the antibiotic resistance of the vector.

2.2.4. Yeast transformation

DNA constructs containing the human MATR3 were transformed into yeast cells by the lithium acetate methods to produce yeast strains that stably express Matrin3 when kept under selective auxotrophic selection in their growth media. In brief, single colonies of yeast cells were inoculated into 3 mL of liquid media and incubated at 30°C with shaking overnight. The liquid culture was then diluted to a 30 mL liquid culture with medium and incubated at 30°C shaking till the cells reached mid logarithmic growth phase (an OD600 of 0.4 to 0.5). The culture was then centrifuged for 5 minutes to pellet the yeast cells. The supernatant was aspirated off, discarded and the pellet was washed with 3 ml of sterile water. The cells were centrifuged again at the same speed and time. The pellet was resuspended in 2 ml of 100 mM Li-Acetate in TE buffer after the wash step and incubated at 30°C shaking for 10 minutes. The culture was centrifuged again after incubation and the pellet was resuspended in 100 µl of Li-Acetate per
transformation. Each transformation was composed of 100 μl cell suspension, 250 μl transformation (1 X TE, 40% PEG, and 100mM Li-Acetate), 12μl salmon sperm DNA, 1μl (0.3–0.5 μg) plasmid DNA, and 25μl DMSO and in the order listed and vortexed thoroughly. The cells then recovered at 30°C shaking for 30 minutes, following a 20-minute of heat shock at 42°C shaking. After heat shock, the cells were centrifuged for 1 minute at 2000 xg, the supernatant was aspirated, and the pellet resuspended in 50 μl TE buffer. The cells were then plated onto selective agar plates.

2.2.5. Yeast spotting assay

Yeast growth (spotting) assays were carried out by inoculating yeast strains overnight in 3ml of selective growth media at 30°C and grown to saturation. The optical density (OD600) was measured for each strain following 16 hours of growth by diluting 100μl of yeast into 900μl of water. Diluted cultures were vortexed and added to a cuvette and the OD600 is measured by a spectrophotometer. To ensure equal number of cells are plated in the assay, cultures were normalized to a standard .02 OD600 across conditions. Normalized cultures were transferred to the top wells of sterile 96-well plates (Greiner bio-one #650161). A five-fold serial dilution was performed using a multichannel pipette to transfer 30μl of the normalized cultures into 120μl of water. Cell dilutions were then transferred to agar plates using a sterilized, 48-pin “replicate - plater”. Cells were spotted on non-inducing, overgrowth controls and inducing, selective plate.

2.2.5.1. Spotting assay quantification

The growth of yeast cells was then quantified using mean gray value densitometry in ImageJ software as described before. The mean gray values of a single spot and of the plate background (no yeast colonies present) were measured on a grayscale, 8-bit image of a spotting assay. The mean gray value of the background was subtracted from the mean gray value of the spot. Those values were divided by the mean gray value of the positive control, yielding a value that represents the relative amount of yeast cell growth for that dilution. Relative mean growth values from three independent biological replicates were statistically analyzed to identify any significant differences in growth between experimental and control groups. Statistical analysis of the quantification
results consisted of a one-way ANOVA followed by a post-hoc Tukey’s HSD test (95% confidence level) and was performed using the RStudio software.

2.2.6. Live cell fluorescence microscopy

Microscopy imaging of YFP tagged constructs was done by first inoculating yeast cells in SD media at 30°C overnight. The cells are then washed twice with ddH2O, resuspended in SD media, or various stress condition media (AZC, H2O2, and Radicicol) incubated at 30°C (37°C to induce heat stress). After inducing for 16 to 20 hours, 2µl of the liquid culture are placed on a microscope slide. Fluorescence microscopy was performed using the Cytation5 cell imaging multi-mode reader (BioTek, Winooski, WI, USA). LED cubes and imaging filter cubes from biotech were employed for YFP tagged proteins: 465 LED 1225001 Rev J, GFP 469/525 1225101 Rev J, BioTekImages were captured on the 20× objective. Image analysis was completed using the Gen5 imaging software (BioTek)

2.2.7. Western blot

2.2.7.1. Alkaline lysis of yeast cells

For preparing total protein lysates in yeast, liquid yeast cultures were grown overnight and centrifuged. The pellet was resuspended in 100µl ddH2O and 100µl of 0.2 M NaOH and incubated at room temperature for 10 minutes. The sample was then centrifuged, and the pellet was resuspended in 50µl of SDS sample buffer (0.06 M Tris-HCl pH 6.8, 5% Glycerol, 2% SDS, 4% BME, 0.0025% BromoPhenol Blue) and boiled for 3 minutes. The sample is then centrifuged for a final time. The supernatant was obtained and loaded into the gel.

2.2.7.2. Mammalian cell lysis

For preparing total protein lysates in mammalian cells, Cells were washed with cold PBS which was then removed by aspiration. A cell scraper was used to detach cells from culture plates into
cold lysis buffer (50 mM Hepes pH 7.5, 150 mM NaCl, 5 mM EDTA, 1% (v/v) Triton X-100, and SIGMA FAST\textsuperscript{TM} Protease Inhibitor), after which cell lysates were frozen at -20°C. After thawing on ice, lysates were subjected to a Branson Sonifier\textsuperscript{®} ultrasonic cell disruptor for two 3-5 sec pulses. Lysates were then centrifuged at 20000×g for 10 min at 4°C. The supernatants were obtained, and protein concentrations were determined using the BCA kit. The samples were then loaded into the gel at the appropriate concentrations.

2.2.7.3. Normalization of protein concentrations (mammalian cells)

We perform a BCA Protein Assay to determine the concentration of protein in the sample. The assay was performed according to the Thermo Scientific Pierce BCA Protein Assay Kit Instructions. After obtaining the concentration and normalizing the total protein amount in each sample per blot, we dilute the samples with 4x reducing SDS buffer (0.25M Trisma Base pH 6.8, 8.0% SDS, 40% sterile glycerol, 10% β-mercaptoethanol, 0.04% bromophenol blue).

2.2.7.4. SDS-PAGE

We run SDS-PAGE with the samples on an 8-16% gradient gel (Bio-Rad Criterion TGX Stain-Free Precast Gels) or 12% acrylamide gels at 220 V for about 30 mins. The gel is then transferred onto a Nitrocellulose or PVDF membranes (BioRad) using the Bio-Rad Trans-Blot Turbo machine following the manufacturer’s protocol. We then block the membrane using 5 or 10% skim milk powder (Carnation) in Phosphate Buffered Saline with 0.01 % (v/v) Tween (PBST) and incubating for 1 hour on a shaker. The membrane is then incubated in primary antibody overnight on shaker. Following incubation with the primary antibody, we wash the membrane with 50mL aliquots of PBST at ten-minute intervals for an hour on shaker and then incubate in the secondary antibody for 1 hour on a shaker. The membrane is washed again with 50mL aliquots of PBST in ten minutes intervals on shaker for an hour. The membrane is then documented using the ChemiDoc MP System (Bio-Rad).

2.2.8. Mammalian cell viability assay

The CellTiter-Glo Luminescent Cell Viability Assay (Promega) was used in this study. This assay delivers a highly sensitive read out of cellular fitness and is particularly useful for studies on the
toxicity associated with misfolded proteins. Transfected SN56 cells are split into 96 well plates and grown in minimum DMEM with 1% fetal bovine serum (FBS, Gibco) and 1 g/L glucose. This minimal medium increase sensitivity to the toxic effects of protein misfolding in many other systems, as well as differentiating N2a cells. The cells are incubated at 37 °C for 20~24 hr. The cell viability assay is then performed according to the supplier’s instructions. Plates were then measured using the Cytation 5 Cell Imaging Multi-Mode Reader (Biotek).

2.2.9. Immunofluorescence microscopy

Following transfection, the cells are split into 8 well chambers (Labtek) and grown in DMEM at 37°C for 20~24 hr. Each well is seeded with ca. 30,000 cells. The media is then aspirated off and the cells washed 2 times with PBS. 4% paraformaldehyde is then added to the wells to fix the cells for 15 min. This is followed by 3 washes in PBS. The cells are permeabilized with 0.5% TritonX-100 in PBS for 15 mins and washed with PBS 3 times, followed by 5 washes in 0.5% BSA (Sigma) in PBS (PBB). We block non-specific binding sites with 20% normal goat serum (Gibco) in PBB for 45 min. The cells are then washed with PBB 5 times and incubated for an hour in rabbit anti-Matrin3 polyclonal antibody (Thermofischer MATR3 antibody 1:1000 dilution). We wash the cells with PBB 5 times and incubate for an hour in α-rabbit Texas red secondary antibody (Thermo Fisher). Following the secondary antibody incubation, we wash in PBB 5 times. All the liquid is then aspirated off and mounted with ProLong® Gold Antifade Mountant with DAPI. SN56 cells were imaged with the Cytation 5 Cell Imaging Multi-Mode Reader (Biotek).

2.2.10. Immunohistochemistry

Immunohistochemistry was used to stain for Matrin3 in human brain and spinal cord. Slides were de-paraffinized in xylene and rehydrated in 100% ethanol, 95% ethanol, 70% ethanol, ddH2O and quench (20 mL 30% H2O2 and 180 mL methanol) to prepare them for staining. Antigen retrieval was conducted in a pH 6 citrate buffer in a decloaking pressure chamber. Slides were blocked for 30 minutes using 2.5% horse serum. They were then incubated overnight in primary antibody solutions at 4°C. Rabbit anti-Matrin3 polyclonal antibody (Sigma MATR3 antibody: 1:800 dilution for spinal cord and 1:1600 for brain samples) was used as a
primary antibody. Slides were then incubated with the ImmPress kit horse-radish peroxidase micro-polymer solution and DNAJC7 was visualized using Diaminobenzidine (DAB) solution incubation for 10 mins. Slides were stained with hematoxylin, washed and then mounted and covered.

2.2.11. Statistical analysis:

Statistical analysis of the viability assays, western blots, and aggregation quantifications were completed using the GraphPad Prism 6 software. Statistical significance was obtained by performing unpaired t-tests to compare the means and standard deviations between the control data set and the experiment data set (each at a minimum of three biological replicas). Significance levels are indicated using asterisks, where * is p<0.05.
Chapter 3
3. Results

3.1. Matrin3 localization in motor neurons of the ventral horn in the spinal cords of 
of sALS, fALS patients, and neurologically unaffected controls

The aim of my first set of experiments is to determine Matrin3 localization in ALS patients in 
motor neurons of the spinal cord, which are the neurons most affected by neurodegeneration in 
ALS. To this end, I performed immunohistochemical analysis on paraffin-embedded formalin-
fixed spinal cord tissue of unaffected controls (n=5), sALS patients (n=11) and fALS patients 
carrying mutations in Super oxide dismutase 1 (SOD1) (n=3) or C9or72 (n=9). My analyses 
focused on motor neurons located in the anterior horn of the spinal cord. Motor neurons were 
identified by their characteristic large size compared to other neurons and glia cells and their 
overall unique morphology.

Immunohistochemical staining of unaffected controls showed weak granular nuclear staining of 
Matrin3 without any cytosolic staining, which can indicate Matrin3 binding to the chromatin or 
highly expressed RNA molecules. Immunohistochemical staining of sALS patient tissue showed 
intense nuclear staining with occasional nuclear inclusions and diffuse cytosolic staining indicating 
Matrin3 mis-localization comparable to TDP-43 in ALS. Immunohistochemical Matrin3 staining 
of fALS tissues with SOD1 mutations showed variable Matrin3 staining depending on the SOD1 
mutation: A4T and I113T mutations did not affect Matrin3 localization compared to unaffected 
controls. By contrast, in a case with a SOD1 D76Y mutation, Matrin3 showed nuclear staining 
with intense diffuse cytosolic staining. Finally, immunohistochemical staining of fALS with 
C9or72 mutations showed highly intense nuclear staining with rare cytosolic staining (Figure 9).
Figure 9: Localization of Matrin3 in the anterior horn of the human spinal cord of unaffected controls, sALS and fALS patients.

A

Ventral horn (motor neurons)

B

Control

C

SALS

D

SOD1

E

C9orf72
Table 8: Summary of the immunohistochemical staining results

<table>
<thead>
<tr>
<th>Type</th>
<th>Granular Nuclear</th>
<th>Intense Nuclear (Inclusions)</th>
<th>Diffuse Cytosolic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls (5)</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>sALS (11)</td>
<td>2</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>C9orf72 (9)</td>
<td>1</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>SOD1 A4T (1)</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SOD1 D76Y (1)</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SOD1 I113T (1)</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

A Schematic presentation of the anterior horn of the human spinal cord. B Immunohistochemistry images of the anterior horn of the human spinal cord from unaffected individuals (control). C Immunohistochemistry images of the anterior horn of the human spinal cord from sALS patients. D Immunohistochemistry images of the anterior horn of the human spinal cord from fALS patients with three different SOD1 mutations. E Immunohistochemistry images of the anterior horn of the human spinal cord from fALS patients with mutations in C9orf72.
3.2. **In-Silico analysis of Matrin3**

After confirming that Matrin3 can misfold and form inclusions and shows aberrant localization patterns in many ALS cases, we investigated the mechanism underlying Matrin3 misfolding. To address this aim, we explored three aspects that can regulate Matrin3 folding and misfolding: ALS associated mutations of MATR3, intrinsically disordered regions, and post-translational modification.

3.2.1. **ALS-associated variants of Matrin3**

We used several databases (ClinVar, Project MinE and ALSKP) to determine the sites of fALS associated mutations. We found a total of 35 different missense mutations in the MATR3 gene in fALS patients (Figure 10).
**Figure 10:** MATR3 fALS mutations.

**A**

![Diagram showing MATR3 mutations](image)

**B**  
**Table 8:** All identified ALS causing mutations in MATR3

<table>
<thead>
<tr>
<th>Mutation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S 61 N</td>
</tr>
<tr>
<td>2</td>
<td>Q 66 K</td>
</tr>
<tr>
<td>3</td>
<td>A 72 T</td>
</tr>
<tr>
<td>4</td>
<td>S 85 C</td>
</tr>
<tr>
<td>5</td>
<td>S 99 C</td>
</tr>
<tr>
<td>6</td>
<td>R 102 H</td>
</tr>
<tr>
<td>7</td>
<td>F 115 C</td>
</tr>
<tr>
<td>8</td>
<td>D 131 E</td>
</tr>
<tr>
<td>9</td>
<td>G 153 C</td>
</tr>
<tr>
<td>10</td>
<td>P 154 S</td>
</tr>
<tr>
<td>11</td>
<td>D 187 E</td>
</tr>
<tr>
<td>12</td>
<td>R 317 C</td>
</tr>
<tr>
<td>13</td>
<td>N 333 S</td>
</tr>
<tr>
<td>14</td>
<td>P 368 T</td>
</tr>
<tr>
<td>15</td>
<td>G 377 D</td>
</tr>
<tr>
<td>16</td>
<td>M 389 V</td>
</tr>
<tr>
<td>17</td>
<td>Q 390 H</td>
</tr>
<tr>
<td>18</td>
<td>G 392 V</td>
</tr>
<tr>
<td>19</td>
<td>V 394 M</td>
</tr>
<tr>
<td>20</td>
<td>R 595 G</td>
</tr>
<tr>
<td>21</td>
<td>S 596 T</td>
</tr>
<tr>
<td>22</td>
<td>S 610 F</td>
</tr>
</tbody>
</table>
A Diagram of fALS-associated variants of Matrin3 caused by missense mutations in the MATR3 gene. Three different databases (ClinVar, Project MinE and ALSKP) were employed to identify these mutations. B List of all ALS-associated amino acid changes caused by missense mutations in Matrin3 and their amino acid positions. The colours represent the predicted severity of the variant: red likely to be severely damaging; black indicates less likely to be severely damaging.
3.2.2. Posttranslational modification sites of Matrin3

Since post-translational modifications can play an important role in protein folding and often characterizes pathological inclusions in neurodegenerative disease \(^{27,83}\), I used Phosphosite plus webserver to determine the post-translational modifications of Matrin3. This analysis indicated that Matrin3 contains many serine/threonine and tyrosine phosphorylation sites. Matrin3 also contains several ubiquitination, methylation, and acetylation sites (Figure 11). The results show that most of the post translational modifications occur in the disordered regions of Matrin3 (Figure 12, C) indicating that these PTMs play a role in the folding of Matrin3.
Figure 11: Post-translational modifications of Matrin3.

A Diagram of Matrin3 PhosphoSitePlus data was plotted for the human Matrin3 entry. Low- and high-throughput numbers were summed for each site. The magnitude of the phosphorylation at residue S188 (denoted with an asterisk) was halved (the total number of references at this site is 388) to fit it within the graph. A high degree of posttranslational modifications is present in the unstructured regions (res. 1-280, and 600-847). Plotting was performed in MATLAB (4).
3.2.3. Intrinsically disordered domains of Matrin3

After using several databases to determine the fALS-causing mutations in Matrin3, we found that the mutations were scattered over the whole length of the protein, unlike the majority of the ALS associated RNA binding protein, such as TDP-43 and FUS, which contain most of their mutations in the intrinsically disordered PrLD. Since there is no evidence in the literature that Matrin3 contains a PrLD, we decided to perform computational analyses to determine if Matrin3 contains intrinsically disordered domains, which can contribute to protein misfolding. First, we generated a charge hydropathy plot using PONDR webserver, which predicted Matrin3 to be a disordered protein. The inclined black line represents the boundary line between ordered proteins (blue) and disordered protein (red). (Figure 12, A). Secondly, we used four different webservers (MFDp2, PONDR-Fit, IUPred2A, and DISOPRED3) to determine which domains of Matrin3 are disordered. We also conducted an analysis of TDP-43 as a control and reference. TDP-43 analysis shows that all the protein is highly ordered except for the carboxy-terminal PrLD (Figure 12, B). Our analysis also showed that Matrin3 protein is almost completely disordered except for the RRM1 and 2 and ZnF1 and 2 domains, which are predicted to be highly structured (Figure 12, C).
Figure 12: In silico analyses predicts Matrin3 to contain intrinsically disordered domains.
Charge-hydropathy plot predicts Matrin3 to be a mostly disordered protein. The plot was produced using the PONDR webserver. A clear boundary mostly separates the ordered (blue) and disordered proteins (red). Matrin-3 and FUS/TLS both fall on the disordered side of the plot, while TDP-43 falls on the ordered side. Plotting was performed in MATLAB. Four webservers MFDp2, PONDR-Fit, IUPred2A, and DISOPRED3 were used to predict intrinsically disordered domains for: B TDP-43 as an example of a well-described misfolded protein in ALS and C Matrin3. A confidence value below 0.5 indicates that a residue likely belongs to an ordered region, while a value above 0.5 indicates that a residue likely belongs to a disordered region. Motifs/domains are drawn and placed approximately to scale. Plotting was performed using GraphPad Prism.
3.3. Matrin3 yeast model

3.3.1. WT Matrin3

We aimed to establish and characterize a yeast model expressing human full-length, wild-type (WT) Matrin3. Therefore, I determined the toxicity and localization of human Matrin3 expressed in yeast. To this end, we established a yeast model expressing full-length WT-Matrin3 by using both a low copy (centromeric) vector with a constitutive promotor (GPD) for low expression and a high copy number 2-micron plasmids with the galactose inducible promotor (GAL) for high, inducible expression. We also fused the human full-length WT Matrin3 to the yellow fluorescent protein (Matrin3-YFP) on its carboxyl terminus to determine its localization in yeast cells by fluorescence microscopy. These Matrin3 constructs were transformed and expressed in yeast (Figure 13, A).

We performed yeast growth assay to determine the effect of Matrin3 expression on yeast cell growth. At low level of expression, WT Matrin3 shows a significant growth defect compared to the vector control. This growth defect was exacerbated with the higher expression level of the Matrin3 (Figure 13, B).

We performed western blot analyses on protein lysates from yeast cells expressing the different levels of Matrin3. Our results show that yeast cells expressing Matrin3 at high levels show a 3-fold increase in the Matrin3 levels compared to low expression cells (Figure 14).

Finally, we used fluorescence microscopy to monitor yeast cells expressing Matrin3-YFP. We observed the formation of both nuclear and cytosolic Matrin3 inclusions (Figure 15).
Figure 13: Expression of human full-length WT Matrin3 causes a growth defect in the yeast model.

**A** Diagram of Matrin3 protein. **B** Growth assay of yeast cells expressing WT Matrin3 under normal growth conditions (30°C). **C** Quantification of the spotting assays as in **A**, (n=3), **p<0.01, ****p<0.0001**, one-way ANOVA with Tukey’s post-hoc test.
Figure 14: High and low expression of WT Matrin3 in the yeast model.

A Western blot prepared with protein lysates from yeast cells expressing WT Matrin3 at different expression levels. B Quantification of western blots as shown A, (n=3), ***p<0.001, t-test with Tukey’s post-hoc test. Irrelevant areas were removed from the blot.
Figure 15: Matrin3 forms nuclear and cytosolic inclusions in the yeast model.

Fluorescence microscopy of human full-length WT Matrin3 fused to YFP (yellow fluorescent protein). A High expression  B Low expression illustrating the formation of nuclear and cytosolic inclusions of Matrin3 in the yeast cells. C Fluorescence microscopy image of WT Matrin3 at higher magnification.
3.3.2. Yeast model expressing ALS-associated Matrin3 variants

Next, we tested the effect of the ALS-associated missense mutations on Matrin3 in the yeast model. To this end, I genetically engineered the expression vectors for the two Matrin3 variants, F115C and T622A, and the corresponding carboxy-terminal YFP fusions and transformed them into yeast cells (Figure 16, A). For these experiments we only used the high expression systems (2-micron plasmid, GAL promoter).

First, we determined if these mutations lead to a change in the growth defect observed for WT Matrin3 expression in yeast. We performed growth assays with the two Matrin3 variants, which showed no significant difference when compared to WT Matrin3 (Figure 16). Then, we also performed a western blot analyses on yeast protein lysates to determine whether there are any differences in the protein levels between WT Matrin3 and the two variants. The results of the western blot demonstrate that there was no significant difference in the steady-state protein levels between the two ALS-associated variants and WT Matrin3 (Figure 17). Finally, fluorescent microscopy further confirmed no changes in the localization between the variants and WT Matrin3 YFP fusions (Figure 18).
Figure 16: Expressing ALS-associated Matrin3 variants in the yeast model.

A Schematic representation of the ALS-associated Matrin3 missense variants F115C and T622A.  

B Growth assay of yeast cells expressing each Matrin3 missense variants (F115C, and T622A) and their YFP fusion under normal growth conditions (30°C).  

C Quantification of spotting assays as in A, (n=3), ****p<0.0001, ns: non-significant, one-way ANOVA with Tukey’s post-hoc test.
Figure 17: ALS-associated MATR3 mutations do not change Matrin3 protein levels in the yeast model.

A Western blot of yeast protein lysates showing WT and Matrin3 ALS associated variants expression levels in yeast. B Quantification of the western blots as shown in A (n=3), ns: non-significant, one-way ANOVA with Tukey’s post-hoc test.
Figure 18: ALS-associated mutations of MATR3 do not affect Matrin3 localization in the yeast model.

Fluorescence microscopy image of A WT Matrin3 and ALS variants, B F115C, and C T622A illustrating the formation of nuclear and cytosolic inclusions in yeast cells expressing WT Matrin3 or ALS variants.
3.4. Matrin3 phospho-variants

3.4.1. Matrin3 phospho-variants in yeast

After establishing the Matrin3 yeast model, I investigated the effect of phosphorylation on Matrin3 toxicity and localization. To this end, I introduced amino acid substitutions at three phosphorylation sites in Matrin3; Serine 188 (S188), Serine 610 (S610) and threonine 622 (T622), using site directed mutagenesis (Figure 19, A). These three phosphorylation sites were chosen because S188A had the highest number of citations according to my in-silico analysis on Matrin3 phosphorylation, and the other two sites are two mutation sites found in fALS patients. I engineered two substitutions for each phospho-site, once with alanine, which prevents this site from being phosphorylated, and once with aspartic acid, which can mimic a phosphorylation site due to it carrying a negative charge (Figure 19, A) and transformed them into yeast cells. First, I determined if these mutations lead to a change in the growth defect caused by WT Matrin3 in yeast. We performed growth assays with the two Matrin3 phospho-variants, which showed that the growth defect caused by the S188A and S610D variants was significantly reduced compared to WT Matrin3 (Figure 19, B).

I performed western blot analyses on yeast protein lysates to determine whether there are any differences in the steady-state protein levels between WT Matrin3 and the phospho-variants. The results of the western blots demonstrated that the S188A variant showed a significant decrease in the steady-state protein levels compared to WT Matrin3, while there was no significant difference between WT Matrin3 and the remaining variants (Figure 20, A).

Finally, fluorescent microscopy demonstrated that the S610D variant showed only nuclear localization and reduced the formation of Matrin3 inclusions. By contrast, the remaining variants showed no changes in the localization compared to WT Matrin3 YFP fusion (Figure 21).
Figure 19: Matrin3 S188A and S610D mutations significantly reduced Matrin3-associated growth defect in the yeast model.
Figure 20: S188A mutation leads to Matrin3 degradation in the yeast model.

A Western blot of yeast protein lysates showing Matrin3 expression level of either WT or Matrin3 phosho-variants (S188A, S188D, S610A, S610D, T622A, T622D). B Quantification of western blots as in A, (n=3), **p<0.01, one-way ANOVA with Tukey’s post-hoc test).
Figure 21: S188A mutations leads to Matrin3 and S610D mutations inhibits Matrin3 inclusion formation in the yeast model.

Fluorescence microscopy image of A WT Matrin3 and B - G engineered phosphorylation variants (S188A, S188D, S610A, S610D, T622A, T622D) illustrating the effect of the phospho-mutations on Matrin3 localization in yeast cells.
3.4.2. Matrin3 phosho-variants in SN65 cells

To further study Matrin3 phosho-variants, I genetically engineered Matrin3 phosho-mutants into mammalian cell expression vectors and transfected them into SN56 cells. SN56 were chosen because it is mouse cholinergic septal neuronal cell line that is frequently used to study neurodegeneration. First, I performed immunofluorescence microscopy on the HA-tagged Matrin3 phosho-variants to determine the effect of the mutations on Matrin3 localization. My results showed that over-expression of WT-Matrin3 led to increased formation of nuclear inclusions compared to endogenous Matrin3 expression. Our results also showed that the S610D variant resulted in decreased inclusions formation while T622D variant showed increased mis-localization to the cytosol. (Figure 22). Second, I performed viability assays to determine the effect of the phosho-variants on Matrin3 toxicity. Our results showed that ectopic expression of WT-Matrin3 resulted in a significant reduction of the viability of SN56 cells compared to empty vector-transfected cells. Our results also showed that the three phosho-variants S610D, T622A and T622D resulted in significant reduction in cell viability compared to WT-Matrin3 (Figure 23).
Figure 22: Matrin3 phospho-variants localization in SN56 cells.

Immunofluorescence microscopy image of SN56 cells expressing A empty vector, B endogenous Matrin3 C HA-tagged-WT Matrin3 and D - I HA-tagged phospho-variants (S188A, S188D, S610A, S610D, T622A, T622D) illustrating the effect of the phospho-variants on Matrin3 localization in SN56 cells.
Figure 23: WT Matrin3 and S610D, T622A and T622D variants reduce SN56 cell viability.

Quantification of viability assay of SN56 over expressing WT Matrin3 and Matrin3 Phospho-variants, (n=4), **p<0.01, ****p<0.0001, one-way ANOVA with Tukey’s post-hoc test).
3.5. Analysis of the intrinsically disordered domains of Matrin3

3.5.1. Intrinsically disordered truncated fragments of Matrin3 in yeast

To determine the contribution of the intrinsically disordered amino-terminal domain on Matrin3 toxicity and localization, we genetically engineered expression vectors of the following Matrin3 amino-terminal truncated fragments 100, 150, 200, 287 and 399 and transformed them into yeast. The numbers indicate the size of each truncated fragment (Figure 24, A). First, I conducted yeast growth assay to determine if expression of the truncated fragments caused growth defects similar to expression of WT full length Marin3 in yeast. My results show that the truncated fragments caused a significant growth defect compared to the negative (Empty vector) control, however the growth defect associated with the expression of all tested truncated fragments is significantly reduced compared to full-length WT Matrin3 (Figure 24, B).

Then, I performed a western blot analyses on yeast protein lysates to determine whether there are any differences in the protein levels between full-length WT Matrin3 and the truncated fragments. The results of my western blots analyses demonstrate that the steady-state protein levels of the 100 and 150 fragments are significantly lower than WT Matrin3. Furthermore, my results show that the steady-state protein level of the 399 fragment is significantly higher compared to WT Matrin3 (Figure 25, A).

My fluorescence microscopy of yeast cells expressing Matrin3 amino-terminal 399 fragment show numerous and small inclusions compared to full-length WT Matrin3. These results also show that introducing F115C mutation to the truncated fragment resulted in bigger and less numerous inclusions in yeast (Figure 26).
Figure 24: Amino-terminal truncated fragments of Matrin3 caused a significant growth defect in the yeast model compared to the empty vector.

A

<table>
<thead>
<tr>
<th>Fragment</th>
<th>WT Matrin3</th>
<th>N. Terminal 100</th>
<th>N. Terminal 150</th>
<th>N. Terminal 200</th>
<th>N. Terminal 287</th>
<th>N. Terminal 399</th>
<th>N. Terminal 287 F115C</th>
<th>N. Terminal 399 F115C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Relative growth</strong></td>
<td>0.6</td>
<td>1.0</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

B

Figure: Matrin3 N. Terminal 399 (-RRM1) aa fragment shows a growth defect in the yeast model compared to the empty vector.
A Schematic representation of the engineered Matrin3 amino-terminal truncated fragments. B Growth assay of yeast cells expressing either WT Matrin3 or Matrin3 truncated fragments under normal growth conditions (30°C). C Quantification of spotting assays as in A, (n=3), **p<0.01, ***p<0.001, ****p<0.0001, one-way ANOVA with Tukey’s post-hoc test).
Figure 25: Amino-terminal 399 aa fragment shows a significant increase in the steady-state protein levels compared to WT Matrin3.

A Western blot of yeast protein lysates showing HA-tagged full-length WT Matrin3 and amino-terminal truncated fragments in yeast cells. Arrow heads indicate bands that specifically detect Matrin3 and its fragments. B Quantification of the western blot in A, (n=3), ns: non-significant, *p<0.05, **p<0.01, ****p<0.0001 one-way ANOVA with Tukey’s post-hoc test.
Figure 26: F115C mutation changes the localization Matrin3 amino-terminal 399 fragment in yeast.

Fluorescence microscopy image yeast cells expressing A WT-Matrin3 and B Matrin3 amino-terminal 399 Fragment C Matrin3 F115C amino-terminal 399 fragment illustrating the localization of WT and Matrin3 F115C amino-terminal 399 truncated fragment in yeast.
3.5.2. **Intrinsically disordered truncated fragments of Matrin3 expressed in SN56 cells**

To further confirm our yeast results, I genetically engineered Matrin3 amino-terminal truncated fragment 399 into mammalian cell expression vectors and transfected them into SN56 cells. These constructs also contain carboxy-terminal HA epitope-tag fusion for detection by western blotting. First, I performed immunofluorescence microscopy on the HA-tagged Matrin3 399 fragment to determine its localization compared to WT Matrin3. My results show that over-expression of the 399 fragment led to the formation of both nuclear and cytosolic inclusions (Figure 27).

Second, I performed viability assays to determine the effect of the amino-terminal truncated fragment on cellular viability. My results show that the expression of the 399 fragment results in a significant reduction in the viability of the SN56 cells compared to empty vector controls. This decrease in cell viability was not significantly different from cells expressing full-length WT Matrin3 (Figure 28).
Figure 27: amino-terminal 399 truncated fragment leads to the formation of nuclear and cytoplasmic inclusions in SN56 cells.

Immunofluorescence microscopy image of A HA-tagged full-length WT Matrin3 and B HA-tagged 399 fragments illustrating Matrin3 localization in SN56 Cells (Passage number = 10).
Figure 28: Overexpression of the amino-terminal 399 truncated fragment significantly reduces SN56 cells viability.

Viability assay quantification of SN56 over expressing full-length WT Matrin3 and Matrin3 amino-terminal 399 fragment (n=5), *p<0.05, one-way ANOVA with Tukey’s post-hoc test.
4. Discussion

4.1. Aberrant Matrin3 localization in motor neurons of the anterior horn in the spinal cords of ALS patients

In this study, I investigated Matrin3 localization in motor neurons in the anterior horn of the spinal cord in unaffected controls, and patients with sALS, and fALS carrying mutations in either C9orf72 or SOD. I found that Matrin3 shows granular nuclear staining, which can indicate binding to the chromatin or specific RNA molecules, in unaffected control without any cytosolic staining (Figure 9, B) as reported before. In ALS patients, Matrin3 shows changes in its localization compared to unaffected controls. There are, however, variation of this aberrant Matrin3 localization depending on the individual patient and the type of ALS.

In sALS, we observed four major discernable patterns of Matrin3 mis-localization: 1) intense nuclear staining without cytosolic staining; 2) intense nuclear staining with weak cytosolic staining; 3) complete mis-localization of Matrin3 from the nucleus showing diffuse cytosolic staining; and 4) Matrin3 cytosolic inclusions (Figure 9, C). In fALS, Matrin3 mis-localization varies according to the gene carrying the mutations, and different mutations within the same gene. Patients with C9orf72 mutations showed a very intense nuclear staining with very rare cytosolic mis-localization (Figure 9, E). In SOD1 patients, A4T and I113T mutations did not affect Matrin3 localization compared to neurologically unaffected controls. Yet, ALS patient carrying the D76Y SOD1 mutation showed increased Matrin3 nuclear localization with intense diffuse cytosolic staining (Figure 9, D).

Our IHC results show that Matrin3 mis-localization in patients with different types of ALS varies from patient to patient. Our results on Matrin3 confirmed previous studies showing Matrin3 pathology in patients with sALS and fALS carrying mutations in C9orf72. Yet, of note, we detected aberrant Matrin3 localization in the majority of the explored ALS specimen expect for the SOD1 mutants A4T and I113T. This indicates that Matrin3 displays a pathological localization pattern in sALS and at least in some cases of fALS. While it remains unknown whether this aberrant Matrin3 localization indeed contributes to ALS pathogenesis, my pathological results suggest that Matrin3 indeed has the capacity to mis-localize and form inclusion as it has been observed for other ALS-associated proteins, such as TDP-43. My results from patients with SOD1 mutations could furthermore indicate that Matrin3 mis-localization does not occur in all cases with SOD1 mutations, but it could be specific to the D76Y mutations. We speculate that this
mutation leads to SOD1 dysfunction and the development of severe oxidative stress inside the cell, which leads to Matrin3 mis-localization and inclusion formation. Our results on Matrin3 pathology in patients carrying mutations in SOD1 is novel and even studies focused on other RNA binding proteins, such as TDP-43 and FUS, did not show any proteinopathy in SOD1 patients and SOD1 inclusions formation is highly controversial in sALS.\textsuperscript{105,106} In any case, my results indicate a possible specific link between SOD1 and Matrin3 that is distinct from other ALS-associated RNA binding proteins.

There are some limitations to our experiment that need to be considered. First, the limited number of specimens, especially in the fALS cases with SOD1 mutations. We need to increase the number of SOD1 mutations in our study and include more patients for each mutation to get statistically significant results. Experiments should also be done to determine the effect of those mutations on the structure and functions of SOD1 and their effect on Matrin3 pathology.

Second, we only used one anti-Matin3 antibody (Sigma Anti-MATR3 antibody). Studies have shown that not all antibodies can detect Matrin3 inclusion. This was indicated in a previous study in which 4 different Matrin3 antibodies were used ALS patient slides, and only one anti-Matin3 antibody (Bethyl laboratories Anti-MATR3 antibody) detected Matrin3 inclusions. The Bethyl antibody in future studies since it is more sensitive to Matrin3 inclusions.

Finally, we did not know the genetic makeup of the spinal cord sporadic ALS (sALS) patient samples were utilized in this study. We don’t know whether they have mutations in any of the ALS causing genes. We also were not able to obtain tissue sample from fALS patients with mutations in Matrin3 to use as a positive control in our study as done before.\textsuperscript{44} Finally, we determined the motor neuron cell population based on their size and morphology. To be more accurate, further studies should be performed by co-staining the protein of interest with specific motor neuron markers such as Motor Neuron and Pancreas Homeobox Protein 1 (MNX1), a nuclear protein that has been used a reliable marker for motor neurons in previous studies.

Nevertheless, our results showed that Matrin3 shows mis-localization in patients with different types of ALS. Future work includes increasing the number of unaffected controls, sALS and fALS carrying mutations in different genes (SOD1, TDP-43, FUS and C9orf72) cases to generate more representative results. Future work should also include tissue samples with other neurodegenerative disorders, such as Parkinson’s and Alzheimer’s disease, to determine whether Matrin3 contributes to the pathogenesis of neurodegeneration, generally, or is it only specific to
ALS. Also, different anti-Matin3 antibodies should be used on tissue sample to make sure that protein inclusions can be detected with different reagents and by recognitions of different regions (antigens) within Matrin3. Using conformation specific antibodies, which can recognize specific three-dimensional antigenic conformation can give more insight on Matrin3 changes occurring in ALS. Finally, subcellular fractionation should be performed to separate nuclear and cytosolic Matrin3. This assessment combined with the use of different detergents and protein lysis methods will further elucidate the biochemical nature of Matrin3 inclusions.

4.2. **Matrin3 yeast model**

Several yeast models have been established to study protein misfolding in neurodegenerative disorders, such as TDP-43 yeast model in ALS. TDP-43 expression leads to a growth defect and the formation of cytosolic soluble aggregates when expressed in yeast cells. Finally, TDP-43 growth defects depend on the level of protein expression, making it an excellent candidate for high-throughput screens to identify genetic interactions that can modulate its toxicity. We established a yeast model for Matrin3 to investigate the mechanisms underlying Matrin3 misfolding and toxicity. Similar to TDP-43, our yeast results show that Matrin3 expression in yeast leads to a growth defect in a dose dependent manner i.e., the higher level of Matrin3 expression, the more prominent the growth defect (Figure 13, B and C). As in human neurons, our results show that Matrin3 overexpression led to the formation of both nuclear and cytosolic inclusions in yeast cells (Figure 15).
4.3. **ALS-Associated Matrin3 Variants**

We transformed the most frequently studied ALS-associated Matrin3 variants, F115C and T622A, into yeast. Expression of these Matrin3 ALS variants, similar to full length WT Matrin3, causes a significant growth defect in yeast cells (Figure 16, B and C). Our western blot and fluorescence microscopy do not show any changes in the steady state protein levels of the variants, or their localization compared to WT Matrin3 (Figure 17 and 18). This indicated that the two ALS-associated variants might not drastically change Matrin3 misfolding in the yeast model. Our results recapitulate finding from previous studies conducted on Matrin3 in mammalian cells showing that the ALS associated variants did not show different localization patterns compared to WT Matrin3.

Yeast models are a powerful genetic tool to study basic cellular processes, protein quality control and the mechanisms of protein folding and misfolding. Several yeast models have been established to study protein misfolding in neurodegenerative disorders, such as TDP-43 yeast model in ALS. The conservation of key cellular pathways between yeast and humans, such as the cellular protein quality control, makes experiments in yeast directly relevant to human and can be used to decipher the mechanisms of the protein misfolding in neurodegenerative diseases. However, similar to any other model system, yeast has limitations that need to be carefully considered. Yeast is a single-celled organism that does not have differentiated cell types such as the highly differentiated motor neurons that are affected in ALS. Furthermore, yeast does not have an ortholog for Matrin3, which also means that naturally occurring protein interactions with Matrin3 do not exist in yeast cells. This poses a limitation since the yeast protein quality control system and molecular chaperones are not normally involved in the processing of the Matrin3 protein. One way to address this limitation was discussed before in the literature, which is to engineer yeast to possess entirely human chaperone pathways. This can help producing results that are more relevant to mammalian cells. Finally, many proteins that interact with Matrin3 and help carry its functions and possibly also maintain its proper, non-toxic conformation do not have orthologs in yeast. This explains why we cannot perform experiments exploring Matrin3 function in the yeast model and might induce over Matrin3 misfolding. Yet the yeast model, like the aforementioned models, allows us testing Matrin3 misfolding and whether this leads to loss or gain of toxic function of the protein.
4.4. **Matrin3 phosphorylation**

Matrin3 is a phosphoprotein with several serine/threonine and tyrosine phosphorylation sites. Here, I used Phosphositeplus to determine the phosphorylation sites in Matrin3. I found 34 serine, 4 threonine and 3 tyrosine residues in Matrin3 that are phosphorylated (Figure 11). Some results from our *in-silico* analysis are obtained from phosphorylation prediction webservers and mass spectrometry on Matrin3 isolated from cancer cells. These webservers depend on certain algorithms or machine learning, which is prone to over-optimization when developed on small data sets. This poses a limitation since some of it can reduce the accuracy of the predicted results, which have to be confirmed by individual experiments in future studies. In any case, Phosphositeplus provides the number of times each site was mentioned in the literature and the **serine 188 (S188)** site has the greatest number of citations, which indicates that this phosphorylation occurs frequently under many different experimental conditions and in different systems. Our analysis also indicated that two additional sites, **serine 610 (S610)** and **threonine 622 (T622)**, are frequently found to be phosphorylated (Figure 11). I therefore focused on these three sites for my project. This focus was based on experimental feasibility. It is possible that there are phosphorylation events than the ones assessed here, and that other posttranslational modifications, such as ubiquitination, acetylation, and Sumoylation, can also alter Matrin3 misfolding and toxicity as show for other misfolded proteins before.\(^\text{108,109}\) These will need to be explored in future studies.

Our results show that the **S188A** mutant variant is expressed at very low steady-state levels compared to wild-type Matrin3. This can be due to either that the mRNA is not expressed, or the mutant proteins undergo rapid proteasomal degradation or degradation by autophagy. By contrast, the **S188D** variant did not have any effect on the growth defect caused by Matrin3 in the yeast model. This can indicate that the negative charge carried by the by the aspartic acid, and us also by a phosphate group, might have a significant role in the folding process of Matrin3 (Figure 19 and 20). Furthermore, our results **S610D** mutations reduced the growth defect caused by Matrin3, the mutations did not change the steady state levels (Figure 19 and 20) and fluorescence microscopy results show that the **S610D** mutation changed Matrin3 localization pattern in the yeast model (Figure 21). The **S610D** variant showed only nuclear localization without forming any nuclear or cytosolic inclusions. This could also provide a link between the growth defect caused
by Matrin3 and inclusion formation. The drastic changes caused by the mutation on Matrin3 toxicity and localization indicate that the S610 phosphorylation site is significant for the proper folding and stability of Matrin3.

We then confirmed our yeast results in SN56 cells. Immunofluorescence microscopy showed that over expression of full-length WT Matrin3 led to nuclear inclusion formation (Figure 22 and 24). Our results showed that S610D variant showed decreased inclusion formation, which is similar to our yeast results further confirming that the phosphate group is required for proper Matrin3 folding.

We also noticed that the T622D variant showed increased mis-localization to the cytosol. I did not notice the same results with the T622A variant indicating a possible role for the phosphorylation of this site in Matrin3 folding (Figure 22). To better understand the effects of the mutations on Matrin3 folding, biochemical approaches, such as vibrational circular dichroism, dynamic light scattering, and analytical ultracentrifugation should be used on purified WT and Matrin3 variants in biophysical experiments.

There is evidence that phosphorylation plays a major role in the pathogenesis of other RNA binding proteins in ALS such as TDP-43 and FUS. Studies suggest that phosphorylation of the PrLD of FUS disrupts its ability to phase separate, aggregate and significantly reduces its toxicity. Studies have also shown that, in ALS, the PrLD of TDP-43 is heavily phosphorylated and disrupting TDP-43 at specific sites was neuroprotective. Taken together, the role of protein phosphorylation in ALS pathogenesis is still controversial and further studies are needed to further confirm the contribution of protein phosphorylation in ALS.

Although further studies are required to elucidate the exact role of phosphorylation on Matrin3 pathogenesis, our results still suggest a significant role of protein phosphorylation on Matrin3 misfolding and pathogenesis.

Even though many human genes have orthologs in yeast, there are many human genes that do not have yeast orthologs. This is particularly relevant to my study since there are differences between the human and yeast kinases. This also poses a limitation on the yeast model since Matrin3 is not an endogenous yeast protein and yeast kinases might not interact with it. Thus, it remains uncertain whether Matrin3 is phosphorylated in yeast or not, and if it is, the phosphorylation pattern might be different compared to mammalian cells. Future experiments using phospho-specific anti-
Matrin3 antibodies and directed mass spec experiments using our yeast model and comparing those results to mammalian systems, will address these open questions.

Finally, to study the effect of phosphorylation on Matrin3 misfolding, we introduced amino acid substitutions. In addition to changing the phosphorylation of Matrin3, these amino acid substitutions change the size of side chain of the amino acid at this particular site; for example, replacing the serine with alanine removes the hydroxyl group in the side chain and using the aspartic acid can only mimic a phosphate group regarding its negative charge, but it cannot be considered a genuine phosphorylated site due to the absence of the phosphate group with its large hydration size and charge compare the carboxyl group of aspartic acid. Future in vitro studies using purified phosphorylated Matrin3 or Matrin3 fragments will allow overcoming these limitations. Nevertheless, the amino acid substitutions we used here have given profound insights into how phosphorylation can alter the three-dimensional conformation and function of many different proteins and might thus also be informative for Matrin3.

We replicate our finding in yeast in SN56 mammalian cell line. Mammalian neuronal cells lines generally represent great models to study human diseases and they helped advance our understanding of protein misfolding and protein quality control in neurodegenerative diseases. SN56 in particular is a very well-established model system to study ALS since they are cholinergic septal neurons which makes it very close to motor neurons affected in ALS. However, SN56 cells are still different from human motor neurons since they are derived from mice. SN56 cells are neuroblastoma cells, so they have to be regarded as cancer cells, which differ profoundly from permanently differentiated, non-diving motor neurons that are affected by ALS. Unlike motor neurons which have a finite life span, immortalized cell lines can divide indefinitely, and sometimes express unique gene patterns not found in any normal cell types in the CNS. To overcome this limitation, future studies should be conducted on primary neurons, or induced pluripotent stem cells (iPS cells) derived from ALS patients.

Importantly, to confirm whether there are indeed changes in Matrin3 phosphorylation in ALS, mass spectrometry should be conducted on misfolded Matrin3 isolated from spinal cord samples of ALS patients and compared to unaffected controls, e.g., by immunoprecipitation. In addition, in vitro experiments should be conducted purified WT and Matrin3 phospho-variants to understand how the mutations affect Matrin3 folding.
4.5. **The intrinsically disordered domains of Matrin3**

Matrin3 is mostly an intrinsically disordered protein, and it can undergo LLPS under certain conditions as shown by my computational analyses and one published study.\(^{100}\) We used PONDR webserver to generate a charge hydropathy plot, which can predict whether the protein is ordered or not based on the amino acid sequence and the total charge of the protein (Figure 12, A). Furthermore, we used four different webservers to determine which regions in Matrin3 are disordered. Our analysis indicate that Matrin3 is almost completely disordered except for the well identified RRM and ZnF domains (Figure 12, C).

Our *In-silico* analysis results were solely obtained from algorithm-based predictions and are not experimentally confirmed. *In-vitro* analyses, such as circular dichroism and nuclear magnetic resonance, should be performed on purified Matrin3 or Matrin3 fragments to demonstrate it intrinsically disordered character. Future *in vitro* experiment will also allow assessing whether the amino-terminal Matrin3 domain undergoes LLPS.

We decided to focus on the amino-terminal intrinsically disordered regions since it is the largest continuous disordered region of Matrin3, contains the majority of the ALS-causing variants, and previous studies indicated that it plays a role in liquid-liquid phase separation.\(^ {100}\) Our results show that all the amino-terminal truncated fragments used in my study (*100, 150, 200, 287 and 399*) caused a significant growth defect in yeast, however, the growth defect was not as strong as the one caused by full length WT Matrin3 (Figure 24, A and B).

Our western blot results show that the steady state protein levels of the 100 and 150 amino acid truncated fragments were significantly lower than full length WT-Matrin3 (Figure 25). We speculate this is due to a decrease in the stability of these fragments or an increase in their degradation. Furthermore, our results show significant increase in the *399* fragment protein levels compared to full-length WT-Matrin3, which indicates increased stability of the fragment, possibly because of its resistance to degradation. Our fluorescence microscopy of yeast cells expressing Matrin3 amino-terminal *399* fragment show numerous and small inclusions compared to full-length WT-Matrin3, which mostly formed one big inclusion inside the cell. Based on these results we speculate that the amino-terminal *399* fragment undergoes LLPS, which can go awry under cellular stress conditions or due to mutations leading to protein misfolding and toxicity. We also
observed that introducing the ALS-associated F115C amino acid substitution to the truncated fragment resulted in drastic changes in protein localization, which indicates that the ALS-associated mutation can affect the LLPS separation of the protein leading its misfolding (Figure 26).

Finally, we confirmed our yeast results in SN56 cells. Immunofluorescence microscopy showed to the formation of nuclear and rare cytosolic inclusions in transfected cells. These inclusions were smaller and more numerous Full-length WT-Matrin3 (Figure 27). There results were, to some extent, similar to our yeast results. Finally, the viability assay results showed that transfection with the amino-terminal 399 fragment resulted in a significant decrease in SN56 cell viability (Figure 28).

Previous studies on TDP-43 have shown that expressing only the carboxy-terminal PrLD did not lead to any growth defect, however expressing the PrLD fused to the RRM resulted in a significant growth defect. Yeast fluorescence microscopy of the GFP tagged PrLD of TDP-43 showed diffuse signaling with very rare cytosolic inclusions, but the PrLD-RRM truncation resulted in numerous cytosolic inclusions in the yeast model. These results indicate a significant role of the RNA binding in TDP-43 toxicity and inclusion formation.

Our results with Matrin3 differ TDP-43. Since all the fragments lack any RNA binding motifs, our results provide evidence that, unlike other RNA binding proteins, such as TDP-43, the RNA binding ability in Matrin3 is not required to produce a toxic effect, at least in yeast models. We speculate that the amino-terminal domain undergoes phase separation into liquid droplets where it misfolds and forms inclusions, which might contribute to the toxic effect we observe in yeast and neuronal cells.
4.6. Overall conclusions and significance

Our results confirms that Matrin3 misfolds in sALS and C9orf72 patients. This thesis thus provides the first evidence of Matrin3 mis-localization in patients with SOD1 mutations. We also established a Matrin3 yeast model to investigate the mechanisms of matrin3 misfolding. Our results signify that protein phosphorylation regulates Matrin3 folding and changes in the phosphorylation can lead to protein misfolding and inclusion formation. Our results also show that amino-terminal disordered regions lead to inclusion formation and toxicity on its own without binding to RNA. This is a novel finding that can help us understand how intrinsically disordered regions drive protein misfolding in different neurodegenerative disorders. **We propose that the protein phosphorylation and the amino-terminal intrinsically disordered region plays a major role in Matrin3 toxicity and misfolding in ALS.** Finally, our study highlights the contribution of protein phosphorylation and intrinsically disordered regions on protein toxicity and misfolding in neurodegenerative disorders. Understanding the role of these two concepts in protein misfolding is an essential step to be able to develop more effective therapeutic approaches for different neurodegenerative disorders, not just ALS.
Figure 29: Proposed mechanism of Matrin3 misfolding in ALS.
A Under normal conditions, Matrin3 is located inside the nucleus. B ALS-associated mutations in the regions encoding the disordered regions of Matrin3 leads to protein misfolding, mis-localization and the formation of nuclear and cytosolic inclusions. C In ALS, changes in Matrin3 phosphorylation pattern may lead to protein misfolding, mis-localization and the formation of nuclear and cytosolic inclusions. We propose that the ALS associated mutations and changes in protein phosphorylation led to protein misfolding by causing defects in the liquid-liquid phase separation of Matrin3.
References


List of Appendices:
Appendix A: Supplementary figures

Figure 30: Localization of Matrin3 in the anterior horn of the human spinal cord of unaffected controls, sALS and fALS patients.

A  Control

B  sALS

C  C9orf72

A Immunohistochemistry images of the anterior horn of the human spinal cord from unaffected individuals (control). B Immunohistochemistry images of the anterior horn of the human spinal cord from sALS patients C Immunohistochemistry images of the anterior horn of the human spinal cord from fALS patients with mutations in C9orf72.
Figure 3: Yeast growth assay non-induced controls

Growth assay of yeast cells transformed with A Full length WT Matrin3 B Full length WT Matrin3 and ALS associated variants C Full length WT Matrin3 and Matrin3 phospho-variants D Full length WT Matrin3 and Matrin3 amino-terminal truncated fragments under normal growth conditions (30°C).
Figure 32: Complete western blot for figure 16.

A Western blot prepared with protein lysates from yeast cells expressing WT Matrin3 at different expression levels

B Western blot showing the protein expression levels of PGK-1 control in yeast cells expressing WT Matrin3.
**Figure 33:** Complete western blot for **figure 22**.

A

![Western blot image](image1)

B

![Western blot image](image2)
A Western blot prepared with protein lysates from yeast cells expressing WT Matrin3 and Matrin3 phospho-variants B Western blot showing the protein expression levels of PGK-1 control in yeast cells expressing WT Matrin3 and Matrin3 phospho-variants.
Figure 34: Complete western blot for figure 26.

Western blot showing the protein expression levels of PGK-1 control in yeast cells expressing Full length WT Matrin3 and Matrin3 amino terminal truncated fragments.
Curriculum Vitae

Name: Ahmed Salem

Post-Secondary Education and Degrees:
University of Western Ontario
London, Ontario, Canada

MSc, Anatomy and Cell Biology
2019-2021

BMSc, Honors specialization in Pathology and laboratory medicine
2016-2019

Honours and Awards:

2016 Transfer Admission Scholarship – University of Western Ontario
2017 Deans Honor List – University of Western Ontario
2018 Deans Honor List – University of Western Ontario
2019 Deans Honor List – University of Western Ontario
2020 Anatomy and Cell biology Travel Award – University of Western Ontario

Related Work Experience:

2017-2018 Teaching Assistant – Helping hands
University of Toronto
Presentations:


Ahmed Salem and Martin Duennwald. 2019. DNAJC7 in ALS. Matrin3 misfolding and the domains mediating its toxicity in Amyotrophic lateral Sclerosis. Pathology 4980 research talk. The University of Western Ontario. London, Ontario. (Research talk)


