The Importance Of Protein Context In Spinocerebellar Ataxia Type 3

Sean Luis Johnson
Wayne State University

Follow this and additional works at: https://digitalcommons.wayne.edu/oa_dissertations

Part of the Biology Commons, Molecular Biology Commons, and the Neurosciences Commons

Recommended Citation
Johnson, Sean Luis, "The Importance Of Protein Context In Spinocerebellar Ataxia Type 3' (2022). Wayne State University Dissertations. 3634.
https://digitalcommons.wayne.edu/oa_dissertations/3634

This Open Access Dissertation is brought to you for free and open access by DigitalCommons@WayneState. It has been accepted for inclusion in Wayne State University Dissertations by an authorized administrator of DigitalCommons@WayneState.
THE IMPORTANCE OF PROTEIN CONTEXT IN SPINOCEREBELLAR ATAXIA
TYPE 3

by

SEAN LUIS JOHNSON

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2022

MAJOR: PHARMACOLOGY

Approved By:

________________________________________
Advisor

________________________________________
Date
DEDICATION

This work is dedicated to Madeline.

You have been my inspiration, my motivator, and my welcome distraction through every up, down, and in-between of this journey. This wouldn’t have happened without your incredible patience and support, and I am beyond lucky to have you as a partner.
ACKNOWLEDGEMENTS

The first person I would like to acknowledge is Dr. Sokol Todi. You took a chance on a student who struggled a bit in grades but had a passion for research. You have always pushed me to be better and because of who you are as a leader and teacher I’ve always felt eager to meet the challenges you presented me. You made it all feel possible, even when I had my own doubts, and I would not be here writing this if it were not for you. In your lab, you make sure that every voice in the room is heard and by doing so allow us all to test ideas, make mistakes, and ultimately become better scientists. Under your tutelage I have grown immensely as a scientist, a thinker, and a person. I could not have asked for a better mentor and will be forever grateful.

I also want to thank the entire Todi lab. So much of what I have learned has come from random conversations around the lab and the comradery shared in failed experiments. A particular thanks to Dr. Wei-Ling Tsou who has played an instrumental role in my lab education and often felt like a second mentor.

Thank you to my committee members, Dr. Michael Bannon, Dr. Andrew Garrett, Dr. Izabela Podgorski, and Dr. Robert Wessells. I have learned so much from all your questions, advice, and support and in doing so have been able to think faster on my feet and approach problems from an improved critical perspective. In addition, I’d like to thank the entire Pharmacology Department, especially my fellow graduate students, for the endless support throughout my PhD training.

To my family, I find it hard to put into the words how important you have been in this process. There was never a moment, in this or any other undertaking, where I felt like you weren’t there for me. You are responsible for the person that I am today and I hope that I can return even an ounce of the support that you have always given me in my endeavors.
# TABLE OF CONTENTS

**DEDICATION** ............................................................................................................................................... ii

**ACKNOWLEDGEMENTS** .................................................................................................................................. iii

**TABLE OF CONTENTS** ................................................................................................................................... iv

**LIST OF FIGURES** .......................................................................................................................................... ix

**LIST OF ABBREVIATIONS** ................................................................................................................................ xii

1 **CHAPTER 1: THE PROTEIN CONTEXT OF POLYGLUTAMINE DISEASES** ... 1

1.1 **Introduction** .............................................................................................................................................. 1

1.1.1 Dentatorubral-Pallidoluysian Atrophy (DRPLA) ......................................................................................... 2

1.1.2 Huntington’s Disease (HD) .......................................................................................................................... 3

1.1.3 Spinal and Bulbar Muscular Atrophy (SBMA or Kennedy’s Disease) ......................................................... 5

1.1.4 Spinocerebellar Ataxia Type 1 (SCA1) ........................................................................................................... 8

1.1.5 Spinocerebellar Ataxia Type 2 (SCA2) ........................................................................................................... 9

1.1.6 Spinocerebellar Ataxia Type 3 (SCA3 or Machado Joseph Disease) .......................................................... 11

1.1.7 Spinocerebellar Ataxia Type 6 (SCA6) ........................................................................................................ 13

1.1.8 Spinocerebellar Ataxia Type 7 (SCA7) ........................................................................................................ 15

1.1.9 Spinocerebellar Ataxia Type 17 (SCA17) ..................................................................................................... 17

1.2 **Advancements in establishing disease protein context** ......................................................................... 19

1.2.1 Dentatorubral-Pallidoluysian Atrophy (DRPLA) ......................................................................................... 19

1.2.2 Huntington’s Disease (HD) .......................................................................................................................... 21

1.2.3 Spinobulbar Muscular Atrophy (SBMA or Kennedy’s Disease) ............................................................ 24

1.2.4 Spinocerebellar Ataxia Type 1 (SCA1) ........................................................................................................... 27

1.2.5 Spinocerebellar Ataxia Type 2 (SCA2) ........................................................................................................... 29

1.2.6 Spinocerebellar Ataxia Type 3 (SCA3 or Machado Joseph Disease) ........................................................ 31

1.2.7 Spinocerebellar Ataxia Type 6 (SCA6) ........................................................................................................ 32

1.2.8 Spinocerebellar Ataxia Type 7 (SCA7) ........................................................................................................ 33
1.2.9 Spinocerebellar Ataxia Type 17 (SCA17) ................................................................. 35

1.3 Conclusions .................................................................................................................. 38

2 CHAPTER 2: DIFFERENTIAL TOXICITY OF ATAXIN-3 ISOFORMS IN DROSOPHILA MODELS OF SPINOCEREBELLAR ATAXIA TYPE 3 .................. 39

2.1 Introduction ................................................................................................................. 39

2.2 Materials and Methods .............................................................................................. 41
  2.2.1 Constructs and plasmids ..................................................................................... 41
  2.2.2 Drosophila stocks and new transgenic lines .................................................... 42
  2.2.3 Longevity and motility assays .......................................................................... 43
  2.2.4 Mammalian cells and assays ............................................................................. 43
  2.2.5 Western blotting ................................................................................................. 44
  2.2.6 Differential Centrifugation ................................................................................ 44
  2.2.7 Filter-Trap Assay .............................................................................................. 45
  2.2.8 Nuclear/Cytoplasmic Extraction ....................................................................... 45
  2.2.9 Antibodies ......................................................................................................... 45
  2.2.10 qRT-PCR ......................................................................................................... 46
  2.2.11 Statistics ............................................................................................................ 46

2.3 Results .......................................................................................................................... 46
  2.3.1 New Drosophila lines to investigate SCA3 ......................................................... 46
  2.3.2 Isoform 2 of pathogenic ataxin-3 is less toxic than isoform 1 in Drosophila .......... 50
  2.3.3 Aggregation of isoforms 1 and 2 of pathogenic ataxin-3 in flies ....................... 54
  2.3.4 Isoform 2 at protein levels comparable to isoform 1 is no less toxic in Drosophila .... 56
  2.3.5 Degradation of ataxin-3 by autophagy and the proteasome in Drosophila .......... 58
  2.3.6 Isoform 2 is turned over more rapidly than isoform 1 in mammalian cells .......... 61
  2.3.7 The role of the C-terminus on the stability of isoform 2 .................................... 64

2.4 Discussion ..................................................................................................................... 66
3  CHAPTER 3: UBIQUITIN-INTERACTING MOTIFS OF ATAXIN-3 REGULATE ITS POLYGLUTAMINE TOXICITY THROUGH HSC70-4-DEPENDENT AGGREGATION ................................................................. 70

3.1  Introduction ......................................................................................................................... 70

3.2  Materials and Methods ...................................................................................................... 72
  3.2.1  Antibodies .................................................................................................................. 72
  3.2.2  Drosophila materials and procedures ......................................................................... 72
  3.2.3  Western blotting and quantification ........................................................................... 73
  3.2.4  Filter-trap assay ......................................................................................................... 74
  3.2.5  Soluble/pellet centrifugation ..................................................................................... 74
  3.2.6  Nuclear/cytoplasmic separation ................................................................................ 74
  3.2.7  Co-immunopurification ............................................................................................ 74
  3.2.8  Mass spectrometry .................................................................................................... 75
  3.2.9  Histological preparation ............................................................................................ 75
  3.2.10 Statistical analyses .................................................................................................. 75

3.3  Results ............................................................................................................................... 76
  3.3.1  The modulatory role of ataxin-3’s domains on its toxicity ........................................... 76
  3.3.2  UIMs enhance toxicity of the isolated polyQ of ataxin-3 ............................................. 79
  3.3.3  UIMs of full-length ataxin-3 regulate toxicity and aggregation ............................... 81
  3.3.4  UIMs mediate ataxin-3 interaction with Hsc70-4 ....................................................... 84
  3.3.5  Hsc70-4 enhances aggregation and toxicity of pathogenic ataxin-3 ......................... 85
  3.3.6  Hsc70-4 knockdown suppresses polyQ toxicity more generally ............................... 88
  3.3.7  UIMs are important for inter-ataxin-3 interactions ................................................. 89

3.4  Discussion ......................................................................................................................... 91

4  CHAPTER 4: TARGETING THE VCP-BINDING MOTIF OF ATAXIN-3 IMPROVES PHENOTYPES IN DROSOPHILA MODELS OF SPINOCEREBELLAR ATAXIA TYPE 3 .................................................. 96

4.1  Introduction ....................................................................................................................... 96

4.2  Materials and Methods .................................................................................................... 98


<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.9</td>
<td>Co-immunopurification assays</td>
<td>128</td>
</tr>
<tr>
<td>5.2.10</td>
<td>Statistics</td>
<td>129</td>
</tr>
<tr>
<td>5.3</td>
<td>Results</td>
<td>129</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Development of new constructs to examine RAN in SCA3</td>
<td>129</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Homomeric and alternating pathogenic ataxin-3 polyQ repeats are differentially toxic in Drosophila melanogaster</td>
<td>131</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Lack of evidence of RAN peptides in Drosophila from pathogenic ataxin-3</td>
<td>134</td>
</tr>
<tr>
<td>5.3.4</td>
<td>New transgenes to investigate mRNA-based toxicity in Drosophila models of SCA3</td>
<td>137</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Minimal survival difference in flies expressing alternating or homomeric, Met-Null CAG/A repeats</td>
<td>138</td>
</tr>
<tr>
<td>5.3.6</td>
<td>Alternative mechanisms of homomeric CAG toxicity</td>
<td>140</td>
</tr>
<tr>
<td>5.4</td>
<td>Discussion</td>
<td>141</td>
</tr>
<tr>
<td>6</td>
<td>CHAPTER 6: FUTURE DIRECTIONS</td>
<td>144</td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
<td>149</td>
</tr>
<tr>
<td>REFERENCES</td>
<td></td>
<td>162</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td></td>
<td>192</td>
</tr>
<tr>
<td>AUTOBIOGRAPHICAL STATEMENT</td>
<td></td>
<td>194</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 2.1 New Drosophila lines for SCA3.................................................................48
Figure 2.2 Expression of the two ataxin-3 isoforms in flies........................................50
Figure 2.3 Differential toxicity from pathogenic ataxin-3 isoforms in developing and adult flies. ..................................................................................................................51
Figure 2.4 Differential toxicity from pathogenic ataxin-3 isoforms in adult neurons and in glia.53
Figure 2.5 Ataxin-3 isoform aggregation in fly neurons. ....................................................55
Figure 2.6 Isoform 2 is no less toxic than isoform 1 when at comparable protein levels........58
Figure 2.7 Knockdown of autophagy or proteasome components leads to increased isoform protein levels in Drosophila.................................................................60
Figure 2.8 Isoform 2 is turned over more rapidly in HeLa cells.........................................63
Figure 2.9 The C-termini of isoforms 1 and 2 are important for host protein levels..............65
Figure 3.1 Ataxin-3 domains impact its polyQ toxicity. ....................................................77
Figure 3.2 UIMs modulate toxicity of the isolated polyQ of ataxin-3.................................80
Figure 3.3 Mutating the UIMs of full-length, pathogenic ataxin-3 decreases its toxicity........82
Figure 3.4 Mutating the UIMs of pathogenic ataxin-3 decreases its aggregation.................83
Figure 3.5 Pathogenic ataxin-3 interacts with Hsc70-4 in a UIM-dependent manner............86
Figure 3.6 Hsc70-4 knockdown improves ataxin-3 toxicity in fly eyes...............................87
Figure 3.7 Hsc70-4 knockdown improves toxicity from independent, truncated polyQ models. 89
Figure 3.8 UIMs are important for inter-ataxin-3 interactions........................................90
Figure 3.9 Proposed model.........................................................................................92
Figure 4.1 N-VCP is non-toxic in Drosophila and interacts with pathogenic ataxin-3.........102
Figure 4.2 N-VCP reduces toxicity from pathogenic ataxin-3 with 77Q............................104
Figure 4.3 N-VCP reduces toxicity from pathogenic ataxin-3 with 80Q............................106
Figure 4.4 The ameliorative effects of N-VCP are specific to ataxin-3 and require its intact VBM.

Figure 4.5 N-VCP improves pathogenic ataxin-3 toxicity in fly eyes.

Figure 4.6 N-VCP reduces pathogenic ataxin-3 aggregation.

Figure 4.7 Dose-dependent N-VCP alleviation of pathogenic ataxin-3 toxicity in fly eyes.

Figure 4.8 N-VCP dose-dependent reduction of ataxin-3 aggregation and inhibition of interaction with VCP.

Figure 4.9 Model of N-VCP-dependent benefits.

Figure 5.1 Designing ataxin-3 constructs for studying RAN translation.

Figure 5.2 Homomeric CAG repeats are more toxic than alternating CAGCAA in fly models of SCA3.

Figure 5.3 No evidence of RAN translation in pan-neuronal Drosophila models of SCA3.

Figure 5.4 No evidence of RAN translation in the glia of Drosophila models of SCA3.

Figure 5.5 ATXN3 constructs to investigate mRNA-toxicity.

Figure 5.6 Homomeric CAG mRNA can be slightly more toxic than alternating CAGCAA in flies expressing Met-Null ATXN3.

Figure 6.1 Heat map visualization of effect of gene knockdowns in Drosophila eyes co-expressing pathogenic ataxin-3.

Figure S2. 1

Figure S2. 2

Figure S2. 3

Figure S2. 4

Figure S2. 5

Figure S2. 6
Figure S3. 1 Ataxin-3 domains impact the toxicity of its polyQ-expanded variant in glia. ....... 152
Figure S3. 2 Levels of polyQ proteins, related to Figure 3.2D..................................................... 153
Figure S3. 3 Uncropped blot from panel 3.6E. ........................................................................... 154
Figure S3. 4 Relative importance of UIMs on fly longevity. .................................................... 154

Figure S4. 1 Motility assays validate non-toxic presence of N-VCP in Drosophila. ............... 155
Figure S4. 2 Daily alive adult numbers from Figure 4.2C.......................................................... 156
Figure S4. 3 Motility assays with the ‘stronger’ pathogenic ataxin-3........................................ 157
Figure S4. 4 VCP increases pathogenic ataxin-3 toxicity in a dose-dependent manner......... 158

Figure S5. 1 Ubiquitous expression longevity data and individual comparisons from survival analysis in figure 2 .......................................................................................................................................... 159
Figure S5. 2 Individual comparisons from survival analysis in figure 6 ................................. 160
Figure S5. 3 Developmental deaths of ATXN3-Q80-expressing flies with glutamine supplementation ................................................................................................................................. 161
LIST OF ABBREVIATIONS

AR: Androgen receptor

DRPLA: Dentatorubral-pallidoluysian atrophy

DUB: deubiquitinase

E1: Ub activating enzyme

E2: Ub conjugating enzyme

E3: Ub ligase enzyme

HD: Huntington's Disease

HMW: Higher Molecular Weight

Hsc: Heat Shock Cognate

HSP: Heat Shock Protein

IP: Immunopurification

MJD: Machado-Joseph Disease

NES: Nuclear export signal

NI: Nuclear inclusion

NLS: Nuclear localization signal

PolyQ: Polyglutamine

RNAi: RNA interference

SBMA: Spinal and bulbar muscular atrophy

SCA: Spinocerebellar Ataxia
TUBE: Tandem Ubiquitin-Binding Entity

UAS: Upstream Activating Sequence

UbS: Ubiquitin-binding Site

UIM: Ubiquitin-Interacting Motif

VBM: VCP-Binding Motif

VCP: Valosin-Containing Protein
CHAPTER 1: THE PROTEIN CONTEXT OF POLYGLUTAMINE DISEASES

1.1 Introduction

The polyglutamine (polyQ) disease family is a group of nine neurodegenerative disorders that are caused by the anomalous expansion of a CAG repeat in the protein-coding region of each disease gene[2-4]. These CAG expansions encode a prolonged polyQ tract in each disease protein that causes a host of shared features: slow progression of neurodegeneration that manifests in the adult years of life, an inverse correlation between the length of the polyQ expansion and the age of onset, and an autosomal dominant inheritance pattern that is shared by all but one of the diseases[2-4]. The current list of polyQ diseases includes dentatorubral-pallidolouysian atrophy (DRPLA), Huntington’s disease (HD), spinal and bulbar muscular atrophy (SBMA or Kennedy’s disease), and spinocerebellar ataxias (SCAs) 1, 2, 3, 6, 7, and 17[2-4]. While each of these disorders shares the same genetic cause and characteristics described above (other than SBMA being an X-linked and not autosomal dominant), they are all clinically distinct[2-4]. Their disease proteins are all widely expressed in both the body and brain, but each disease features accumulation and degeneration in particular brain regions resulting in unique sets of symptoms and presentations[2-4].

This work will highlight the advancements that have been made in the study of polyQ disorders and what differentiates the members of the polyQ family from one another. It is an examination of the factors outside of the polyQ expansion that play a role in the toxicity of each disease protein. For each of these proteins, there is a network of protein-protein interactions and domains that exist outside of the polyQ that help to establish the ‘context’ of pathogenesis for every polyQ disorder. This ‘protein context’ is critical in building our understanding of these
diseases and paints a more comprehensive picture of the toxic pathways and interactions that can be targeted as potential therapeutic entry points.

In this chapter, I will introduce each member of the family of polyQ diseases and then highlight some of the prominent domains and interactions involved in establishing the protein context of each disease protein. This work will be expanded into a future comprehensive review paper covering the entire established protein context of each disease.

1.1.1 Dentatorubral-Pallidoluysian Atrophy (DRPLA)

DRPLA is a progressive ataxia characterized by movement, cognitive, and emotional abnormalities[2,3,5-8]. Additional manifestations vary depending on the age of onset with those developing the disease before the age of 20 typically experiencing epilepsy, myoclonus, and progressive intellectual deterioration and those with onset after age 20 exhibiting cerebellar ataxia, choreoathetosis and dementia with the most clinical similarity to HD[2,3,5]. Repeat expansion occurs in the ATN1 gene located on chromosome 12p which encodes the atrophin-1 protein, a transcriptional corepressor whose function is relatively unknown[2,3,5,9-13]. The normal CAG tract of this protein ranges from 3-38 repeats with a pathogenic threshold beginning at expansions to 49 or more repeats[2,3,5].

Atrophin-1 is a hydrophilic, 1184 amino acid protein with several simple repetitive motifs including a serine-rich region, the variable polyQ tract, a polyproline tract, and a region of alternating acidic and basic residues[2,3,5,14,15]. It also possesses a reported nuclear localization signal (NLS) in its N-terminus (amino acids 16-32) and a reported nuclear export signal (NES) in its C-terminus (amino acids 1033-1041) [2,3,5,16,17]. With these intrinsic signals, atrophin-1 has been shown to localize to both the cytoplasm and the nucleus of neurons and appear in both neuronal nuclear inclusions (NI) as well as cytoplasmic polyQ aggregates, neither of which
necessarily correlate directly with toxicity[2,3,5,14,18-21]. While it is expressed widely in the brain of both normal and affected individuals[13,15,18], the primary targets for degeneration and cell death upon expansion are the dentate nucleus of the cerebellum, red nucleus, globus pallidus, and subthalamic nucleus[2,3,5]. This range of pathological of clinical and pathological presentations among patients with similar mutations may hint at additional genetic and environmental factors playing a role in DPRLA pathogenesis[2,3,5].

1.1.2 Huntington’s Disease (HD)

Huntington’s disease (HD) is the most common among the polyQ disorders[2,22-24]. While the first descriptions of it came in 1872 by George Huntington, the actual mutation in the disease protein, huntingtin, was not discovered until 1993[25]. It is a single gene disorder with autosomal dominant transmission and while most cases have come from families with a history of HD, 6-8% of patients have no family history indicative of somatic expansion[2,22-24]. Motor, psychiatric, and cognitive symptoms begin to appear in middle age and progressively get worse over the next 10-15 years leading to patient death[2,22-24]. Patients will often experience psychiatric and cognitive symptoms, along with subtle motor deficits, for years before official disease onset correlated with the length of the CAG repeat[22]. The areas of the brain preferentially impacted in this disease are the striatum and the deep layers of the cortex[2,22-24]. This includes atrophy of the cerebral cortex and the subcortical white matter and, as the disease progresses, spread to other brain regions like the hypothalamus and hippocampus[2,22-24]. The massive losses in degeneration include up to 95% of the GABAergic medium spiny neurons that project to the globus pallidus and substantia nigra[22].

The polyQ repeat expansion in the huntingtin protein originates from the huntingtin gene (HTT) on the short arm of chromosome 4[22]. A non-HD allele of this gene could have anywhere
from 6 to 35 CAG repeats and the upper limit of that non-pathogenic threshold from 27-35 repeats referred to as an “intermediate allele” [2,22-24]. Beyond this intermediate allele is a range from 36-39 repeats that not fully penetrant, and 40 to as many as 121 repeats representing a fully penetrant mutation[2,22-24]. As is the case with almost every polyQ disorder, huntingtin expansions exhibit anticipation and longer repeats are inversely correlated with the age of onset. It is suspected that this anticipation is a result of intrinsic instability in the CAG repeat of the HTT during meiosis in dividing cells and DNA repair in somatic cells that leads to further expansion with successive generations[22].

The 350 kDa huntingtin protein has an N-terminal polyQ tract that beings at residue 18[2,22-24]. This repeat tract is immediately followed by two consecutive polyproline repeat domains of 11 and 10 residues making up the external loop structure of the N-terminus[22-24]. The rest of the protein is mostly comprised of a series of four ~40 amino acid HEAT (huntingtin elongation factor 3) repeats that are found in a variety of proteins involved in intracellular transport and chromosomal segregation and help mediate huntingtin’s protein-protein interactions[22-24]. Although no NLS has been detected in the protein, huntingtin does have a NES near the C-terminus[23].

Huntingtin is ubiquitously expressed and is present in nearly all tissues[24]. Its widespread subcellular localization shows association with the nucleus, endoplasmic reticulum (ER), Golgi complex, synaptic vesicles, and mitochondria[24]. Many of huntingtin’s interactions with other proteins occur through its N-terminal region (amino acids 1-588) and are associated with either gene transcription or the pathways of intracellular signaling, trafficking, endocytosis, and metabolism[2,23,24]. The role of huntingtin in all these interactions may be that of scaffold, organizing sets of proteins for signaling processes and intracellular transport[24]. In addition to
protein interactions, huntingtin is also cleaved by multiple proteases, resulting in several N-terminal fragments[22-24,26-28]. These N-terminal fragments tend to appear in higher concentrations in the nucleus than the full-length protein and are more susceptible to accumulation and aggregation following polyQ expansion[23]. Whether these aggregates are harmful or protective remains up for debate but may involve several proteins and pathways that will be discussed later in this chapter. Whatever its exact function is, huntingtin does appear to be an essential protein, indicating a loss-of-function mechanism of neurodegeneration; however, there is also evidence that the expanded polyQ is also linked to a toxic gain-of-function that could play a role in HD pathogenesis adding to the complexity of this disease.

1.1.3 Spinal and Bulbar Muscular Atrophy (SBMA or Kennedy’s Disease)

Spinal and bulbar muscular atrophy (SBMA), also known as Kennedy’s Disease, is a late-onset neuromuscular disorder caused by the expansion of a polymorphic polyQ repeat in the androgen receptor (AR) protein. In 1991, SBMA became the first identified polyQ disease and has several features that make it unique among the family of disorders[29]. It is the only X-linked polyQ disorder and thus the only polyQ disorder that is not inherited in an autosomal dominant manner[2,30,31]. The existence of the polyQ expansion in the AR protein also means that SBMA is the only ligand-dependent polyQ disorder and shares the disease protein with many other diseases including androgen insensitivity syndrome (AIS) and prostate cancer[30-33]. On top of the late-onset, progressive neurodegeneration, dysarthria, and dysphagia shared among most polyQ disorders, the overlap of disease proteins means that men with SBMA also experience mild signs of AIS like gynecomastia and some degree of infertility[30,31]. Some of these effects, in both SBMA and AIS, are suspected to be a result of AR loss-of-function, but the presence of AIS
symptoms in SBMA, and of any neuromuscular phenotypes in AIS suggests that the polyQ expansion of AR must lead to a gain-of-function that is selectively harmful to motor neurons[30,31].

While the X-linked nature of the disease means that full penetrance only occurs in males, women who are heterozygous for the polyQ AR mutation may experience subclinical effects like muscle cramps and electrophysiological abnormalities, but their low circulating levels of AR ligands protects them from neurodegeneration[2,30,31]. For males, SBMA is also detrimental to skeletal muscle and can cause progressive muscle cramps, weakness and wasting, and twitching[30,31]. These neuromuscular symptoms have resulted in a large fraction of SBMA patients being initially misdiagnosed with other neuromuscular disorders, including amyotrophic lateral sclerosis (ALS) or autosomal recessive spinal muscular atrophy (SMA)[30]. Distinguishing between these diagnoses is critically important for the patient because while the mean survival time after an ALS diagnosis is only a few years, the life expectancy for SBMA patients is normal or only minimally reduced[30,31].

Pathologically, SBMA is characterized by the loss of lower motor neurons of the brainstem and anterior horn of the spinal cord, and to a lesser extent the sensory neurons of the dorsal root ganglia[2,30,31]. This degeneration leads to weakness and atrophy of the bulbar, facial, and limb muscles in concordance with the identified role of skeletal muscle in SBMA disease pathogenesis[30,31]. It has been theorized that the toxicity of polyQ expanded AR may originate in the skeletal muscles and result in secondary motor neuron degeneration[30,31].

As mentioned above, the polyQ expansion in SBMA occurs in the AR protein, the male steroid receptor encoded by the AR gene on the short arm of the X chromosome. This gene contains 8 coding exons consisting of three functional domains that are shared among the super-family of steroid-binding transcription factors[30,31]. The highly polymorphic CAG repeat is in the first exon
of the AR gene and would normally possess 5-34 repeats but gets expanded to 37 or more repeats in alleles that have reached the SBMA pathogenic threshold, while shorter CAG tract lengths have been associated with an increased risk of prostate cancer[30,31]. The AR sequence actually has 3 polyQ tracts, but the first one (located in exon 1) is by far the longest[30].

Another unique feature of SBMA is that unlike other polyQ disorders, whose native disease protein function is mostly unknown, the biology of the AR is amply studied and well understood, allowing for greater understanding of the relationship between native AR function and polyQ expanded AR dysfunction. AR is a ligand-activated transcription factor and member of the nuclear receptor super-family[30,31]. AR consists of an N-terminal domain (NTD) that modulates transcriptional activation, a central DNA-binding domain (DBD) that binds androgen-responsive elements (AREs), and a C-terminal ligand-binding domain (LBD) [30,31]. It also has several subdomains that are involved in nuclear localization, dimerization, and interaction with heat shock proteins (HSPs), co-activators and other proteins[30,31]. In the absence of androgens, AR usually resides in the cytoplasm in a multi-heteromeric inactive cytosolic complex with HSPs[30,31]. Upon binding of the AR-ligand testosterone, or its metabolite 5α-dihydrotestosterone (DHT), the AR undergoes a conformational change that promotes its dissociation from the HSP chaperone complex and exposes the AR NLS, DBD, and dimerization domains[30,31]. This exposure then promotes dimerization and nuclear translocation, allowing AR to then bind AREs directly in target promoter regions and selectively recruit transcriptional co-activators and machinery for transcription initiation of androgen-sensitive genes[30,31]. Wild-type AR also interacts with proteins involved in several pathways including RNA-splicing, protein translation, proteasome/protein ubiquitination, transcription, and male sexual differentiation and
Each one of these normal mechanisms have been implicated in SBMA pathogenesis upon polyQ expansion of AR, some of which will be explored later in this chapter.

1.1.4 Spinocerebellar Ataxia Type 1 (SCA1)

Spinocerebellar ataxia type 1 (SCA1) was first described in 1993 when an unstable CAG repeat was identified in the ATXN1 gene[34]. Like many polyQ disorders, SCA1 is an adult-onset and progressive inherited ataxia that presents with cognitive impairments, difficulty with speaking and swallowing, and eventually chronic lung infections and respiratory failure[2,35-37]. Diagnosis is typically in the third or fourth decade of life, with the disease duration ranging from 10-30 years depending on the size of the repeat expansion[2,35-37]. Symptoms typically begin with gait ataxia and progress to the point of wheelchair necessity within 15 years[2,35-37]. Patients of SCA1 make up approximately 6% of all autosomal dominant cerebellar ataxias, with large variances across ethnic groups[35]. The primary pathological pathways of SCA1 are degeneration of the cerebellum and brain stem, as well as atrophy of the ventral pons and middle cerebellar peduncles[2,35-37]. There is major loss of cerebellar Purkinje cells and neurons of the inferior olivary nucleus, cortical, subcortical, and spinal structures[2,35-37].

The ATXN1 gene is located on chromosome 6p and encodes the 816 amino acid, 87 kDa, ataxin-1 protein[2,35,38]. This protein is widely expressed across neuronal and glial populations throughout the central nervous system. Of the 9 exons in the ATXN1 gene, only exons 8 and 9 contain protein-coding sequences and exon 8 contains the polyQ tract[35]. This polyQ tract normally contains anywhere from 4 to 36 repeats and any alleles with over 21 repeats also possessing 1 to 3 interrupting CAT codons encoding histidine[2,35-37]. These interrupting histidine residues reduce the aggregation kinetics of expanded ataxin-1 and are often lost upon fully penetrant pathogenic expansion[35]. With that said, uninterrupted CAG repeats of over 39, and
interrupted CAG repeats of 43 or more, result in the highly-penetrant, late-onset, and progressive ataxia described above, and does display the typical correlation between polyQ tract length and disease severity[35]. Tracts with 36-38 repeats have an intermediate effect that can result in ataxia without some of the other SCA1 specific clinical features[35].

Wild-type ataxin-1’s function remains somewhat unknown although, like many polyQ disease proteins, has suspected involvement in the regulation of transcription factors[2,35,37,39]. Without histidine interruptions, polyQ expanded ataxin-1 is highly stable and much more resistant to degradation[35]. While the polyQ expansion has been identified as the central cause of SCA1, ataxin-1 has several domains that mediate both its normal function and its pathogenicity upon expansion that will be discussed in future sections. In terms of suspected normal function, these domains allow ataxin-1 to shuttle between the nucleus and cytoplasm as well as bind RNA and other proteins like transcription and splicing factors[35,37]. In addition to SCA1, ataxin-1 is also implicated in other neurodegenerative diseases[35]. Intermediate polyQ expansions of 32 or more repeats that do reach the SCA1 pathogenic threshold are associated with a greater risk of developing sporadic amyotrophic lateral sclerosis (sALS)[35]. Additionally, a SNP in an intronic region of the ATXN1 gene has been associated with an increased risk and earlier onset of late-onset Alzheimer’s disease (LOAD)[37].

1.1.5 Spinocerebellar Ataxia Type 2 (SCA2)

While spinocerebellar ataxia type 2 (SCA2) was initially described in India throughout the 1960s and 1970s[40], it was not until 1996 that the disease gene, ATXN2, along with the causative CAG expansion encoding a polyQ repeat in the causative protein, ataxin-2, were identified[41-43]. As is the case with other SCA polyQ disorders, SCA2 is an autosomal dominantly inherited neurodegenerative disorder characterized primarily by gait ataxia, onset coinciding with muscle
cramping, and other symptoms indicative of cerebellar degeneration[2,3,44]. Unlike other SCAs, whose symptoms tend to involve or spread to other brain regions, SCA2 signs and symptoms are almost entirely of cerebellar origin and include presentations such as appendicular ataxia, dysarthria, and ocular deficits including nystagmus and ocular dysmetria[2,3,44]. Other frequently reported symptoms are dystonia, frontal-executive dysfunction, myoclonus, muscle spasticity, neuropathy, and slow or absent saccades[44]. Like SCA1 and SCA3 there is major pathological involvement in the cerebellum and brainstem[2,3,44]. There is evidence of atrophy in the pontine gray matter, middle cerebellar peduncles, cerebellar white matter and folia, and inferior olives[2,3,44].

The ATXN2 gene is located on chromosome 12q24 and normally possesses a CAG tract of 15-32 repeats[2,3,44]. SCA2 pathogenicity is reached when repeats are expanded beyond 32 and can reach as high as 77, with the possibility of reduced penetrance in those possessing repeats of 32-34 CAGs in length[2,3,44]. As with many polyQ disorders, there is a strong inverse correlation between the repeat length and the age of onset of the disorder along with anticipation as the gene is passed from one generation to the next. Consisting of 25 exons, ATXN2 has a transcript of 4699 base pairs and two in-frame start codons[44]. The purpose of these two translational pathways remains undetermined, but it is known that the translation from the first start codon produces a 144kDa protein while translation from the second, only four codons upstream of the CAG repeat, produces an ataxin-2 protein that is 17kDa smaller[44]. In addition, an even smaller 42kDa ataxin-2 fragment has been observed in patient brain extracts and SCA2 mice[45-47].

It is thought that expansion within these transcripts causes a toxic gain of function of the ataxin-2 protein resulting in abnormally slow Purkinje cell firing frequency and eventual degeneration. Normally, ataxin-2 is cytoplasmic protein with proposed functions in RNA
metabolism[2,3,44]. It has reported interactions with multiple RNA binding proteins and contains N-terminal Lsm and LsmAD domains that are common in spliceosomal small nuclear ribonucleoproteins and aid in ataxin-2’s ability to bind RNA and other RNA-associated proteins[46,48]. In addition, ataxin-2 can localize to the trans-Golgi network, and can be phosphorylated as well as cleaved by caspase-3 at cleavage consensus site (aa 396-399)[49,50]. These traits and others open the door for additional roles for ataxin-2 in stress granule formation/regulation, secretion and endocytosis, calcium homeostasis, and apoptotic and receptor-mediated signaling.

Apart from SCA2, ataxin-2 is also implicated in multiple diseases with variant ataxin-2-associated phenotypes that are entirely outside of the cerebellar spectrum: most prominent being L-DOPA responsive Parkinsonism and ALS[44].

### 1.1.6 Spinocerebellar Ataxia Type 3 (SCA3 or Machado Joseph Disease)

Spinocerebellar ataxia type 3 (SCA3), also known as Machado Joseph Disease (MJD), is the most common dominantly inherited ataxia worldwide[51-54]. Like other polyQ SCAs, SCA3 is a progressive ataxia defined by cerebellar and brainstem dysfunction[51-54]. Onset is typically in the young to mid adult years and initially manifests as progressive gait imbalance along with vestibular and speech difficulties[51-57]. As the disease progresses, visual and oculomotor problems appear including nystagmus, jerky ocular pursuits, slowing of saccades, disconjugate eye movements, ophthalmoplegia, and apparent bulging eyes[51-57]. The most advanced stages of the disease leave the patient wheelchair bound with severe dysarthria and dysphagia[51]. Other potential symptoms include facial atrophy, dystonia, spasticity, amyotrophy, and non-dementia mild cognitive impairment[51,58-60]. Life expectancy following disease onset is approximately 20-25 years[51,61].
The polyQ expansion occurs in a gene first identified in 1994 as MJD1, also referred to as the ATXN3 gene, that encodes the deubiquitinating enzyme (DUB) ataxin-3 protein[62]. The disease protein has been shown to accumulate within nuclear inclusions of specific populations of neurons in several brain regions that experience degeneration; particularly pontine neurons but also other brainstem neuronal populations, dentate nuclei, thalamus, substantia nigra, globus pallidus, cranial motor nerved nuclei, and, rarely, in the striatum[51-54,63-65]. The increase in CAG-tract length in the ATXN3 gene causes an expanded polyQ in the ataxin-3 protein that misfolds, aggregates, and forms the intranuclear inclusions in the affected regions listed above[51-54,62,66]. Whereas normal alleles possess anywhere from 12 to 43 repeats, pathogenic ATXN3 alleles contain 60-87 repeats with milder forms of ataxia and restless leg syndrome appearing in patients with mid-range repeats[51-54,67]. In the repeat range that is full penetrant, the largest repeats will often lead to more pyramidal signs and dystonia[68] while repeats on the lower end of the pathogenic range (<73) are associated with peripheral neuropathy[51,69,70].

Ataxin-3 is a 42 kDa protein that is widely expressed in the central nervous system and, like many of the neurodegenerative diseases covered in this chapter, has selective tissue vulnerability that cannot be explained by restricted protein expression[51-54]. As a DUB, it is believed that ataxin-3 plays a role in the clearance of misfolded proteins in the ubiquitin-proteasome system (UPS), and while the exact role of ataxin-3 remains unclear, it is thought that the expansion of its polyQ domain somehow burdens the protein quality control of cells in the affected regions[51-54]. The ataxin-3 protein consists of several domains and multiple splice isoforms outside of the polyQ tract that are crucial to its native function and its pathogenesis upon expansion in SCA3. These domains, isoforms, and the protein-protein interactions associated with them will be the focus of the following chapters of this dissertation.
1.1.7 Spinocerebellar Ataxia Type 6 (SCA6)

Spinocerebellar ataxia type 6 (SCA6) is caused by small polyglutamine expansions in the bicistronic CACNA1A gene that encodes for the α1A subunit of the P/Q-type voltage-gated calcium channel called Cav2 and the transcription factor α1ACT[2,3,71]. Identified in 1997, it is the first polyQ disease attributed to mutations in an ion channel gene[72]. The disease has been found with variable frequency in many populations worldwide and has particular prominence in the Netherlands and Japan where it is the second most common SCA behind SCA3[2,71,73]. Clinically, SCA6 is a pure cerebellar ataxia with a dominant inheritance pattern and typical appearance later in life[2,3,71]. Patients show pure cerebellar dysfunction, slowly progressive gait ataxia and imbalance, dysarthria, and late-onset nystagmus[2,3,71]. While the age of onset can range anywhere from 19 to 71 years age, the mean age of onset is 43-52 and is inversely correlated with the size of the expansion[71]. After onset, SCA6 manifests much slower than other SCAs and does not usually shorten the patient’s lifespan, but even with slow progression there can be large variances in symptom severity, even within the same family[2,3,71]. In addition to SCA6, mutations in the CACNA1A gene are associated with at least two other dominantly inherited neurological disorders, episodic ataxia type 2 and familial hemiplegic migraine, that both have clinical overlap with SCA6[2,71,74].

The pathological presentation of SCA6 is defined by diffuse loss of Purkinje cells, primarily in the cerebellar vermis with very limited degeneration in associated cortical and cerebello-olivary regions[2,3,71]. The remaining Purkinje cells in the most degenerated regions undergo extensive morphological changes including heterotypic nuclei, unclear cytoplasmic outline, somatic sprouts, dendritic swelling with increased numbers of spine-like protrusions, and disorganization of axonal arrangements[71].
The SCA6 disease gene, *CACNA1A*, is located on chromosome 19p13 and contains 47 exons, the last of which is home to the expanded CAG repeat[3,71]. Normally this repeat exists in a range of 4-18 CAGs but is expanded to a range of 19-33 repeats in the disease state[2,3,71]. SCA6 does show variable penetrance within its pathological range where heterozygous individuals with 19 repeats are usually asymptomatic while those possessing 20 or more repeats showing full penetrance[71]. The *CACNA1A* gene is bicistronic and has a novel internal ribosomal entry site (IRES) that allows it to encode two structurally unrelated proteins. These proteins have distinct functions and have overlapping open reading frames within the same mRNA[71]. The first of these proteins is the α1A subunit of the Cav2 P/Q-type voltage-gated calcium channel. The second is a newly recognized transcription factor known as α1ACT that possesses the C-terminal polyQ repeat. Both proteins are expressed abundantly in the Purkinje cells of the cerebellum and the C-terminal α1ACT peptide with an expanded polyQ tract has been shown to be toxic in tissue culture in manner that is dependent on nuclear localization[71].

α1ACT is a functional transcription factor that coordinates the expression of neuronal genes involved in cerebellar and Purkinje cell development. In addition to the expanded polyQ and an NLS, it contains residues that are typically associated with channel inactivation and modulation by intracellular signaling proteins[2,3,71]. One hypothesis is that SCA6 is an ion channel disorder and the polyQ expansion of α1ACT causes ataxia by altering calcium channel function, but the investigations into this theory have yielded conflicting findings[71]. Other evidence has shown nuclear accumulation of the α1A subunits which suggests that pathogenesis could be more closely related to the nuclear accumulation of expanded polyQ proteins/fragments that could lead to transcriptional dysregulation as seen in other polyQ disorders[75]. This accumulation and expansion could also promote a toxic gain-of-function that could exacerbate disease progression.
through several mechanisms including sequestration of housekeeping genes and transcription factors, disruption of UPS pathways, direct DNA damage, or even the induction of apoptotic pathways[76,77]. Although SCA6 was identified as the first ion channel polyQ mutation, the evidence of that polyQ mutation existing in the second cistron encoded transcription factor α1ACT points to a more likely mechanism of toxicity involving transcriptional dysregulation because of expansion instead of the typical mechanisms of channelopathies seen in many other CACNA1A disorders.

1.1.8 Spinocerebellar Ataxia Type 7 (SCA7)

Spinocerebellar ataxia type 7 (SCA7) is caused by a polymorphic polyQ expansion in the ATXN7 gene encoding the ataxin-7 (ATXN7) protein[78]. The gene and expansion were identified in 1997, making SCA7 the eighth of nine currently known polyQ disorders[79]. Clinical presentation is similar to that of SCAs 1, 2, and 3: spastic ataxia, dysarthria, dysphagia, slow eye movement, ophthalmoplegia, prominent hyperflexia with crossed supraclavicular, pectoral, and hip adductor reflexes, spasticity, and pyramidal signs[78]. What makes SCA7 unique is that it is the only SCA and polyQ disorder that is characterized by retinal degeneration in addition to ataxia[78]. Visual impairment begins with cone photoreceptor degeneration that progresses toward cone-rod dystrophy and eventually complete blindness[78]. Postmortem pathology from patient cerebellum shows substantial loss in the Purkinje cell layer, dentate nuclei, and to a lesser extent the granule cell layer[78]. The initial, ataxia-like symptoms that appear early in SCA7 match well with the primary neuropathology and then become increasingly diverse and the degenerative process becomes more widespread as the disease progresses[78].

Repeat expansions in the ATXN7 gene have some of the widest ranges among polyQ disorders. The ATXN7 gene is widely expressed in neuronal and non-neuronal tissue and the CAG
tract of wild type ATXN7 typically has 4-35 repeats, with an estimated 70-80% of all ATXN7 genes carrying 10 CAGs and [78,79]. Alleles with 28-35 repeats are prone to expansion and repeats of 34-36 are associated with reduced penetrance and a mild- and late-onset of disease symptoms [78,79]. SCA7 repeats show the highest tendency to expand upon transmission and show strong anticipation so while 36 repeats are the pathological threshold, these repeat expansions can reach as high as 460 repeats [78,79]. This means that there is a wide range of fully penetrant manifestations of SCA7: 36-55 repeats present as adult-onset SCA7; greater than 70 repeats present as a juvenile-onset form of SCA7 with an accelerated disease progression; and greater than 100 repeats present as a severe infantile form of the disease with death in the first few years or months of life [78,79]. Longer repeats are not only inversely correlated with the age of onset, but also result in an earlier appearance of retinal degeneration and vision loss prior to any cerebellar ataxia [78].

The ATXN7 protein is a member of the multiprotein Spt-Ada-Gcn5 (SAGA) complex [78,80,81]. The SAGA complex is responsible for interactions with transcriptional machinery and chromatin modifications though both histone deacetyltransferase (HAT) and DUB activities [78]. In this complex, ATXN7 functions in several proposed roles. The first is in the DUB module where it anchors the DUB to the rest of the SAGA complex [78]. ATXN7 is also a functional scaffold of SAGA that plays a role in the regulation of H2B Ub levels and has been suggested as being important for the full differentiation of Purkinje neurons in the cerebellum and maintenance of differentiated photoreceptors [78].

The protein is localized to the cytoplasm and nucleus of neurons, and the ratio between the two varies in different brain regions [78]. It is also present in all neurons of the retina [78]. PolyQ expanded ATXN7 has been shown to form large nuclear inclusions and aggregates in cell nuclei across degenerating and spared brain regions but does seem to accumulate faster in vulnerable
nuclei (Purkinje cells and photoreceptors) indicating that accumulation, not localization, is a better indicator of toxicity[78]. The polyQ domain is in the N-terminal portion of the ATXN7 protein and is followed by a small proline repeat that resembles the polyQ/proline-rich regions of HD and other transcription factors[78,79]. Additional domains include 3 NLSs and 1 NES[78,82]. The effects of polyQ expanded ATXN7 have been shown in mice to be both cell autonomous and non-cell autonomous[78].

1.1.9 Spinocerebellar Ataxia Type 17 (SCA17)

The most recently discovered polyQ disorder, spinocerebellar ataxia type 17 (SCA17) is a rare disease accounting for approximately 0.5-1% of dominantly inherited ataxias[83-85]. It is caused by a polyQ expansion in the TATA-box binding protein (TBP), a general and essential transcription initiation factor[83-85]. Clinically, it can resemble the presentations of Parkinson’s and Huntington’s diseases as well as psychiatric symptoms ranging from psychosis to multitudinous cerebellar syndromes[83,86,87]. Patients experience many of the early signs of ataxia with gait imbalance, dysarthria, dystonia, and choreic movements, but also encounter more intellectual decline, dementia, psychiatric abnormalities, seizures, and basal ganglia symptoms that can help diagnostically distinguish it from other SCAs and HD[83,86,87]. At the pathological level, SCA17 manifests as marked cerebellar atrophy and degeneration, as well as moderate and diffuse cortical, subcortical, and brainstem atrophy[83,86,87]. Purkinje cell loss and gliosis are also prominent features of this disease[83,86,87].

The 18,528bp TBP gene is located on chromosome 6q and the 1867bp transcript spanning from exon 2 to exon 8 encodes the TBP protein[83]. The polyQ repeat sequence of TBP is in exon 3 of this transcript, and wild-type alleles typically have 25-42 repeats, with the most common range in human alleles being 32-39 repeats[83,86,87]. The TBP polyQ tract contains two polymorphic
(CAG)\textsubscript{n} stretches interrupted by polyQ-encoding CAAs in the following pattern: (CAG)\textsubscript{3} (CAA)\textsubscript{3} (CAG)\textsubscript{n} CAA CAG CAA (CAG)\textsubscript{n} CAA CAG\cite{83}. These interruptions stabilize the repeat sequence, and it is though that intergeneration instability and anticipation result from the loss of the interrupting CAA CAG CAA element and fusion of the (CAG)\textsubscript{n} blocks, also seen in SCA1 and SCA2\cite{83}. The stabilization by CAA interruptions in tandem with the narrow gap between normal and abnormal repeats makes it difficult to determine an exact pathological cutoff. Our current understanding is that the range for SCA17 is 46-55 repeats, with 43-48 repeats constituting an intermediate range with incomplete penetrance\cite{83,86,87}. Juvenile-onset SCA17 has been reported in patients with repeats exceeding 62 and experience different set of symptoms including ataxia, intellectual deterioration, muscle weakness, growth retardation, faster disease progression, and early death\cite{86}.

TBP is a saddle shaped protein that is recruited to the basal promoter elements of RNA polymerases I, II, and III and can directly bind the TATA boxes in DNA or, in the case of TATA-less promoters, have protein-protein interactions in order to initiate transcription\cite{83,86}. It is also one of the more well reported polyQ disease proteins in the family as its dysfunction is implicated in several neurodegenerative disorders, including other polyQ disorders\cite{83,86}. Bound TBP functions as the ‘nucleus’ of the initiation complex and provides docking sites for several additional transcription factors in the RNA Pol II in particular\cite{83,86}. TBP has a variable N-terminal domain and highly conserved C-terminal domain and it has been suggested that the C-terminal core domain houses the large DNA-binding domain that interacts with the DNA directly while the N-terminal assists in modulating that DNA-binding activity\cite{83,86}. The polyQ stretch of TBP is located at the N-terminus starting at amino acid 58 and is followed by the remaining 244 residues of TBP\cite{83}. As it is in the N-terminal half of the protein, another suggestion is that the
subregions surrounding the polyQ synergistically promote the formation of the TBP-DNA complex, and that neurodegeneration upon polyQ expansion could arise from a toxic gain-of-function rather than a loss-of-function[86]. Multiple models have shown an age-dependent accumulation of polyQ expanded TBP aggregates in neuronal nuclei and cerebellar degeneration, especially Purkinje cell death, but their actual role in the degenerative process remains unclear[83,86].

1.2 Advancements in establishing disease protein context

1.2.1 Dentatorubral-Pallidoluysian Atrophy (DRPLA)

The DRPLA protein, atrophin-1, is reported to have multiple interactions with caspases that cleave the protein and form atrophin-1 fragments that may alter its regulatory mechanisms upon polyQ expansion and increase cellular toxicity[5,88,89]. Many of these cleavages and their purported toxic effects center around separating the atrophin-1 N-terminal nuclear localization signal (NLS) or C-terminal nuclear export signal (NES) from the remainder of the fragment and polyQ expansion[5]. Caspase cleavage separating the NLS from the rest of the protein results in a C-terminal fragment that is more prone to cytoplasmic accumulation and has been observed in DRPLA patient tissue[5,88]. Without an NLS, this fragment is normally exported to the cytoplasm as a nucleocytoplasmic protein, and upon exportation, caspases play a direct role in its cleavage to regulate accumulation[5,88]. However, polyQ expansion of atrophin-1 (and subsequently its fragments) reduces caspase activity leading to accumulation of the fragment in the cytoplasm, suppression of nuclear transport of full-length atrophin-1, and a furthering of cytoplasmic aggregation and potential toxicity[5,88]. The N-terminal Asp109 residue of atrophin-1 represents an additional cleavage site separating the NLS from the fragment protein. Cleavage at the Asp109 site (N-terminal 106(DSLEDG)110) in an expanded state by caspase-3 separates the NLS from the
rest of the protein and produces a pro-apoptotic fragment that could contribute to aggregation but does not effect atrophin-1 localization[16,89]. Mutation of this site to prevent cleavage blocks the creation of this fragment and dramatically reduces cellular toxicity; however, whether this caspase-3 specific effect occurs in vivo has not been confirmed[89].

The NES also seems to be important, with fragments generated by unknown caspases that separate the NES from the polyQ and NLS showing increased nuclear retention and toxicity[5]. Both patients and DRPLA transgenic mice have nuclear accumulation of atrophin-1, particularly a 120 kDa fragment that would be representative of a cleavage product[5]. Mutation of full length polyQ expanded atrophin-1 to render the NES non-functional in Neuro2a cells resulted in increased nuclear labeling of the protein and an increase in cellular toxicity. Furthermore, mutating the NLS in the fragment lacking an NES shifted atrophin-1 more to the cytoplasm and reduced cellular toxicity[5]. So far, it has been shown that this fragment is not a result of caspase 2, 3, 6, 7, or 8 cleavage and the exact enzyme that creates it has not been determined[5]. These studies indicate that the NLS and NES both play critical roles in normal atrophin-1 function and polyQ expanded atrophin-1 toxicity, and that multiple caspase cleavages involving these sites act as modulators of aggregate formation and toxicity in DRPLA. They also show that both nuclear and cytoplasmic localization of expanded atrophin-1 fragments are factors in pathogenesis.

On top of cleavage products, atrophin-1 also has a few recorded direct interactions that could play a role in toxicity. The first is an interaction through one of two arginine-glutamic acid (RE) dipeptide repeats in the atrophin-1 protein[90]. The interaction is with a protein, RERE, that shares 67% sequence homology with atrophin-1, including the proximal RE-repeat through which they can form a heterodimer[90]. RERE and atrophin-1 interact and colocalize with one another in what appers as a speckled pattern in the nucleus, and this interaction is strengthened (~3.3 fold) when
the atrophin-1 polyQ is expanded[90]. This increased interaction results in the formation of larger atrophin-1 aggregates and, as RERE over-expression is toxic to cells, could mean that further interaction between polyQ expanded atrophin-1 and RERE is also toxic[90]. Additional interactions that will require further investigation to determine their exact physiological relevance to the disease include atrophin-1’s interaction with the transcriptional regulator ETO/MTG8[91]; Drosophila atrophin-1 homolog Atro’s interaction with Drosophila even-skipped[92]; and the Drosophila brakeless homolog ZNF608 interaction with atrophin-1 that all play a part in atrophin-1’s proposed function as a transcriptional corepressor[93].

1.2.2 Huntington’s Disease (HD)

The interactions and protein context of the expanded huntingtin protein can be sorted into a few major categories: caspase cleavage, post-translational modification, changes to intracellular trafficking and signaling interactions, and gene transcription alterations. There have been several sites within the huntingtin protein, primarily the first 550 amino acids, that have been identified as caspase cleavage sites[27,94-96]. In HD mouse models, modification of huntingtin to prevent caspase-6 cleavage at amino acid 586 helped maintain normal neuronal function and eliminate the striata neurodegeneration and motor dysfunctions normally observed in HD phenotypes[97,98]. This indicated that caspase-6 cleavage may be a potentiating event in HD. In conjunction with post-translational modification changes in HD, it has been shown that huntingtin is phosphorylated by Cdk5, and that this phosphorylation reduces the cleavage of huntingtin by capsae-3 at amino acid 513[99,100]. Phosphorylation, and resulting protection from cleavage, reduces aggregation and toxicity in cell models of HD[99,100]. However, in HD there are lower levels of Cdk5 which leads to reduced phosphorylation of huntingtin and any afforded protections from cleavage and subsequent aggregation and toxicity are lost[99,100]. Given the more toxic nature of N-terminal
huntingtin fragments and the evidence of natural generation and accumulation of N-terminal expanded huntingtin fragments in HD mouse brains[23,24,101-106], it stands to reason that cleavage at these sites plays a role in enhancing toxicity in HD.

Post-translational modifications play additional roles in huntingtin toxicity beyond changing accessibility to caspase-3 cleavage sites. Huntingtin undergoes extensive post-translational modification and has sites for phosphorylation, SUMOylation, acetylation, ubiquitination, and palmitoylation. Each of these modifications can alter the toxicity of the expanded huntingtin protein. The first of these modifications, phosphorylation, generally seems to protect cells from the toxic effects of expanded huntingtin. The Akt-mediated phosphorylation of S421 of expanded huntingtin has been shown to relieve neurotoxicity in cell models of HD[107-112]. The same amelioration was observed upon phosphorylation of expanded huntingtin at sites S536, S13, and S16[100,113,114]. The mechanisms of these protections are not yet fully understood but it seems that N-terminal phosphorylation targets huntingtin for degradation[113,114], while S536 phosphorylation limits calpain cleavage[100] and S421 phosphorylation increases trophic support in intracellular transport[107,112]. Acetylation and palmitoylation also appear to be protective against huntingtin-induced neurodegeneration. Mutations to prevent the acetylation of huntingtin K444 resulted in increased accumulation and neurodegeneration in both cultured neurons and mouse brains[115]. Similar results were observed upon mutation of huntingtin’s C214 palmitoylation site[116]. At this site, palmitoylation is performed by huntingtin-interacting protein 14 (HIP14) [116]. The association between huntingtin and HIP14 is reduced when the huntingtin polyQ is expanded resulting in reduced palmitoylation and increased inclusion formation, a potential contributor to HD pathogenesis[116].
Huntingtin’s K6 and K9 residues are sites of post-translational modification with influence on two fronts. SUMOylation at both sites reduced the ability of an expanded N-terminal huntingtin fragment to aggregate and promoted its capacity to repress transcription in a cell model[117]. However, in a Drosophila model expressing exon 1 of huntingtin, SUMOylation at the same sites enhanced neurodegeneration while ubiquitination was protective[117]. These conflicting results point to a role for SUMOylation in HD pathogenesis that requires further elucidation and further establishes the importance of understanding the protein context surrounding and influencing these factors in toxicity.

In addition to cleavage and post-translational modification, expanded huntingtin also has several direct protein-protein interactions that influence its toxicity, particularly with proteins involved in intracellular trafficking and signaling. Huntington-associated protein 1 (HAP1), among its many influential associations and interactions, was found to interact with both huntingtin and the type 1 inositol (1, 4, 5)-trisphosphate receptor (IP_3_1) to form an IP_3_1-HAP1-huntingtin ternary complex[118,119]. The presence of an expanded huntingtin in this complex enhances the sensitivity of IP_3_1 to inositol (1, 4, 5)-trisphosphate and provides an explanation for the changes in calcium signaling observed in HD patients[118]. Another huntingtin associated protein, huntingtin-interacting protein (HIP1), has decreased binding to mutant huntingtin compared to wild-type[120-122]. It has been theorized that this decreased binding increases the amount of free HIP1, which can then associate with HIP1-protein interaction (HIPPI) and induce caspsae-8 cleavage and apoptosis[122]. Mutant huntingtin also has a decreased interaction with postsynaptic density 95 (PSD-95), a scaffold protein that causes clustering and activation of postsynaptic membrane receptors[123]. HD patients release more PSD-95 than unaffected individuals which
results in an over activation of N-methyl-D-aspartate (NMDA) receptors and associated abnormal synaptic transmission that has been validated in HD mouse models[124,125].

The final set of huntingtin interactions that will be discussed in this section are those that impact transcriptional activity. Wild-type huntingtin interacts with repressor element-1 transcription factor in the cytoplasm[126,127]. REST, in association with neuron restrictive silencer element (NRSE), affects the nuclear transcription of neuronal genes, including brain-derived neurotrophic factor (BDNF) [126,127]. The interaction between huntingtin and REST-NRSE is much weaker when the polyQ if huntingtin is expanded[126,127]. So, while wild-type huntingtin may interact with REST-NRSE in the cytoplasm to prevent entry into the nucleus and reduce the REST-NRSE-induced suppression of gene expression, mutant huntingtin limits this prevention which subsequently inhibits the expression of BDNF and other neuronal genes[126,127]. PolyQ-expanded huntingtin inhibits the function of several transcriptional coactivators, in a similar manner to REST-NRSE, including p53 and CBP[128,129], and represents a pathway to toxicity through transcriptional dysregulation that connects several polyQ disorders and will be useful in understanding polyQ toxicity more generally.

1.2.3 Spinobulbar Muscular Atrophy (SBMA or Kennedy’s Disease)

The unique nature of native AR’s function compared to other polyQ disease proteins makes SBMA the only ligand-dependent polyQ disease. Many of the protein-protein interactions and domains of influence in expanded AR toxicity are only possible when AR-LBD is in a ligand-bound state with either testosterone or DHT. The binding of a ligand to AR along with polyQ expansion in the AR-NTD promotes conformational changes in the AR protein, ligand-dependent unfolding, and nuclear translocation of the mutant protein[130]. All these steps have been shown to be necessary in SBMA pathogenesis and can alter the AR interactome[130].
In the presence of the AR ligand DHT, retinoblastoma protein (Rb) showed a much stronger association with polyQ expanded AR compared to wild-type AR[131]. Rb normally acts as a neuroprotective factor that recruits the histone deactyltransferase 1 (HDAC) in the co-repression of E2F1 – a transcription factor whose target genes regulate apoptosis, development, and differentiation[131]. The stronger association between Rb and expanded AR suppresses Rb’s ability to recruit HDAC1 and leads to aberrant E2F1 transcriptional activation and neurotoxicity. Another ligand-dependent AR interaction involves DHT-induced cell cycle arrest and how polyQ expansion alters the function of AR in cell cycle regulation[131]. Under both normal and expanded conditions, AR interacts in complex with anaphase promoting complex/cyclosome (APC/C) and its adaptor protein, Cdh1, in a DHT-dependent manner[132]. Cell studies revealed the possibility that expanded AR acts as a pseudo-inhibitor, competing with substrates for binding with APC/C-Cdh1 through a toxic gain-of-function upon polyQ expansion that results in abortive neuronal differentiation followed by mitotic entry[132].

The AF-2 domain also plays a prominent role in the interactions that facilitate expanded AR toxicity. Within the AR protein itself, the interdomain interaction between the C-terminal AF-2 domain and N-terminal FXXLF motif (N/C interaction) is important for AR aggregation and toxicity[133-135]. While it is a normal step in wild-type AR metabolism, it is also thought to be a critical early step in the abnormal metabolism, stabilization, aggregation, and toxicity of the polyQ expanded AR protein[133-135]. This is supported by work in an SBMA cell model where both ligands and mutations were introduced that prevented the interaction of FXXLF and AF-2 and subsequently abolished toxicity and aggregation of AR and rescued primary motor neurons from testosterone-induced toxicity[135]. There was an additional suggestion that there are coactivators of this N/C interaction including GRIP1, F-src-1, and CBP that could further stabilize the mutant
AR protein[135]. The AF-2 domain is also the site of AR’s interaction with transcription factor EB (TFEB)[136]. AR normally functions as a co-activator of TFEB; however, the expansion of the AR polyQ tract changes this interaction to an inhibitory one causing TFEB dysregulation that is responsible for the autophagic flux impairments observed in models of SBMA[136].

In a disease mechanism of activation instead of inhibition, the AR AF-2 domain also interacts with the protein arginine methyltransferase 6 (PRMT6)[137]. It has been previously shown that phosphorylation of polyQ expanded AR by Akt reduced hormone binding and AR transactivation while protecting from neurodegeneration[137]. On the other hand, PRMT6 catalytic activity is required for the arginine methylation at those Akt consensus sites and subsequent transactivation of AR[137]. So, while phosphorylation by Akt is generally neuroprotective, an enhanced interaction between PRMT6 and expanded AR and increased arginine methylation at those same Akt sites leads to neurodegeneration as demonstrated in both cell and fly models of SBMA[137]. PRMT6 is not the only post-translational modifier of mutant AR toxicity. Nemo-like kinase (NLK) can proliferate the accumulation and toxicity of polyQ-expanded AR through its binding and AR-S81 phosphorylation of an N-terminal region of AR and aberrantly increasing its activation of particular genes[138]. In another set of posttranslational studies, phosphorylation at S516 via the p44/32 MAP kinase pathway was found to be required to induce cell death in an SBMA cell model[139,140]. This finding suggests that phosphorylation enhanced the ability of caspase-3 to cleave polyQ-expanded AR and generate toxic polyQ fragments[139,140]. Ligand-dependent hyperacetylation has also been reported in expanded AR compared to wild-type and seems to be associated with toxicity in SBMA[141]. Finally, there is a differential interaction between wild and expanded AR with the post-translational modifier, deubiquitinating enzyme USP7[142]. PolyQ-expanded proteins tend to preferentially associate with USP7, which could be
an underlying common component of pathogenesis to which polyQ-expanded AR is no exception[142]. Lysine residue 17 of AR is a confirmed substrate of USP7, and an enhanced interaction with expanded AR could result in changes in the AR ubiquitination pattern that would stabilize the mutant protein and make it more prone to aggregation and toxicity[142].

Like other transcription-associated polyQ disease proteins, DNA-binding and nuclear translocation appear to be critical to expanded AR toxicity[133,143].

1.2.4 Spinocerebellar Ataxia Type 1 (SCA1)

It has been well documented that SCA1 pathology does not depend solely on the polyQ tract of the ataxin-1 protein. In fact, there are several domains that are essential for the induction of disease phenotypes even in the presence of an expanded polyQ including the ataxin-1/HBP-1 (AXH) domain, the ataxin-1 NLS, and the phosphorylation site at serine 776[2,35-37]. These three domains function both individually and in coordination with one another to shape the interactome of ataxin-1 and understanding their functions in a polyQ expanded state is essential in building the protein context ataxin-1 toxicity in SCA1

The AXH domain of ataxin-1 mediates many of its protein-protein interactions and RNA-binding toxicity and as such is responsible for many of the interactions that impact toxicity upon polyQ expansion. Possibly the most important of these AXH-mediated interactions in SCA1 pathogenesis is the polyQ-expanded ataxin-1’s interaction with the transcriptional repressor Capicua (CIC)[35,144-151]. In studies using mouse cerebellum, it was found that a majority of wild-type and expanded ataxin-1 assembles into large stable complexes with CIC[144]. This assembly seems to stabilize ataxin-1 and allow for the induction of neurotoxicity and the direct interaction between CIC and expanded ataxin-1 alters CIC’s transcriptional repression activity[144]. As further evidence of the important role of CIC interaction in ataxin-1 toxicity, mutagenesis of key amino
acids to prevent the CIC-ataxin-1 interaction also prevented toxicity in a Purkinje-cell specific SCA1 transgenic mouse model[146,151]. Additionally, haploinsufficiency of Cic reduced disease severity in SCA1 knock-in mice[150]. Ataxin-1’s incorporation into a CIC-containing large protein complex appears critical for its stabilization, particularly in an expanded state, and confers toxicity through this mechanism as well as through the alteration of CIC transcriptional repression activity.

CIC is not the only protein to interact with ataxin-1 through the AXH domain and contribute to SCA1 pathogenesis. The AXH domain also mediates ataxin-1’s interaction with Gif-1[152]. This transcription factor is necessary for the survival of Purkinje cells and is destabilized by its interaction with the AXH domain of ataxin-1[152]. Destabilization enhances Gif-1 degradation via the proteasome which contributes to the selective Purkinje cell degeneration seen in SCA1[152]. Like CIC and Gif-1, the AXH domain facilitates ataxin-1’s interaction with Tip60/retinoic acid orphan related receptor alpha (RORα)[153,154] and ATXN1L[149] in a manner that involves incorporation into larger complexes that alter the stability and activity of the complexes when the polyQ of ataxin-1 is expanded. These changes can alter the levels of expanded ataxin-1 and contribute to SCA1 pathogenesis.

The final interaction of interest through the AXH domain is with the leucine-nuclear protein (LANP). This interaction occurs at both the AXH domain and polyQ tract of ataxin-1 and occurs in a repeat-dependent manner[155]. It is thought that expanded ataxin-1 competes for LANP binding with the transcriptional repressor P120E44F (E4F)[156] and that this gain-of-function in conjunction with LANP expression pattern may explain the Purkinje cell-specific vulnerability observed in SCA1[157,158]. Further evidence of this was provided when LANP reduction reversed molecular layer thinning in SCA1 knock-in mice[157]. Apart from direct protein-protein interactions, the AXH domain itself has been found to be a key player in the aggregation of ataxin-
1. The isolated AXH domain from the ataxin-1 protein induces aggregation and forms β-plaques in vitro while expanded ataxin-1 with its AXH domain replaced with a non-aggregation prone sequence significantly reduced protein aggregation[159], further emphasizing the importance of the AXH domain in SCA1.

The next domain in establishing protein context in SCA1 is the ataxin-1 NLS. The C-terminal NLS of ataxin-1 has been shown to be necessary for disease pathogenesis, and mutation of the NLS eliminated ataxin-1-based toxicity in an SCA1 transgenic mouse model[160]. Near the C-terminal NLS is another important domain in influencing expanded ataxin-1 toxicity, the 14-3-3 binding domain. Ataxin-1 interacts with 14-3-3 proteins through this 14-3-3 binding domain and these interactions are enhanced by polyQ expansion[161]. The ataxin-1-14-3-3 interaction stabilizes ataxin-1 and enhances its protein levels allowing for increased NI formation[161].

Amidst all these interactions is a post-translational modification with critical significance in mediating ataxin-1’s ability to bind other proteins. The phosphorylation of serine 776 (S776) is necessary for the association of ataxin-1 with CIC and 14-3-3 proteins[144,162,163]. S776 phosphorylation and polyQ expansion allow ataxin-1 to more readily bind RBM17 which, in turn, interferes with RNA polymerase 2 and cell death through a toxic gain-of-function mechanism[148]. In contrast to S776 phosphorylation, SUMOylation of expanded ataxin-1 targets it for nuclear degradation via the SUMO-dependent ubiquitin ligase RNF4[164], providing an aspect of protein context with interesting potential for therapeutic investigation.

1.2.5 Spinocerebellar Ataxia Type 2 (SCA2)

The ataxin-2 protein has been implicated as having a role in several neurodegenerative diseases, including other polyQ disorders like SCA1, SCA3, and HD. In terms of its SCA2 specific protein context, ataxin-2 interacts with multiple RNA binding proteins in its wild-type and polyQ
expanded forms that provide hints towards understanding how those interactions alter its toxicity in a disease state. Ataxin-2 binding protein 1 (A2B1) is a regulator of RNA alternative splicing and was the first RNA binding protein that was discovered as having an interaction with ataxin-2[44,165]. A2B1 interacts with the C-terminal half of ataxin-2 in the residue range of 760-1313 and has been labeled in granules in SCA2 patient dentate neurons and Purkinje neurons[44]. This interaction with wild-type ataxin-2 hints at the protein’s potential role in alternative splicing, and the tissue specific sequestration in patient samples points to that role occurring in a tissue-specific manner or subset of RNAs that renders those areas or genes particularly vulnerable upon polyQ expansion[44].

Another altered ataxin-2 interaction upon polyQ expansion is with the E3 ubiquitin ligase Parkin. This ligase directly interacts with the N-terminal domain of ataxin-2 in the first 396 amino acids[165]. This interaction has been shown to occur with ataxin-2 possessing a polyQ tract in wild-type, patient, and hyperexpanded ranges in both cell and knock-in mouse SCA2 models[165-167]. In these models, ubiquitination was far more pronounced when the polyQ tract of ataxin-2 was expanded and overexpression of Parkin resulted in increased turnover of ataxin-2[165-167]. These studies indicate that Parkin may act as a regulator of ataxin-2 toxicity through ubiquitination and triggering proteasomal degradation.

Staufen1 is a regulator of stress granule formation that also has a proposed interaction with ataxin-2 and altered function upon ataxin-2 polyQ expansion[44]. In a wild-type state, ataxin-2 is thought to be involved in multiple staufen and stress granule mediated pathways; however, staufen1 expression increases in SCA2 cell and mouse models[44]. This information in combination with the finding that staufen1 overexpression causes systems to constitutively present
stress granules suggests that staufen1’s interaction with polyQ expanded ataxin-2 could be contributing to aggregate formation and increasing cellular toxicity[44].

Some final pathways of toxic influence in SCA2 are the calcium signal production and slow excitatory postsynaptic potential that are essential for neuronal communication. Calcium influxes via the transient receptor potential ration channel 3 facilitate slow excitatory postsynaptic potentials, and that channel is gated by diacylglycerol and IP3R[168]. Through an interaction with mutant ataxin-2, IPR3 is aberrantly activated and triggers an abnormal release of calcium from intracellular stores[169]. This interaction and overactivation of IPR3 by ataxin-2 has only been shown to occur when the polyQ of ataxin-2 has been expanded and does not occur with normal ataxin-2[169]. All of this is in association with the reduced mRNA translation of Rgs8 observed in SCA2 that could be a result of sequestration by stress granules[168,169]. Rgs8 is an inhibitor of mGluRS whose normal functioning in Purkinje neurons and motor neurons includes the production of local dendritic calcium signal and slow excitatory postsynaptic potential that is disrupted by the direct interaction of ataxin-2 and IP3R[168,169].

1.2.6 Spinocerebellar Ataxia Type 3 (SCA3 or Machado Joseph Disease)

While the bulk of the work that is presented in this dissertation covers the recent advancements made by our lab in establishing protein context in ataxin-3 toxicity, there is a great deal of work being done in the SCA3 field that compliments what has been uncovered here. Much of this work will be referenced and discussed in the following chapters of this work. There are, however, a few studies published more recently that have added to growing body of work surrounding ataxin-3 protein toxicity. Some highlights from these studies include, but are not limited to, the identification of pre-mRNA processing factor (Prpf19/prp19) as a regulator of expanded ataxin-3 ubiquitination and degradation[170]; the recruitment of ataxin-3 to
promyelocytic leukemia protein (PML) nuclear bodies by PML isoform II (PML-II) and reduction of its deubiquitinase activity[171]; and the proposal that, through its interaction with autophagy adaptors LC3C and GABARAP, ataxin-3 plays a role in stimulating autophagy in response to proteotoxic stress or nutrient starvation that could be disrupted following polyQ expansion[172]. These findings further the idea that protein context is critically important in understanding ataxin-3 toxicity and, like all the works discussed throughout this chapter, expand our knowledge of polyQ expansion toxicity as a whole.

1.2.7 Spinocerebellar Ataxia Type 6 (SCA6)

The protein context of SCA6 begins with the bicistronic mRNA of the disease gene CACNA1A and its IRES. While the first cistron encodes the well-characterized 𝛼1A subunit of the Cav2 P/Q-type voltage-gated calcium channel, it is the second cistron encoding the 𝛼1ACT transcription factor that harbors the polyQ expansion that causes SCA6[71,173-177]. Upon expansion, 𝛼1ACT l loses its transcription factor function and neurite outgrowth properties leading to cell death in culture and ataxia and cerebellar atrophy in transgenic mice[176]. Without the IRES, the CACNA1A gene would no longer encode the second cistron 𝛼1ACT and no protein product from the disease gene could be translated that possessed a polyQ expansion and could confer SCA6-specific toxicity[176]. In addition to the necessity of the IRES for translation of the pathogenic second cistron, there is also evidence that the polyQ-mediated toxicity is dependent on the nuclear localization of the pathogenic protein[178].

After translation of the polyQ-containing 𝛼1ACT protein, there are a few known protein-protein interactions that can alter 𝛼1ACT toxicity. The first is with the J-protein co-chaperone DnaJ-1. Expression of DnaJ-1 greatly reduces polyQ expanded 𝛼1ACT degeneration and lethality,
along with a reduction in aggregation and nuclear localization of the toxic protein[179]. The same goes for the nuclear importer karyopherin α3. Mutating this protein in the eye of Drosophila also reduced pathogenic α1ACT toxicity[179]. Overall, it seems that DnaJ-1 and karyopherin α3 are suppressors of α1ACT toxicity in SCA6.

1.2.8 Spinocerebellar Ataxia Type 7 (SCA7)

Although disruptions to transcriptional regulation and activation have already been heavily implicated in many aspects of SCA7 pathogenesis[78], there are several domains and protein-protein interactions with the ATXN7 protein itself that are worth highlighting in this review. The first of which is the cleavage of the expanded ataxin-7 protein by caspase-7. Aspartic acid residues at positions 266 and 344 of ataxin-7 have been identified as caspase-7 cleavage sites[82]. Caspase-7 cleavage at these sites appears to be a critical step in SCA7 pathogenesis as preventing cleavage through mutation of polyQ-expanded ataxin-7 mitigated the cell death, aggregate formation, and transcriptional interference normally observed in SCA7[82]. In addition, expression of the caspase-7 truncation product of ataxin-7 with two different pathogenic repeat lengths that also removes the ataxin-7 NES and NLS increased cellular toxicity[82]. One hypothesis is that the active form of caspase-7 cleaves ataxin-7 in the nucleus rather than the cytoplasm, and that the location of ataxin-7’s cleavage sites would separate the polyQ tract from the NES and disrupt export of ataxin-7 out of the nucleus[82]. This nuclear retention of polyQ-expanded ataxin-7 could lead to the accumulation of toxic N-terminal fragments of ataxin-7 that could impair or sequester ataxin-7’s normal interactors in the SAGA complex[82]. The accumulation of expanded ataxin-7 fragments has also been associated with an increased level of acetylation at lysine 257 (K257)[180]. This acetylation site is close to one of the two caspase-7 cleavages sites and seems to prevent the clearance of expanded ataxin-7[180].
K257 is also the site for ataxin-7 SUMOylation and polySUMOylation[181,182]. In this case, SUMOylation is having a contrasting effect at this site compared to acetylation. SUMO2 is added to the polyQ-expanded ataxin-7 substrate and this modification leads to the recruitment of SUMO-targeted ubiquitin ligase RNF4[181,182]. Ataxin-7 is then ubiquitinated by RNF4 and, following recruitment of clastosomes, is subsequently degraded by the proteasome[181,182]. SUMO2 has been colocalized in ataxin-7 NIs, and while SUMOylation at K257 seems to promote the clearance of expanded ataxin-7 it is possible that chronic expression of this polyQ protein could produce enough misfolded proteins to eventually overwhelm the SUMO pathway and clastosomes, especially as they become compromised with age[181,182]. An overwhelmed processing system would cause aggregate sizes to increase and possibly explain the delayed onset of the disease.

Along with caspase cleavage and posttranslational modification, the interface between ataxin-7 and components of the SAGA complex presents numerous interactions that all play a role in expanded ataxin-7 toxicity. Upon polyQ expansion, ataxin-7’s interaction with the SAGA complex alters SAGA function through the sequestration of SAGA components into NIs[183-187], the undermining SAGA integrity[188,189], and aberrantly interacting with SAGA components and associated transcription factors[81]. The cone-rod homeobox protein (Crx) is a photoreceptor transcription factor that has been shown to have a polyQ-dependent interaction with ataxin-7[190,191]. The Crx and STAGA (human ortholog to yeast SAGA) interaction is mediated by ataxin-7 and controls Crx-dependent gene activation in the retina[188,192]. PolyQ expansion of ataxin-7 disrupts this pathway leading to the downregulation of retinal specific genes and contributing to disease progression, as seen in transgenic SCA7 knock-in mice[188,190-193]. PolyQ expanded ataxin-7 has also been shown to aberrantly bind the Gcn5 catalytic core along with its adaptor proteins, Ada2 and Ada3, through a highly conserved zinc binding domain between
residues 311 and 406 of ataxin-7\[81\]. This interaction caused significant decreases in Gcn5 histone acetyltransferase activity in vitro and in vivo\[81\]. These findings in SAGA complex components suggests that polyQ expanded ataxin-7 dominantly inhibits the function of the SAGA complex and helps explain the prominence of transcriptional dysregulation in SCA7 pathogenesis.

Other interesting protein-protein interactions with expanded ataxin-7 include histone deacetylase 3 (HDAC3), whose activity increases in SCA7 mouse retinas and whose interaction with ataxin-7 stabilizes the protein in a deacetylase-independent manner\[194\]. Ataxin-7, through its N-terminal region of residues 1-230, is also able to interact with the SH3P12 gene products, R85 and CAP, in both a wild-type and expanded state\[195\]. It is suggested that in a wild-type state, this interaction helps to facilitate the physiological regulation, ubiquitination, and degradation of ataxin-7; however, when ataxin-7 is expanded, ubiquitinated proteins still accumulate in NIs in SCA7 brains suggesting that the normal function of these gene products is disrupted and they may be contributing to SCA7 progression\[195\]. The final ataxin-7 interaction that will be discussed in this section is with the ATPase subunit S4 of the 19S regulatory complex of the proteasome. Uncovered in a yeast-two-hybrid assay, the ataxin-7-S4 interaction is inversely correlated with the length of the polyQ tract, and it is suggested that S4 functions in the ubiquitin-mediated proteasomal degradation of ataxin-7\[196\]. As the polyQ tract becomes expanded, and the interaction between S4 and ataxin-7 weakens, it may create abnormalities in the recognition or degradation of ubiquitin protein conjugates by the 26S proteasome complex and contribute to accumulation of ataxin-7/ubiquitin-protein conjugates and SCA7 pathogenicity\[196\].

1.2.9 Spinocerebellar Ataxia Type 17 (SCA17)

Dysfunction of TBP results in broad clinical presentations ranging from Parkinson’s to Huntington’s to psychiatric symptoms and may have impacts beyond just SCA17\[83,86,87\]. In
Drosophila, loss of TBP expression not only caused age-related neurodegeneration in SCA17 disease fly models, but also exacerbated polyQ-expansion-induced retinal degeneration in disease models of SCA3 and HD[197]. Findings like these suggest that TBP dysfunction may play a larger role in polyQ-induced neurodegeneration and highlight the significance of understanding the protein context of TBP toxicity to clarify the causes of SCA17 and the entire family of polyQ diseases.

One such TBP protein context change upon polyQ expansion comes from a potential conformational change that results in a disruption of one of TBP’s native functions, binding DNA. It has been reported that the polyQ expansion of TBP reduces its in vitro binding to DNA and that mutant TBP fragments that lack an intact C-terminal DNA-binding domain have been identified in transgenic mouse models of SCA17[198]. In addition to this reduced binding, polyQ-expanded TBP with a deletion in the DNA-binding domain that prevents DNA-binding has been shown to form nuclear aggregates and inhibit TATA-dependent transcription in cultured cells as well as causing early death and neuronal nuclear inclusions in transgenic mice with the same double mutant[198]. All these findings suggest that the expanded polyQ tract of TBP negatively affects its ability to bind promoter DNA and that this disruption can induce neuronal aggregation and toxicity in SCA17.

In addition to DNA-binding, TBP also has several new or altered protein-protein interactions that occur because of its polyQ expansion. The first is with the transcription factor nuclear factor-Y (NF-Y). It was found that in TBP transgenic mice, expanded TBP bound more tightly to NF-Y and seemed to sequester it into TBP nuclear inclusions, possibly altering its function as a master regulator of the chaperone system and impairing NF-Y-mediated expression of chaperones like Hsp70, Hsp25, and HspA5[199]. Expanded TBP has also been found to
abnormally bind another transcription factor, Sp1, and seems to sequester it away from its normal function in a manner similar to NF-Y[200,201]. In this case, it was shown in SCA17 knock-in mice that the mutant TBP-Sp1 interaction leads to reduced transcription of Inpp5a, a gene that is critical in the maintenance of Purkinje cells and neurite growth pathways and whose knockout in wild-type mice results in Purkinje cell degeneration[200,201]. A final example of the polyQ expansion enhancing TBP’s interaction with a protein is another transcription factor, RBP-J/Su(H)[202]. This transcription factor participates in Notch signaling and was shown in Drosophila to interact more efficiently with expanded TBP than wild-type TBP[202]. Its dysfunction, and subsequent alteration of the Notch signaling pathway, contributed to TBP expanded polyQ-induced phenotypes[202]. These three interactions form a trend that points to the possibility of enhanced interaction of mutant TBP with other transcription factors which – like NF-Y, Sp1, and RBP-J/Su(H) – could be sequestering them away from their normal functions leading to toxic transcriptional alterations.

While polyQ expansion increases several TBP interactions, there are also cases where it does the opposite. The transcription factor XBP1 recognizes the ER stress response element in the promoter of MANF, an ER-stress inducible protein that is enriched in Purkinje cells[203,204]. XBP1 and TBP are normally present in the same transcriptional complex to help mediate the expression of MANF; however, mutant TBP has decreased association with XBP1 leading to reduced expression of MANF and a lowered ER-stress response[203,204]. Another normal interaction that is impacted by the polyQ expansion of TBP is with the muscle-specific transcription factor MyoD[205]. The TBP-MyoD interaction stabilizes MyoD’s binding to a DNA promoter, but the association between TBP and MyoD is reduced upon TBP polyQ expansion causing decreased transcriptional activity and protein levels because of destabilization and enhanced degradation[205]. The lowered MyoD levels and activity negatively impacts muscle-specific gene
expression and leads to muscle degeneration in SCA17 knock-in mice[205]. This model also showed that larger polyQ repeats in TBP created an even greater reduction in association with MyoD and more severe muscle degeneration[205].

1.3 Conclusions

What has been covered in this review only begins to scratch the surface of all the great investigations that have helped advance the understanding of the protein context surrounding the nine polyQ disease proteins. There are countless examples of toxic effects specific to each disease that are outside the scope of this work but certainly provide valuable information about the contributors to polyQ protein toxicity. These contributors may not involve direct protein interactions, but mechanisms of toxicity such as transcriptional dysregulation, sequestration of essential proteins away from their normal functions and into aggregates, and non-protein level contributions like RAN translation and mRNA toxicity all play an important role in establishing the distinct pathogenesis of each disease. Additionally, there are several areas of overlapping protein context between these disease proteins that share functions, interactions, and pathway disruptions that should be investigated further as they may hold clues to broader therapeutic approaches that could target polyQ disorders more generally.

The ever-growing list of investigations into protein context have been a continual inspiration for this dissertation and have laid the groundwork for the studies building the protein context of the ataxin-3 protein highlighted over the next five chapters. Included verbatim are four manuscripts that have either been published or submitted for publication at the time of this writing.
2 CHAPTER 2: DIFFERENTIAL TOXICITY OF ATAXIN-3 ISOFORMS IN DROSOPHILA MODELS OF SPINOCEREBELLAR ATAXIA TYPE 3

2.1 Introduction

Several ataxias are caused by specific genetic mutations. Among these is the most common ataxia in the world, Spinocerebellar Ataxia Type 3 (SCA3), also known as Machado-Joseph Disease[53,206]. An age-related disorder, SCA3 is a member of the family of polyglutamine (polyQ)-dependent diseases alongside other SCAs (1, 2, 6, 7 and 17), Huntington’s Disease, Kennedy’s Disease, and Dentatorubropallidoluysian Atrophy. PolyQ diseases are caused by abnormal lengthening of a CAG triplet repeat in the respective genes, which translates into an expanded polyQ stretch in the encoded proteins. Expanded polyQ in each of these disease proteins causes misfolding and the formation of aggregates, neuronal dysfunction and death [2-4,207-210].

SCA3 is a progressive ataxia accompanied by dystonia, dysarthria, spasticity, rigidity, ophthalmoparesis, dysarthria, dysphagia and neuropathy. Pathology includes degeneration of cerebellar pathways and nuclei, pontine and dentate nuclei, substantia nigra, globus pallidus, cranial motor nerve nuclei and anterior horn cells. SCA3 is caused by abnormal triplet CAG repeat expansion in the gene ATXN3 that is normally 12-42 repeats in length, but is expanded to ~52-84 in patients[53,206]. Expansion occurs in the protein ataxin-3, a deubiquitinase (DUB) implicated in quality control[53,206]. Although ataxin-3 is widely expressed, degeneration is constrained to specific areas of the brain. Histopathologically, the SCA3 protein is found in macroaggregates/inclusions in the cytoplasm and in the nucleus[51,53,211].

Extensive work by various laboratories has provided insight into enzymatic properties and potential functions of the SCA3 protein. As a DUB, ataxin-3 binds and cleaves ubiquitin chains of a specific length and type[53,206,212,213]. Ataxin-3 may have different types of substrates,
depending on the needs of the cell, and might even function as a monitor of ubiquitin homeostasis. In cultured mammalian cells and in the fruit fly, *Drosophila melanogaster*, ataxin-3 is important during proteotoxic stress[214-216]. *In vivo*, according to different *atxn3* knockout lines, this protein appears dispensable[206,217-220].

PolyQ expansions in ataxin-3 most likely cause disease though a toxic gain-of-function, although the precise nature of the dominant toxic properties remains elusive. What is clear is the fact that nuclear ataxin-3 is markedly more toxic than its cytoplasmic portion, and that aggregation of disease-causing ataxin-3 precedes and correlates with cellular dysfunction and disease [64,221,222]. Preventing aggregation of the SCA3 protein is beneficial to cells, tissues and organisms harboring polyQ-expanded ataxin-3 [53,206,216,220,222,223]. Still, our understanding of SCA3 biology remains incomplete.

Towards understanding SCA3, we reasoned that it will be beneficial to comprehend the toxicity of the products of its causative gene, *ATXN3*. Two major isoforms of ataxin-3 arise by alternative splicing of *ATXN3* mRNA [1]. The original ataxin-3 variant isolated from SCA3 human brain[62], which we refer to as isoform 2 in this work, encodes an isoform containing 2 Ubiquitin-interacting motifs (UIMs), the polyQ repeat and a C-terminal stretch of hydrophobic amino acids. Subsequent studies identified another variant, referred to as isoform 1, that encodes a third UIM at its C-terminus [224] (figure 2.1). Isoform 1 of ataxin-3 is the predominant form in the mammalian brain; isoform 2 is also present in this organ, but at lower levels than isoform 1 [1].

Due to a lack of reagents that can readily and reliably distinguish these isoforms, it is unclear whether specific forms are enriched in certain types of cells or tissues in the normal and diseased brain.
Isoform 2 appears to aggregate more rapidly \textit{in vitro} and in cultured mammalian cells. In cultured cells, isoform 2 protein is also less stable than isoform 1\cite{1}. These informative observations notwithstanding, a comprehensive understanding of the relative toxicity and pathogenic contribution of each isoform is presently lacking \textit{in vivo}. Here, we investigated the toxicity of each isoform of pathogenic ataxin-3 by generating novel, isogenic \textit{Drosophila melanogaster} transgenics that encode ataxin-3 protein with a polyQ of 80, within human disease range. We observe that isoform 1 of ataxin-3 is much more toxic than isoform 2 as a result of higher turnover and lower steady state levels of the latter variant. We also find that both isoforms are degraded by the proteasome and autophagy. Collectively, our results yield new information on the pathogenicity of the ataxin-3 isoforms \textit{in vivo}, present new genetic tools to be used to further assess toxic products of the \textit{ATXN3} gene and expand overall understanding of SCA3 biology.

2.2 Materials and Methods

2.2.1 Constructs and plasmids

We used the company Genscript (genscript.com) to generate human ataxin-3 cDNA with a CAGCAA repeat that encodes 80Q. The nucleotide sequences used for isoforms 1 and 2 were based on ataxin-3 sequences described in prior publications\cite{1,212,223,225}, but with a CAGCAA repeat instead of a pure CAG tract to circumvent the possibility of RAN translation and mRNA toxicity \cite{226-233}. Constructs were sub-cloned into pcDNA3.1 for mammalian expression and pWalium-10.moe for fly line generation by using EcoRI (5’) and XbaI (3’) restriction digest sites. A Kozak sequence, shown in figure 2.1D, was added 5’ onto the ataxin-3 cDNA sequence for fly expression and the endogenous Kozak of pcDNA3.1 was utilized for expression in mammalian cells. An HA tag-encoding sequence was added immediately after the last nucleotide of the ataxin-3 sequence, as shown in figure 2.1A. Each construct was sequence-verified before utilization for
its intended purposes. Genscript was also used to synthesize the ataxin-3 constructs in figure 2.9 and sub-clone them into pcDNA3.1 through EcoRI (5’) and XbaI (3’) digestion/ligation.

2.2.2 Drosophila stocks and new transgenic lines

All fly stocks and crosses were maintained at 25°C and ~60% humidity in diurnal incubators with 12 h light/dark cycles. Common stocks were from Bloomington Drosophila Stock Center (BDSC). New transgenic lines were generated through phiC31-dependent integration of transgenes into the attP2 site of the fly’s third chromosome, as we have done in the past[174,216,222,223]. New lines generated were verified to ensure insertion at the right site (PCR-based assays), in the correct orientation (PCR-based assays) and sequence-verified for transgene nucleotide sequence. The lines generated for the work in this manuscript all contain a CAGCAA repeat to encode the expanded polyQ of ataxin-3. The only exception is the fly line that we denote here as “weaker expressing isoform 1”, which contains a pure CAG repeat and was described in our prior publications[222,223,224]. This latter transgene is also inserted on chromosome 3 of Drosophila, and on the same genetic background as the other two isoform lines.

Fly lines used here are as follows: GMR-Gal4 (#8121) and the isogenic host strain attP2 (#36303) were from the Bloomington Drosophila Stock Center (BDSC). RNAi lines were from BDSC and the Vienna Drosophila RNAi Center (VDRC): RNAi to Atg7 (BDSC34369 (#1), BDSC27707(#2)); Atg8a (BDSC34340 (#1), BDSC28989(#2)); Atg12 (BDSC27552 (#1), BDSC34675 (#2)); Rpn1 (BDSC34348 (#1), VDRC25549 (#2), VDRC103939 (#3)); prosα5 (BDSC34786); prosβ2 (BDSC67363). The following lines were gifts: sqh-Gal4 driver (Dr. Daniel P. Kiehart, Duke University), elav-Gal4(GS) driver (Dr. R. J. Wessells, Wayne State University),
elav-Gal4 driver and repo-Gal4 driver (Dr. Daniel F. Eberl, University of Iowa). Unless otherwise stated in the figure legends and text, all flies were heterozygous for driver and transgene.

2.2.3 Longevity and motility assays

Adult flies were collected after eclosion from their pupal case and were aged at 25°C on conventional cornmeal fly media or RU486-containing media, as noted in the text and figure legends. Flies were switched to vials with fresh food every 2-3 days. Each vial was examined for death each day, until all flies were dead. Motility was tested weekly through negative geotaxis. Approximately 10 flies per vial were forced to the bottom of the vial by gentle tapping on the bench. The total number of adult flies that reached the top of the vial at 5, 15 and 30 seconds was recorded and expressed as a percentage of total flies. Adults were transferred to fresh vials 1 hour before assessment, and every 2-3 days for the duration of the experiment; their longevity was monitored daily.

2.2.4 Mammalian cells and assays

HeLa and HEK-293T cells were purchased from ATCC and cultured in DMEM supplemented with 10% FBS and 5% Penicillin-Streptomycin, under conventional conditions. The day before transfection, cells were seeded in 12-well plates. Cells were transfected using Lipofectamine LTX (Invitrogen) per manufacturer’s directions. Twenty-four hours after transfection, cells were treated as indicated in figure legends with cycloheximide (100 µg/ml; A.G. Scientific), and/or MG132 (20 µM; R&D Systems), or chloroquine (100 µM, A. G. Scientific). Cells were harvested in hot SDS lysis buffer for western blotting (see below).
2.2.5 Western blotting

Five whole flies, 10 pupae or 10 dissected adult fly heads per group were homogenized in boiling lysis buffer (50mM Tris pH 6.8, 2% SDS, 10% glycerol, 100mM dithiothreitol), sonicated, boiled for 10 minutes, and centrifuged at 13,300 RPM at room temperature for 10 minutes. The same lysis buffer and procedures were used for extraction of protein from mammalian cells. Samples were electrophoresed on 4-20% gradient gels (BioRad). Western blots were developed using the charge-coupled device-equipped VersaDoc 5000MP system (BioRad) or Syngene PXi 4. Blots were quantified using Quantity One software (BioRad) or GeneSys (Syngene). Direct blue staining was used for loading: PVDF membranes were submerged for 5 minutes in 0.008% Direct Blue 71 (Sigma-Aldrich) in 40% ethanol and 10% acetic acid. PVDF membrane was then rinsed briefly in 40% ethanol and 10% acetic acid solvent, then ultrapure water, air dried, and imaged.

2.2.6 Differential Centrifugation

Ten whole flies per group were lysed in 200 µL of NETN buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.5% Nonidet P-40), supplemented with protease inhibitor cocktail (PI, SigmaFast Protease Inhibitor Cocktail Tablets), sonicated at 50% for 15 seconds with a probe solicitor (Thermo Scientific) and centrifuged at 20,000 x g for 30 minutes at 4°C. The supernatant was transferred to a new microfuge tube and quantified by the BCA assay (Thermo Scientific). The pellet was resuspended in 200 µL of PBS + 1% SDS, vortexed, and boiled. Thirty µg of the supernatant fraction and 7 µL of the pellet fraction were supplemented with 6x SDS, boiled, and loaded onto SDS-PAGE gels gels for western blotting.
2.2.7 Filter-Trap Assay

Three whole flies per group were lysed in 200 µL of NETN+PI. Lysates were diluted with 200 µL PBS containing 0.5% SDS, sonicated at 50% for 15 seconds with a probe sonicator, and centrifuged at 4,000 x g for 1 minute at room temperature. 100 µL of supernatant were diluted with 400 µL PBS. 70 µL of sample were filtered through a 0.45 µm nitrocellulose membrane (Schleicher & Schuell) using a Bio-Dot apparatus (BioRad) and analyzed by Western blotting.

2.2.8 Nuclear/Cytoplasmic Extraction

Fractionation was conducted using the ReadyPrep Protein Extraction Kit (Cytoplasmic/Nuclear; BioRad) to separate cytoplasmic and nuclear proteins. Five whole adult flies were lysed in cytoplasmic extraction buffer (BioRad). Nuclei were resuspended in Protein Solubilization Buffer (BioRad). Samples were analyzed by western blotting, where, proportionally, 3X nuclear fraction was loaded compared to cytoplasmic fraction to eliminate need for over-exposure. Through this protocol, sometimes we are able to observe a main band for ataxin-3, and other times we are not, likely as a result of the buffers used. This is highlighted through two different examples in figures 2.2C and S2.2. For quantification purposes, we quantified signal from the entire lane, including main band when visible, as well as the smear above it.

2.2.9 Antibodies

Anti-Ataxin-3 (mouse monoclonal 1H9, 1:500-1,000; Millipore) (rabbit polyclonal, 1:15,000; [211]), anti-GAPDH (mouse monoclonal MAB374, 1:500; Millipore), anti-tubulin (mouse monoclonal T5168, 1:10,000; Sigma-Aldrich), anti-HA (rabbit monoclonal, 1:500-1,000; Cell Signaling Technology), anti-lamin (mouse monoclonal ADL84.12-5, 1:1,000; Developmental Studies Hybridoma Bank), peroxidase conjugated secondary antibodies (goat anti-mouse and goat anti-rabbit, 1:10,000; Jackson Immunoresearch).
2.2.10 qRT-PCR

Total RNA was extracted from whole adult flies using TRIzol (Life Technologies). Extracted RNA was treated with TURBO DNase (Ambion). Reverse transcription was carried out using the high capacity cDNA reverse transcription kit (ABI). mRNA levels were quantified using the StepOne Real-Time PCR system with Fast SYBR Green Master Mix (ABI). Primers used:

ATXN3 F: 5’-GAATGGCAGAAGGAGGAGTTACTA- 3’;
ATXN3 R: 5’-GACCCGTCAGAGAGAAATTCAAGT- 3’;
rp49 F: 5’-AGATCGTGAAGAAGCGCACCAAG- 3’;
rp49 R: 5’-CACCAGGAACCTTTCTTGAATCCGG- 3’

2.2.11 Statistics

Student's t-tests, ANOVA with post-hoc assays and log-rank tests were conducted for statistical comparisons, as indicated in the figure legends. Software utilized was GraphPad Prism and Apple Numbers. Graphs were generated in Apple Numbers.

2.3 Results

2.3.1 New Drosophila lines to investigate SCA3

As summarized in figure 2.1A, earlier studies[62-64,235-237] established that two major, full-length isoforms arise from alternative splicing of ATXN3 mRNA; a polymorphism in what we term isoform 2 in this study also gives rise to a slightly shorter variant, denoted by the arrowhead in figure 2.1A. The ataxin-3 variants are identical, except for their respective C-termini. As noted above, both of these forms are present in mammalian brain, but isoform 1 is predominant. Here we focus on the pathogenic property of the two full-length isoforms, depicted in the lower part of
We generated constructs that encode either human ataxin-3 isoform, with an expanded polyQ (80), which is within patient range. The transgenes are used through the Gal4-UAS binary expression system in the fly [238], which enables expression in a tissue- and time-dependent manner [174, 216, 222, 223].

We designed the polyQ of ataxin-3 to be encoded by CAGCAA doublet repeats. Expanded ATXN3 may lead to neuronal toxicity not only through the ataxin-3 protein, but also via mRNA toxicity as a result of a long, homogenous CAG trinucleotide repeat [228]. Moreover, in addition to encoding polyQ, the expanded CAG repeat of ATXN3 could lead to repeat-associated non-AUG (RAN) translation of non-polyQ homopolymers, as observed in other polyQ disorders [239-241]. For the purposes of this study, we sought to focus on toxicity from the polyQ-expanded ataxin-3 protein. Therefore, to avoid the potential confound of RAN translation and mRNA toxicity, the polyQ of ataxin-3 is encoded by alternating CAGCAA repeats. This approach precludes the production of proteins in alternative frames and eliminates potential ATXN3 mRNA toxicity [226-233]. The polyQ-reading frame has an HA epitope tag at its C-terminus and we capped the other two sense reading alternatives with in-frame MYC and V5 epitope tags (figure 2.1B). As shown in figure 2.1B, we do not observe specific protein signal from unintended reading frames. Blotting for the HA epitope tag for the polyQ frame yields clear protein signal, whereas blotting for MYC and V5 tags (alternative reading frames of the sense strand) does not yield signal above background, indicating that RAN translation is unlikely to be occurring with this design. We say “unlikely” since there is still the formal possibility of RAN from anti-sense frames, for which we were unable to generate in-frame tags. Nonetheless, based on the results in figure 2.1B and other similar data not included, as well as on prior work showing that CAGCAA repeats circumvent RAN [226-233], we are confident that any RAN translation occurring from our constructs is likely
negligible. We therefore proceeded with using the CAGCAA design to encode the polyQ of both isoforms. We reiterate here that both isoforms contain 80Q, which is within the range of polyQ repeats observed in SCA3 patients[206].

The new *Drosophila* lines were generated with the phiC31 site-specific integrase system, which enables insertion of a single construct in the same orientation at site attP2 on the third chromosome of *Drosophila* [242]. The new ataxin-3 isoform 1 and 2 lines were verified for

---

**Figure 2.1 New Drosophila lines for SCA3.**

A) ATXN3 isoforms. U: untranslated region; DUB: deubiquitinase; polyQ: polyglutamine; UIM: ubiquitin-interacting motif. 2-UIM transcript is termed here isoform 2 since it is less common[1]. Isoform 1 results from alternative splicing linking exon 10 to 11 for UIM3 (●). (★) Isoform 2 with a different tail, lacking UIM3. (▲) A SNP results in an early stop within isoform 2. In this work, both isoforms are capped by an HA tag. B) Ataxin-3 constructs were designed with a CAGCAA repeat to encode polyQ. PolyQ frame is HA-tagged. Isoform 1 is shown. The other, non-polyQ sense frames are MYC and V5 tagged. NS: non-specific. C) qRT-PCR from adults expressing either isoform in all neurons. Means ± SEM. P value from student's t-test. Flies were one day old. D) Start and end sequences for the ataxin-3 lines from genomic sequencing. Italics: nucleotides. Sequence between nucleotides is ataxin-3 amino acid sequence. Bold: sequences added for Kozak (front) and HA tag (end).
insertion at the proper site, in the same orientation and as a single copy through PCR, and were also sequence-verified from genomic DNA [174,179]. The phiC31 integrase system enables expression of transgenes at a similar level, as shown in figure 2.1C and as demonstrated in our prior work with other proteins and transgenes[174,216,222,223]. Figure 2.1D shows that the “start” and “end” genomic DNA sequencing results of the ataxin-3-encoding areas in the new fly lines are identical.

When we examine the protein levels of isoforms 1 and 2 when expressed in all fly neurons as well as in all fly tissues, we notice clear differences between the two. Even though at the mRNA level the two isoforms are similar and their “start” regions are the same (figure 2.1C and D), at the protein level they differ markedly. Isoform 1 protein is present at much higher levels than isoform 2 when expressed in all fly neurons (throughout development and in adults; figure 2.2A) as well as in all fly tissues (figure 2.2B). This difference is not due to the antibody used, as anti-HA blots show a similar difference between the two isoforms at the protein level (supplemental figure S2.1).

The significantly lower levels of isoform 2 protein compared to isoform 1, coupled to our observations that their mRNA levels are similar, suggest that isoform 2 protein is degraded more rapidly in flies. This difference in protein levels between the two full-length ataxin-3 isoforms is not entirely unexpected, as prior work had shown that isoform 2 is more rapidly degraded in cultured mammalian cells [1].

Next, we examined whether the two ataxin-3 isoforms are similarly partitioned in the nucleus and the cytoplasm. This is an important point since nuclear pathogenic ataxin-3 is significantly more toxic than cytoplasmic ataxin-3 [221]. Figure 2.2C and supplemental figure 2.2 summarize our findings from cytoplasmic/nuclear fractionations: both isoforms are found
similarly in these two fractions. With these new tools and an initial impression about their expression, protein levels and localization, we next investigated their toxicity in the fly.

Figure 2.2 Expression of the two ataxin-3 isoforms in flies.
A & B) Western blots when isoforms are expressed in all neurons (A) or ubiquitously (B). Adults were used for panel A and pupae for panel B. For quantification, the entire signal in each lane was used. Means ± SD. P values from student’s t-tests. Black arrow: unmodified ataxin-3. Blue arrows: ubiquitin-inated ataxin-3[215,225,243,244]. Black bar: SDS-resistant species. In all blots, isoform 2 migrates slightly lower than isoform 1. C) Western blots from cytoplasmic (C) / nuclear (N) separation of fly lysates expressing isoforms in all neurons. The entire signal - the main band as well as smear above it - was quantified, as denoted by the bar on the right side of the blots. Isoforms were loaded on separate gels to avoid over-exposure. Means ± SD. N.S.: non-significant P value from student's t-test. Flies were one day old.

2.3.2 Isoform 2 of pathogenic ataxin-3 is less toxic than isoform 1 in Drosophila
We assessed the toxicity of isoforms 1 and 2 of pathogenic ataxin-3 in Drosophila with the aid of several Gal4 drivers that we have employed in the past to examine the toxic properties of various proteins[174,179,215,216,222,234]. We began by expressing either isoform in all tissues, throughout fly development and in adults, via the sqh-Gal4 driver [245-248]. As shown in figure 2.3A, ubiquitous expression of isoform 1 leads to developmental lethality in flies: flies die during
late pupal and early pharate adult stages, before they can emerge as adults. Expression of isoform 2 instead leads to successful eclosure of adults from their pupal cases; these flies survive for nearly 60 days. Control flies, which contain the sqh-Gal4 driver on the background line of the chromosome targeted for transgene insertion, live up to ~90 days. Co-expression of both isoforms in all tissues leads to a phenotype that is more severe than expression of isoform-1 alone (supplemental figure S2.3).

Figure 2.3 Differential toxicity from pathogenic ataxin-3 isoforms in developing and adult flies.
A & B) Adult fly longevity when isoform 1 or 2 of pathogenic ataxin-3 was expressed in the noted tissues. Ubiquitous driver was sqh-Gal4, pan-neuronal driver was elav-Gal4. P values from log-rank tests. C) Negative geotaxis assays of flies with the noted genotypes. P values from student's t-tests comparing motility of isoforms 1 and 2 to control flies for their respective time course. None of the flies set aside for motility assays for isoform 1 survived to day 21. At least 50 flies per group were used.

These data were corroborated by expression of either pathogenic ataxin-3 isoform in other tissues. As shown in figure 2.3B, expression of isoform 1 in all fly neurons, during development
and in adults, leads to higher lethality than isoform 2. Flies expressing isoform 1 in all neurons successfully emerged as adults, but die within one month. Isoform 2-expressing adults live for about 80 days, and the control flies, which contain the elav-Gal4 driver but without ataxin-3 expression, die by ~100 days. Observations from longevity assays were complemented with motility tests, performed with the negative geotaxis assay. Through this method, flies in a vial are forced to the bottom by tapping and the number of flies that reach the top of the vial at 5, 15 and 30 seconds is calculated and expressed as a percentage of total flies in the vial. As shown in figure 2.3C, flies expressing isoform 1 pan-neuronally are impaired in their climbing ability compared to control flies and flies expressing isoform 2. Slower mobility from isoform 1-expressing flies is noticeable as early as day 7 and escalates throughout testing. Isoform 2-expressing flies also show reduced ability to climb, but later than isoform 1 counterparts. Still, for most of the time that we conducted this study, isoform 2 flies resembled controls (figure 2.3C).

Data from figures 2.3B and C were obtained using a driver that expresses ataxin-3 isoforms throughout development and adulthood. We next assessed whether expression of either isoform in all fly neurons solely at adult stages shows a different pattern of lethality. The rationale behind this approach rests on the fact that SCA3 is an adult-onset disease. What if the differential toxicity we observe with isoforms 1 and 2 of SCA3-causing ataxin-3 is due to developmental problems? We utilized a pan-neuronal driver, elav-Gal4-GS (Gene Switch), which requires the compound RU486 to initiate expression of ataxin-3 (figure 2.4A) [249-251]. Flies are raised in media without RU486. On day 1 after they emerge as adults, flies are placed in media with RU486 to induce protein expression [222,252]. Flies are then maintained in media with RU486 until they die and longevity is tracked daily. As summarized in figure 2.4B, expression of either isoform leads to increased
lethality compared to controls, with lethality occurring earlier with isoform 1. Control flies contained the elav-Gal4-GS driver on the genetic background used to generate ataxin-3 lines.

Figure 2.4 Differential toxicity from pathogenic ataxin-3 isoforms in adult neurons and in glia.
A) Western blots from whole fly lysates. Flies heterozygous for both elav-Gal4-GS driver and either isoform were raised in media without RU486 until they emerged as adults, when they were placed in RU486-containing media for the indicated amount of time. NS: non-specific bands we sometimes observe with the indicated antibody. NS*: another non-specific band present only in lanes 1 and 5, running immediately below/near the ataxin-3-positive band. In the isoform 2 blot, 4/5th less lysate was loaded for isoform 1 compared to the other lanes. Iso: isoform. B, D) Adult fly longevity when isoform 1 or 2 of pathogenic ataxin-3 was expressed in the noted tissues. Pan-neuronal driver dependent on RU486 was elav-Gal4-GS, and glial cell driver was repo-Gal4. P values from log-rank tests. C) Western blots from adult flies expressing isoform 1 or 2, driven by repo-Gal4. Shown in histograms are means ± SD. P value from student's t-test. Flies were one day old.

Finally, we targeted either isoform to glial cells using the Repo-Gal4 driver, which leads to transgene expression in development and adults. Here, we notice again markedly lower protein levels in glial cells for isoform 2 compared to isoform 1 (figure 2.4C). As shown in figure 2.4D, expression of isoform 1 again leads to accelerated incidence of death than expression of isoform 2; only a handful of adult flies are able to emerge and die within 12 days. Most isoform 1-expressing flies do not manage to eclose from their pupal cases. By contrast, isoform 2 is not toxic to glial cells; adults that express isoform 2 only in glia emerged successfully as adults and live
similarly to control adults that contain the Gal4 driver but do not express isoform 2. Collectively, these observations demonstrate that isoform 1 of pathogenic ataxin-3 is significantly more toxic than isoform 2 in flies.

2.3.3 Aggregation of isoforms 1 and 2 of pathogenic ataxin-3 in flies

A defining property of pathogenic ataxin-3 is its propensity to aggregate[206]. This feature presents as macro-inclusions in SCA3 mammalian brain and in cultured cells, its partitioning into soluble and insoluble (pellet) species through centrifugation protocols, and the retardation of aggregated species on filter-trap assays [53,213,220,222,253-256].

We first utilized centrifugation to separate each isoform of pathogenic ataxin-3 into a pellet and soluble fraction. For this set of experiments, we used the elav-Gal4 driver to express either ataxin-3 isoform in all neuronal cells, and adults were collected on days 1, 7 and 14. Due to the much higher levels of isoform 1 protein than isoform 2, blots in figure 2.5A were imaged separately to avoid the need for over-exposure to obtain isoform 2 signal. We notice that each isoform of pathogenic ataxin-3 partitions both into soluble and pellet fractions on day 1 (figure 2.5A). However, unlike isoform 1, which by two weeks is mostly found in the pellet fraction, isoform 2 continues to be present in both fractions. Additionally, at every time point more isoform 1 fractionates into the pellet than isoform 2, indicating a higher aggregative propensity of this variant of the SCA3 protein (figure 2.5A). Another feature of isoform 1 that we do not observe with isoform 2 is the clear presence of species at higher molecular weight (HMW) regions present in the pellet fraction (denoted by black lines in figure 2.5A). These HMW are SDS-resistant species of pathogenic ataxin-3 that we have reported before and which accumulate in fly tissue [222]. Their unambiguous presence with isoform 1 underscores the higher propensity of isoform 1 to aggregate than isoform 2.
Figure 2.5 Ataxin-3 isoform aggregation in fly neurons.
A) Western blots from adult flies expressing isoform 1 or 2 of pathogenic ataxin-3 in all neurons throughout development and in adults, harvested on the indicated days. Fly lysates were separated into soluble (S) and insoluble (P) fractions and loaded onto SDS-PAGE gels. Signal from the entire length of each isoform lane was quantified, normalized to loading control and expressed as percent ataxin-3 signal in the pellet fraction compared to the total signal. Black bars on the right side of longer exposure blots indicate SDS-resistant species whose migration over time shifts to higher molecular weight markers for isoform 1. Shown in histograms are means ± SD. P values from ANOVA with Tukey’s post-hoc. Samples from isoforms 1 and 2 were loaded onto different gels to avoid the need for over-exposure of isoform 1 to visualize and quantify isoform 2. Black arrow: main ataxin-3 band. Red arrow: potentially phosphorylated form of ataxin-3. Blue arrows: ubiquitinated species of ataxin-3. B) Filter-trap assay of flies expressing isoform 1 or 2 in neurons, harvested at the indicated days. C) Independent filter-trap assays from isoforms 1 and 2 were loaded separately to avoid the need for over-exposure of isoform 1 to visualize and quantify isoform 2. Shown in histograms are means ± SD. P values from ANOVA with Tukey’s post-hoc. For panels (B) and (C): Iso: isoform.
We next examined the aggregation of both isoforms through filter-retardation. We loaded samples of each isoform side by side (figure 2.5B) as well as separately (figure 2.5C) to show how much lower isoform 2 protein levels are compared to isoform 1 in this assay, similar to what we observe in western blots. For quantification, however, we loaded preparations from both isoforms on separate filter papers and membranes, which were then processed simultaneously. Over the course of 14 days, we observe a clear and statistically significant increase in the amount of aggregated isoform 1 trapped on the membrane (figure 2.5B, C). We do not observe the same trend with isoform 2 over the same time course. In fact, we notice lower levels of filter-trapped species on days 7 and 14 compared to day 1 (figure 2.5C). Reduced levels of filter-trapped ataxin-3 isoform 2 from day 1 to day 7 could be due to different types of aggregates formed over time by isoform 1 than 2, as hinted at by the SDS-resistant smears in figure 2.5A. Future work, utilizing pure isoforms, is necessary to precisely differentiate the types of oligomeric and higher-order species that each isoform generates. At this point, we can conclude that over the course of two weeks in fly neurons we observe increased aggregation of pathogenic ataxin-3 isoform 1, but not of isoform 2, in two independent assays. Increased aggregation of isoform 1 over this time course could be due to the much lower protein levels of isoform 2—if there is less protein available to aggregate, quality control processes might more easily “handle” it, rendering the protein less toxic.

2.3.4 Isoform 2 at protein levels comparable to isoform 1 is no less toxic in *Drosophila*

Above, we observed that isoform 2 is much less toxic in flies than isoform 1 of pathogenic ataxin-3. This property of isoform 1 coincides with its higher protein levels. If both isoforms are expressed at the same protein level, is isoform 2 still less toxic to *Drosophila*? To address this question, we used fly genetics to express isoform 2 so that its protein levels are similar to isoform 1.
We used a different isoform 1-expressing fly line for these assays. As we showed in figure 2.3A, expression of isoform 1 in all tissues leads to developmental lethality. However, for work that we published before[223,234] we prepared another fly line that expresses isoform 1 of pathogenic ataxin-3 at protein levels lower than the current transgene (this line is denoted as “weaker line” figure 2.6A and supplemental figure S2.4). The ataxin-3-encoding transgene in this line is also inserted on the third chromosome of the fly and is on the same genetic background as the other two isoform lines we described above. Flies expressing the “weaker” version of isoform 1 in all tissues can eclose as adults and survive for a number of days. We generated flies that contain one or two copies each of the driver and isoform 2 and used the sqh-Gal4 driver to examine their relative toxicity in all fly tissues. We selected this driver, because ataxin-3 is a ubiquitously expressed protein and because in the above studies isoform 1 was consistently more potently toxic than isoform 2, regardless of expression pattern. Thus, we proceeded with ubiquitous expression for this set of investigations.

Increased copy number of driver and isoform 2 leads to higher isoform 2 protein levels (figure 2.6B) and increased lethality (figure 2.6C) compared to trans-heterozygous flies, as expected. Based on western blotting, double-homozygous driver;isoform 2 flies express ataxin-3 at protein levels comparable to the weaker isoform 1 (figure 2.6B). Flies that harbor two copies each of driver and transgene are able to emerge as adults and survive for ~25 days. When we compare the longevity of these flies to those expressing the weaker isoform 1 line, we observe comparable survival curves (figure 2.6D; statistical analysis indicates that there is a significant difference in overall longevity of flies expressing isoform 2 compared to the weaker isoform 1-expressing line). Based on these results, isoform 2 is decidedly toxic in the fly when at protein levels similar to isoform 1.
Figure 2.6 Isoform 2 is no less toxic than isoform 1 when at comparable protein levels.
A) Western blots of flies expressing the noted ataxin-3 isoform transgenes in all neurons. Flies were one day old. Shown in histograms are means ± SD. P values from student’s t-tests comparing isoform 1 and isoform 2 to the weaker isoform 1. For these blots we utilized the pan-neuronal Gal4 driver, since use of the ubiquitous one results in late pupal and pharate adult lethality from isoform 1. For a blot showing levels of the three ataxin-3 transgenes when expressed ubiquitously in flies, see Supplemental fig. 4 where we show results from pupal lysates. B) Western blots of one-day-old adult flies with the noted genotypes. Each lane is an independent repeat. Shown in graph are means ± SD. P values from ANOVA with Tukey’s post-hoc. Iso: isoform. C) Longevity assays when flies express isoform 2 in all tissues. sqh-Gal4 was the driver. For perspective, flies without pathogenic ataxin-3 live >70 days (e.g. Fig. 3). P values from log-rank tests. D) Longevity of adult flies with the noted copies and transgenes. Curve from isoform 2 is the same as in panel C, since these assays were conducted at the same time. P value from log-rank test.

2.3.5 Degradation of ataxin-3 by autophagy and the proteasome in Drosophila

Isoforms 1 and 2 mRNA levels are similar in the fly (figure 2.1C), but isoform 2 protein levels are much lower in all tissues tested (figures 2.2, 2.3, 2.4, 2.5, 2.6) suggesting that it is turned
over more rapidly than isoform 1. The two major degradative routes in the cell are the proteasome and autophagy. We used fly genetics to examine the role of each pathway in determining the levels of ataxin-3. We expressed ataxin-3 isoform 1 or 2 by itself or in the presence of RNAi constructs targeting autophagy or proteasome components. For these analyses we limited expression to the eye, because knockdown of these genes more widely is lethal in flies, whereas targeting in fly eyes is tolerated ([252] and other observations from the Todi lab). Our findings from these assays are summarized quantitatively in figure 2.7A. Expression of isoform 2 in fly eyes results in its lower protein levels compared to isoform 1 (figure 2.7B), similar to what we observe in all other fly tissues examined. Where available, we used more than one RNAi line per gene, but in some cases we only had one line available. We reasoned that this approach is not unreasonable to explore the role of autophagy and of the proteasome on isoform protein levels, since we are focusing on a general pathway rather than a specific gene.

For autophagy we targeted Atg7, Atg8a and Atg12, which are critical for the formation and expansion of the autophagosome [257,258]. Atg8, a member of the ubiquitin-like proteins of the LC3 family, becomes conjugated to phosphatidylethanolamine at autophagosome-forming sites, where it plays a role in membrane dynamics and substrate recruitment [257]. Atg12 is constitutively conjugated to Atg5. The Atg12-Atg5 conjugate is essential for Atg8 lipidation and thus essential for autophagosome formation [257,258]. We observe that targeting Atg12 and Atg8a though RNAi leads to a significant increase in protein levels for both isoforms (figure 2.7A and blots 2.7B-G).
Figure 2.7 Knockdown of autophagy or proteasome components leads to increased isoform protein levels in Drosophila.
A) Summary of quantification of the data in panels B–J and other experimental repeats not included. Histograms show means ± SD. P values from student's t-tests. N ≥ 3 for each group. B–J) Western blots of dissected fly heads expressing either isoform 1 or 2 of pathogenic ataxin-3 on RNAi control background or with RNAi constructs targeting the noted genes. Numbers associated with RNAi lines indicate different constructs targeting the same gene. Lack of numbers means that only one RNAi line was at hand for that gene. Black arrow: main ataxin-3 band. Red arrow: potentially phosphorylated form of ataxin-3. Blue arrows: ubiquitinated species of ataxin-3. Black bar in panel (B): SDS-resistant species. In panels D and E, images are from the same membrane, same exposure, cropped and rearranged for viewing. C: controls, containing the driver (GMR-Gal4) and either isoform of ataxin-3 on the genetic background of the RNAi line. Signal was quantified from the entire lane for each well and used in panel (A).
Overall, the effect is not very robust, ranging from 50%-100% increase in protein levels. These findings are limited to the effect from one RNAi line for each gene; results from the other RNAi line do not reach significance, probably due to differential capacity from each RNAi construct to reduce mRNA levels of the intended gene.

Next, we turned to the proteasome. The proteasome comprises a 20S degradative core and one or two 19S regulatory components [259-263]. We focused on Rpn1 (part of the 19S, where it helps with proteasome substrate binding [264]) and the 20S proteolytic core components α5 and β6[259-263]. Targeting of each of these genes by RNAi leads to a statistically significant and robust increase in the protein levels of isoform 2, reaching as high as 400% (figure 2.7A, F, H, I). For isoform 1, targeting of Rpn1 and β6 also leads to higher protein levels (figure 2.7A, C, E). For Rpn1, three different RNAi lines lead to statistically increased protein levels for isoform 2, but only one line leads to higher levels for isoform 1 (figure 2.7A, C, E, F, H, I, J).

Collectively, we interpret these results to suggest that the proteasome and autophagy are both involved in the degradation of isoform 2 of pathogenic ataxin-3. The results are not as significant with isoform 1, but since targeting of at least one gene in each pathway leads to increased protein levels of this isoform as well, we are confident in concluding that isoform 1 is also a target of both pathways, in concurrence with prior reports[53,206,220]. Based on our observations, the proteasome is critically important for ataxin-3 degradation, with contribution from autophagy.

2.3.6 Isoform 2 is turned over more rapidly than isoform 1 in mammalian cells

All of our assays thus far employed an intact organism, the fruit fly. The isoforms we examined are of human origin and thus would benefit from investigations in a mammalian environment. We used the same constructs from fly-based work to examine the levels of both
pathogenic ataxin-3 isoforms in mammalian cells. The objective for this next set of experiments was to determine whether isoform 2 of pathogenic ataxin-3 is degraded more rapidly than isoform 1 and whether the proteasome and autophagy are involved.

As shown in figure 2.8A, when either isoform is transiently transfected in HeLa cells, we observe much lower protein levels of isoform 2 than isoform 1. Inhibiting the proteasome (MG132) markedly increases the levels of isoform 2, approaching the protein levels of isoform 1 (figure 2.8A). Similarly, when we inhibit autophagy (chloroquine), we observe modestly, but statistically significantly increased levels of isoform 2 (figure 2.8B). We do not see a clear upregulation of isoform 1 in the presence of MG132 or chloroquine (figure 2.8A and B). This is likely because of the long half-life of ataxin-3 \([243,265,266]\) requiring prolonged incubation with inhibitors to observe an effect, as shown in supplemental figure S2.5.

We next determined that the lower levels of isoform 2 protein are due to its expedited degradation in cells. We inhibited new protein translation with cycloheximide (CHX). Isoform 1 is rather stable over the course of 6 hours whereas isoform 2 is degraded quite rapidly (figure 2.8C). Considering the strong effect of proteasome inhibition on the levels of isoform 2 (figure 2.8A), we examined whether MG132 halts the degradation of isoform 2. As shown in figure 2.8C, in the presence of MG132 degradation of isoform 2 is essentially halted.
Figure 2.8 Isoform 2 is turned over more rapidly in HeLa cells.
A & B) Western blots from HeLa cells transiently transfected with plasmids encoding either isoform of pathogenic ataxin-3 and treated, or not, with the proteasome inhibitor MG132 (20 μM) or with chloroquine (100 μM) for 6 h before harvesting. Every lane is from an independent repeat. Each well of cells received 1 μg plasmid of the indicated isoform. Arrow in (B): main ataxin-3 band. Histograms are from blots above and other independent repeats. P values from student's t-tests comparing treated lanes to respective untreated lanes of the same construct. C) Western blots of whole cell lysates from HeLa cells transiently transfected as noted then treated with cycloheximide (CHX; 100 μg/ml) for the indicated amounts of time, as well as MG132, where applicable (20 μM). Wells received a higher amount of isoform 2 than isoform 1 (4:1 ratio; isoform 1 plasmid was supplemented with empty vector to bring it to the same total amount as isoform 2) to more closely approximate the amount of isoform 2 protein to isoform 1 at time 0 h. The two sets of blots are from independent experiments. Quantifications are from blots on the top and other, independent repeats. P values from student's t-tests comparing isoform 2 to isoform 1 for the middle graph, and the two right columns to the left-most column in the bottom graph. Shown in graphs are means ± SD.

The experiments above were conducted in HeLa cells. We conducted similar studies in HEK-293T cells in order to recapitulate our findings in another mammalian cell line. Again we
observe lower levels of isoform 2 than isoform 1 (supplemental figure S2.6), mirroring the results we observed in HeLa cells (figure 2.8A-C). Proteasome inhibition has a strong impact on the levels of isoform 2, whereas application of chloroquine does not lead to noticeable changes in its protein levels. As in HeLa cells, we again observe that isoform 1 is not readily impacted by the application of proteasome or autophagy inhibitors in HEK-293T cells. Based on these results, we conclude that isoform 2 is degraded more rapidly by the proteasome than isoform 1 in mammalian cells, with additional, minor contribution from autophagy.

2.3.7 The role of the C-terminus on the stability of isoform 2

The only difference between the two ataxin-3 isoforms is at their C-termini and isoform 2 is degraded decidedly more rapidly than isoform 1. Rapid turnover of isoform 2 could be due to the recognition of its tail end by degradative machineries, with its C-terminus essentially functioning as a degron (amino acid sequences that signal host protein turnover). Degron sequences function alone or in conjunction with internal lysine residues that enable their ubiquitination to target host proteins for degradation, and can be found at either terminus or internally in a protein [267-269]. If the tail of isoform 2 acts as a degron, then its addition to isoform 1 should lead to lower protein levels. It is also possible that the C-terminus of isoform 1 instead stabilizes its host protein; thus, its fusion onto isoform 2 might increase its protein levels.

We fused the C-terminus of isoform 1 onto the end of isoform 2 and vice versa (figure 2.9A). As shown in figure 2.9B, C, and F, addition of the C-terminus of isoform 1 onto isoform 2 leads to higher levels of this protein in a proteasome-dependent manner; autophagy does not seem critical. The levels of the new construct now approach those of isoform 1 (figure 2.9B, C and F), and are much higher than the levels of the original isoform 2. Conversely, fusion of the C-terminus of isoform 2 to the end of isoform 1 leads to significantly lower protein levels (figure 2.9D, E); the
new resulting protein is at levels comparable to those of isoform 2 (figure 2.9F). Proteasome inhibition results in higher levels of isoform 1 with the C-terminus of isoform 2 (figure 2.9D, F); autophagy again does not seem to have a noticeable impact (figure 2.9E, F).

From these assays, it appears that the C-terminus of isoform 2 has a destabilizing effect on its host protein. At the same time, the C-terminus of isoform 1 might also protect ataxin-3 protein from degradation. The current setup does not entirely exclude the possibility that the stabilizing effect of the C-terminus of isoform 1 onto isoform 2 prevents the recognition of the tail-end of...
isoform 2 as a potential degron. It should be noted, however, that all of the ataxin-3 constructs used here are capped by an HA epitope tag at their C-termini (figure 2.9A); if the C-terminus of isoform 2 must be at the very tail-end of ataxin-3 to destabilize it, then the HA tag should have also precluded its recognition for degradation and would have presumably led to protein levels comparable to isoform 1. Thus, it is still quite possible that the C-terminus of isoform 1 has a stabilizing effect on the rest of ataxin-3. Altogether, these data suggest that the C-terminus of ataxin-3 is important for the overall levels of the SCA3 protein, unlocking new paths to consider for the handling of this protein in the cellular environment.

2.4 Discussion

The work presented here sheds new light on the relative pathogenic contribution of the two major ataxin-3 isoforms in SCA3. Two full-length ataxin-3 protein variants arise from alternative splicing of the ATXN3 gene, in which the abnormal expansion of a polyQ-encoding CAG triplet repeat causes SCA3. Each isoform contains the same N-terminal portion and polyQ domain but differs at the C-terminus. The objective of this study was to investigate the possibility of differential toxicity from the two main, full-length isoforms that are expressed in the mammalian brain. Our fly-based investigations show that polyQ-expanded isoform 2 is less toxic than polyQ-expanded isoform 1, as a result of lower protein levels. Isoform 2 is more rapidly degraded, principally by the proteasome, leading to reduced aggregative propensity and toxicity in all fly tissues tested. Thus, polyQ-expanded isoform 2 of pathogenic ataxin-3 is not likely to be a major contributor to SCA3. (We note here that expression of ataxin-3 with a normal polyQ repeat (22-25Q) is not toxic when present throughout the fly or in select tissues (e.g. neuronal, eye or glial) using the same drivers as in this work ([215,216] and additional, unpublished observations by the Todi lab).)
Prior work with isoforms 1 and 2 in vitro and in mammalian cells observed higher aggregation of isoform 2 than isoform 1 [1], which differs from what we observe in flies. The tail of isoform 2 is inherently more hydrophobic than that of isoform 1 [1], most likely accounting for its enhanced aggregation in the earlier study. However, the isoforms might behave differently in an intact organism, leading to higher aggregation of isoform 1 than 2. Another key point that may explain this difference is the fact that isoform 2 is present at much lower protein levels in fly models, thus presenting a reduced challenge for quality control pathways that regulate its folding and aggregation. There may also be cell type- and tissue-specific differences in the aggregative propensity of one isoform over the other in the mammalian brain. Addressing this last point will require the generation of additional tools that can reliably and consistently identify one species over the other in the mammalian brain. Based on our work in fly neurons, we can only conclude that the two isoforms behave differently, with isoform 1 more likely to aggregate than isoform 2.

A recent study investigated ataxin-3 isoform properties in cultured mammalian cells (HEK-293T) by utilizing ATXN3 knockout cells and transiently expressing the ataxin-3 isoforms described here [213]. This other work included yet another variant of ataxin-3 with a premature stop codon in isoform 2 (figure 2.1A), a variant which we did not investigate, choosing instead to focus on the full-length proteins for this manuscript. Both versions of isoform 2 (early-stop and full-length) exist in SCA3 patients. Results from this prior publication do not mirror our observations: it concluded that full-length isoforms 1 and 2 harboring the polyQ expansion are degraded similarly in HEK-293T cells and that autophagy, rather than the proteasome, primarily mediates their turnover [213]. Instead, we observed clear proteasome involvement and rapid degradation rates for isoform 2 compared to 1. Supporting the present observations, an earlier investigation also found that isoform 2 is more rapidly cleared in a proteasome-dependent manner.
in cultured Cos7 cells [1]. How might we resolve these differences? Autophagy likely participates in the degradation of pathogenic ataxin-3 and has been reported before as a key component in ridding the cell of the SCA3 protein [53, 54, 206, 220, 270, 271]. Depending on specific conditions and types of stress, one pathway may predominate over the other in importance for ataxin-3 degradation. With respect to the degradation rates of each isoform, our overall results and prior published work [1] lead us to propose that under most cases isoform 2 is less stable than isoform 1 of pathogenic ataxin-3, regardless of the mechanism of clearance.

Our work raises an intriguing question: what property of isoform 2 renders it more rapidly degraded? We suggest that the C-terminus of isoform 2 might act as a degron, signaling the degradation of the host protein. Whether the tail of isoform 2 is directly recognized by degradative machineries as a signal, or whether its ubiquitination is required for the protein to be eliminated, awaits determination. We have shown before that ataxin-3 does not require its own ubiquitination to be degraded [266]; it may well be that the C-terminus of isoform 2 directly targets the host protein to the proteasome through specific protein-protein interactions. Our studies also present the additional possibility that the C-terminus of isoform 1 stabilizes its host protein. The tail of isoform 1 could protect the rest of the protein from degradation through specific binding partners that preclude the degradation of ataxin-3, similar to what we have observed for ataxin-3 and Rad23 [223, 266]. As argued in work by other scientists [213], the two different C-termini of ataxin-3 have differing binding partners—e.g. ubiquitin can bind the tail of isoform 1 but not isoform 2 due to the lack of UIM3—that may dictate their stability and functions. Additional explorations into the precise nature of isoform 2 degradation may unlock new mechanisms of protein quality control that could be of importance not only to understanding SCA3 but also more generally the principles guiding protein turnover in this and other disease states.
In conclusion, we investigated whether the two main, full-length ataxin-3 isoforms are differentially toxic \textit{in vivo} when harboring an abnormally expanded polyQ repeat. Our results reveal marked differences in isoform toxicity due to differences in the stability of the two variants of the SCA3 protein. Defining the properties and propensities of various products expressed by a disease gene, as done here for SCA3, is necessary to comprehend how that disease arises and how to treat it.

Citation:

CHAPTER 3: UBIQUITIN-INTERACTING MOTIFS OF ATAXIN-3 REGULATE ITS POLYGLUTAMINE TOXICITY THROUGH HSC70-4-DEPENDENT AGGREGATION

3.1 Introduction

Spinocerebellar Ataxia Type 3 (SCA3; also known as Machado-Joseph Disease) is the most frequent dominant ataxia worldwide. SCA3 is caused by CAG repeat expansion in the gene ATXN3 that is normally 12-42 repeats long, but is expanded to ~60-87 repeats in patients [2,206]. This triplet repeat encodes a polyglutamine (polyQ) tract in the protein ataxin-3, a deubiquitinase (DUB; figure 3.1) implicated in protein quality control and DNA repair [206,272]. The precise molecular details of SCA3 pathogenesis remain unclear.

SCA3 is one of nine neurodegenerative disorders caused by abnormal polyQ expansion in diverse proteins [2,206]. The polyQ family includes Huntington’s Disease (HD), SCAs 1, 2, 3, 6, 7, and 17, Dentatorubral–pallidoluysian atrophy, and Kennedy’s Disease. Together, HD and SCA3 are the two most common among the nine known polyQ diseases. Although each polyQ disorder is caused by aberrant lengthening of the same domain, they are clinically distinct diseases [2,3,37,206,273,274]. Fundamental to polyQ disorders is the role of protein context: regions outside of the polyQ that modulate properties of the elongated tract [2,3,37,216,222,223,266,273,274]. It is protein context that differentiates polyQ diseases; why, for example, expansions in ataxin-3 cause SCA3 instead of another disorder. While studies of other polyQ diseases provided insight into the role of protein context in polyQ neurodegeneration [2,3,37,273,274], for SCA3 protein context has not been clarified. For instance, in SCA1 the nuclear localization signal directs its disease protein, ataxin-1, to the nucleus and is necessary for pathogenesis [37,160]. Studies in SCA7 also identified interactions with other proteins that show a role for non-polyQ domains in the toxicity of its disease protein, ataxin-7 [3,194,273]. Another good case of polyQ protein biology and expression pattern
being connected to phenotype is Kennedy's Disease. In this disorder, polyQ expansion occurs near the amino terminus of the androgen receptor (AR) resulting in a partial loss of AR function that leads to androgen insensitivity. AR is expressed in motor neurons of the spinal cord and the brainstem, where it mediates neurotrophic responses. It is these neurons that degenerate in Kennedy's Disease [2,3,130,275,276].

To understand the role of protein context in SCA3, we focused on the domains of ataxin-3. Ataxin-3 contains a ubiquitin-protease (Josephin) domain on its N-terminal half [206,272] (figure 3.1A). Within it are the catalytic triad that enables ataxin-3 to cleave ubiquitin bonds, as well as two ubiquitin-binding sites (UbS) that either interact with ubiquitin (UbS1), or with ubiquitin and the proteasome-associated protein, Rad23 (UbS2) [206]. Downstream are two ubiquitin-interacting motifs (UIMs 1 and 2), followed by a site that binds the AAA ATPase, VCP (VCP-binding motif; VBM), and the polyQ tract [206]. At the end resides a third UIM (UIM3) [206,272]. In another, less common isoform UIM3 is replaced by a hydrophobic C-terminus [1,62,277]. UIMs enable ataxin-3 to bind poly-ubiquitin [212,225,244].

We previously examined the role of three domains of ataxin-3 in its pathogenicity: UbS2, VBM and the catalytic site. UbS2 regulates ataxin-3 turnover and toxicity [216,223,266]; the VBM controls its aggregation and toxicity, but not turnover [216,222,266]; and the catalytic site is necessary for ataxin-3's inherent neuroprotective functions in [215,216,223]. These findings led us to examine the role of other domains of ataxin-3 by using novel, isogenic lines of the model organism, *Drosophila melanogaster*. Here, we show that toxicity from ataxin-3's polyQ tract is markedly impacted by the UIMs. Whereas the isolated polyQ is decidedly pathogenic, UIM addition enhances toxicity. Mutating the UIMs of full-length, pathogenic ataxin-3 renders it less pathogenic. UIMs interact with heat shock protein cognate 4 (Hsc70-4), which enhances ataxin-3
aggregation and toxicity in a manner dependent on these ubiquitin-binding domains. Additional studies indicate that Hsc70-4 enhances the aggregation of other polyQ proteins and provide a comprehensive view of the relative impact of various ataxin-3 domains on its pathogenicity. We introduce unique genetic tools to understand SCA3 biology and to discover and optimize therapeutic options for it. Our studies provide novel insight into protein context in SCA3 and highlight Hsc70-4 as a new target for intervention.

3.2 Materials and Methods

3.2.1 Antibodies

Anti-ataxin-3 (mouse monoclonal 1H9, MAB5360, 1:500-1,000; Millipore), anti-MJD (rabbit polyclonal, 1:15,000 [21]), anti-HA (rabbit monoclonal C29F4, 1:500-1,000; Cell Signaling Technology), anti-V5 (mouse monoclonal R960-25, 1:500-1,000; ThermoFisher), anti-tubulin (mouse monoclonal T5168, 1:10,000; Sigma-Aldrich), anti-lamin (mouse monoclonal ADL84.12-5, 1:1,000; Developmental Studies Hybridoma Bank), anti-HSPA8/Hsc70-4 (rabbit monoclonal D12F2, 1:1,000; Cell Signaling Technology), peroxidase conjugated secondary antibodies (goat anti-mouse, goat anti-rabbit, 1:5,000-10,000; Jackson Immunoresearch).

3.2.2 Drosophila materials and procedures

Flies were housed at 25°C in diurnal environments on conventional cornmeal or RU486-containing media. Common stocks were from Bloomington Drosophila Stock Center: GMR-Gal4 (#8121); isogenic host strain attP2 (#36303); Hsc70-4 RNAi lines #1, 2, and 3 (#28709, #34836, and #35684, respectively), GMR-QF2W (#59283), and ATXN3TR-Q78 (#8141). Gifts included: sqh-Gal4 (Dr. Daniel Kiehart, Duke University), elav-GS-Gal4 (Dr. R. J. Wessells, Wayne State University), elav-Gal4 and repo-Gal4 (Dr. Daniel Eberl, University of Iowa). All flies were heterozygous for driver and transgene.
Ataxin-3 cDNAs were based on sequences from previous publications \[1,2,12,216,222,223,225,265,277\]. These include Ataxin-3(SCA3): -Intact, -Catalytically Inactive, -UbS2 mutated, -VBM mutated; as well as isolated polyQ80, UIM1-2-Q80, Q80-UIM3, UIM1-2-Q80-UIM3, and Ataxin-3(SCA3)-UIMs mutated. Transgenes were sub-cloned into pWalium-\[10.moe\]. Transgenic lines were generated via phiC31 integration into attP2 on chromosome 3 \[174,216,222,223,277\]. An additional ataxin-3 line was generated for figure 3.8 with the same sequence as in figure 3.1A, but with a V5 tag, using plasmid pQUASp and driven by GMR-QF2W. All transgene insertions were validated by PCR and genomic sequencing and Western blotting, using procedures described before \[174,216,222,223,252,277\].

For longevity and motility, adults were collected on day of eclosion. Deaths were recorded daily. Motility was tested weekly. For motility, adults were transferred into fresh vials 1h before assessment. Then, the percentage of flies per vial to reach the top at 5, 15, and 30s was recorded after flies were forced to the bottom.

### 3.2.3 Western blotting and quantification

Unless otherwise specified, 3 or 5 flies (depending on experiment), or 10 dissected adult heads per group were homogenized in boiling lysis buffer (50mM Tris pH 6.8, 2% SDS, 10% glycerol, 100mM dithiothreitol), sonicated, boiled for 10m, and centrifuged at 13,300xg at room temperature for 10m. Western blots were developed using PXi 4 (Syngene), or ChemiDoc (Bio-Rad). Blots were quantified with GeneSys (Syngene), or ImageLab (Bio-Rad), respectively. For direct blue staining, PVDF membranes were submerged for 10m in 0.008% Direct Blue 71 (Sigma-Aldrich) in 40% ethanol and 10% acetic acid, rinsed in 40% ethanol / 10% acetic acid, air dried, and imaged.
3.2.4 Filter-trap assay

Three adult flies, or 10 dissected heads per group were homogenized in 200μl NETN buffer (50mM Tris, pH 7.5, 150mM NaCl, 0.5% Nonidet P-40), supplemented with protease inhibitor cocktail (PI; S-8820, Sigma-Aldrich). Lysates were diluted with 200μl PBS containing 0.5% SDS, sonicated briefly, then centrifuged at 4,500xg for 1m. One hundred μl supernatant was diluted with 400μl PBS. Thirty or 70μl (depending on experimental setup) of each sample was filtered-vacuumed using Bio-Dot (Bio-Rad) through a 0.45μm nitrocellulose membrane (Schleicher & Schuell) that was pre-incubated with 0.1% SDS in PBS. Membrane was rinsed twice with 0.1% SDS/PBS and analyzed by Western blotting.

3.2.5 Soluble/pellet centrifugation

Ten whole flies per group were lysed in 300μl NETN buffer with PI, sonicated, then centrifuged at 20,000xg at 4°C for 30m. Supernatant was quantified with the BCA assay (ThermoFisher). Pellet was resuspended in 200μl of PBS/1% SDS. Thirty μg of supernatant fraction and 7μl of resuspended pellet was each supplemented with 6X SDS, boiled, and loaded for Western blotting.

3.2.6 Nuclear/cytoplasmic separation

Fractionation was performed using the ReadyPrep Protein Extraction Kit (Bio-Rad) using 5 whole flies per group that were lysed in Cytoplasmic Extraction Buffer (Bio-Rad). Three times as much nuclear fraction was loaded onto gels compared to cytoplasmic fraction to eliminate the need for over-exposure.

3.2.7 Co-immunopurification

Fifteen whole flies or 30 dissected fly heads per group were lysed in 600μl NETN+PI, tumbled at 4°C for 30m, then centrifuged for 5m at 10,000xg at 4°C. Supernatant was incubated
with bead-bound antibody for 2-4h. Then, beads were rinsed 5X with NETN+PI. Bead-bound complexes were eluted by boiling in Laemmli buffer.

3.2.8 Mass spectrometry

Forty whole flies per group were homogenized in 1mL NETN+PI, sonicated, centrifuged at 4°C for 5m at 10,000xg and supernatant was transferred into a fresh microfuge tube. Half of each sample was incubated with anti-HA beads (ThermoFisher) for 2h, rinsed 4X with NETN and bead-bound proteins were eluted with Laemmli buffer. The other half was tumbled for 30m with an equal mixture of Tandem Ubiquitin Binding Entities (TUBEs) 1 and 2 (Lifesensors), centrifuged to isolate TUBE beads and supernatant was incubated for 2h with anti-HA beads, rinsed and eluted as above. Eluates were resolved on an SDS-PAGE gel that was stained with Sypro Ruby (ThermoFisher). Bands of interest were excised and examined by LC/MS/MS at Wayne State University Mass Spectrometry Core.

3.2.9 Histological preparation

The proboscises and wings of adult flies were removed before fixing overnight in 2% glutaraldehyde/2% paraformaldehyde in Tris-buffered saline/0.1% Triton X-100. Fixed bodies were dehydrated in a series of 30, 50, 70, and 100% ethanol for 1h each, washed in propylene oxide overnight, embedded in Poly/Bed812 (Polysciences), sectioned at 5μm, and stained with toluene blue.

3.2.10 Statistical analyses

Statistical tests used are specified in figure legends. Log-rank tests with Holm-Bonferroni adjustments, ANOVA, and Kruskal-Wallis tests were conducted in Prism 8 (GraphPad); student's t-tests were conducted in Excel (Microsoft) or Numbers (Apple). P values are noted as reported by the software used. The number of biological replicates is noted on figures and corresponding
legends. We have also noted in legends whether two-tailed or one-tailed tests were used. Two-tailed tests were used where a directionality of change was not necessarily predicted. One-tailed tests were used where we expected a specific directionality in change between control and experimental groups. In figures, non-statistically significant outcomes are noted in italicized font, whereas statistically significant outcomes are in normal font.

3.3 Results

3.3.1 The modulatory role of ataxin-3’s domains on its toxicity

Protein context is a key determinant of polyQ degeneration [2,3,37,130,206,272-275], but its importance in SCA3 has not been investigated systematically. We previously examined the role of non-polyQ domains of ataxin-3, UbS2 [216,223,266] and VBM [222] (figure 3.1A). Prior work from us, spearheaded by pioneering studies from the Bonini lab, also showed that the catalytic site of ataxin-3 is important for toxicity [214-216,223]. However, the contribution of different domains to ataxin-3 pathogenicity has not been investigated side-by-side.

Using isogenic fly lines (figure 3.1B), we examined the toxicity of pathogenic ataxin-3 when expressed pan-neuronally in adults using the binary, Gal4-UAS system of expression [238,278] and a driver that requires the drug RU486 to induce transgenes [250,251,279,280]. We selected this approach since SCA3 is adult-onset and progressive, and because neurons are the type of cell impacted. We compared the following full-length, human ataxin-3 versions: intact domains, catalytically inactive, mutated UbS2 (UbS2* [222,223,266]) and mutated VBM (VBM* [222]; figure 3.1B). The polyQ repeat is 78-80, within patient range [206]. Each transgene is inserted into the same chromosomal location, attP2, as one copy and in the same orientation, leading to similar expression [174,179,222,242,252].
Figure 3.1 Ataxin-3 domains impact its polyQ toxicity.

(A) Diagram of ataxin-3. Josephin: deubiquitinase domain necessary for ataxin-3’s ability to cleave ubiquitin bonds. UbS: ubiquitin-binding site; UIM: ubiquitin-interacting motif; RKRR: amino acid sequence necessary for ataxin-3 to bind directly to VCP, known as VCP-binding motif. QQQ: polyQ that is expanded in SCA3. Also shown is the amino acid sequence of human ataxin-3 used. (B) Summary of transgenic fly lines, mutations, and abbreviations. (C) Longevity results when pathogenic ataxin-3 versions are expressed pan-neuronally in adults. P values: log-rank tests with Holm-Bonferroni adjustment. Italicized P value: not statistically significant. (D) Means +/- SD from negative geotaxis tests. Groups are the same as in (C). P values: one-tailed, Student’s t-tests comparing lines to control flies of the same day and time point. ‘One-tailed’ tests were chosen based on the outcomes from longevity data in panel (C), which indicated specific directionality in expectations. If no P value is noted, it was equal to or greater than 0.05. Exact P values are given in other cases. The total number of flies in each group is shown in panels.
For panel (D), the number of independent groups of flies per genotype was: Ctrl: 5, Intact: 5, Inactive: 6, UbS2*: 5, VBM*: 5. Repeats were biological replicates.

Adult, pan-neuronal expression of pathogenic ataxin-3 with intact domains caused early lethality (figure 3.1C). Catalytically inactive ataxin-3 led to significantly shorter lifespan, as did the version with mutated UbS2. Mutating the VBM led to markedly longer lifespan compared to “intact”. Still, flies expressing pathogenic ataxin-3 with mutated VBM did not reach normal longevity (figure 3.1C). These data were complemented by negative geotaxis assays (figure 3.1D). We observed reduced motility with flies expressing pathogenic ataxin-3 “intact” and further impaired motility with those expressing inactive ataxin-3, or ataxin-3 with mutated UbS2. On the other hand, pan-neuronal expression of pathogenic ataxin-3 with mutated VBM did not significantly impact motility.

Next, we examined the effect of the same ataxin-3 transgenes in glial cells to obtain information on toxicity from pathogenic ataxin-3 in this cell type. Overall trends were similar to neuronal cells: catalytically inactive and UbS2-mutated ataxin-3 led to increased toxicity compared with pathogenic ataxin-3 with intact domains. But, unlike in neurons, in glial cells we did observe a statistical difference in toxicity between pathogenic ataxin-3 with mutated catalytic site and the one with mutated UbS2. There was no statistical difference between controls not expressing pathogenic ataxin-3 and flies expressing pathogenic ataxin-3 with mutated VBM both in terms of longevity (figure 3.1 - figure supplement 3S.1A) and motility (figure 3.1 - figure supplement 3S.1B). These findings highlight a need to explore the role of glia in SCA3.

Based on these collective data, we conclude that domains outside of ataxin-3’s polyQ play significant modulatory roles in pathogenicity. These outcomes led us to ask: what is the role of the most common domain of ataxin-3, the UIM, in SCA3?
3.3.2 UIMs enhance toxicity of the isolated polyQ of ataxin-3

To conduct an initial examination of ataxin-3 UIMs in SCA3, we synthesized transgenes that encode its isolated polyQ by itself or with the sequential addition of UIMs, including a full-length version (figure 3.2A). We generated isogenic lines that integrated each transgene into locus attP2, same as flies in figure 3.1. The polyQ for these new flies is encoded by alternating CAGCAA repeats. Long, pure CAG tracts can lead to mRNA-based toxicity [228] and unconventional translation of non-polyQ frames [239-241]. Our approach mitigates both possibilities [226-233], allowing us to focus solely on toxicity from the polyQ and specific domains. Insofar as we know, ours is the first instance that this approach is undertaken to dissect in detail the role of regions surrounding a polyQ in an intact organism.

The isolated polyQ80 was more toxic than full-length, pathogenic ataxin-3 when expressed in all tissues or only in neurons (figure 3.2B, C). This pattern was reversed in glia; full-length ataxin-3 was more problematic than polyQ80 in these cells, highlighting a need to better understand glial response to polyQ proteins. Intriguingly, polyQ80 was generally less toxic than the same peptide with appended UIMs when expressed everywhere, selectively in neurons, or in glia; addition of UIM3 appeared particularly pathogenetic (figure 3.2B, C). Eye-restricted expression of these transgenes also led to variable toxicity, with UIM3 addition being more problematic than the isolated polyQ80 or polyQ80 with all UIMs (figure 3.2D).

Gal4 drivers above enable transgene expression during development and in adults. Because, as mentioned above, SCA3 is adult-onset, we also investigated the effect of polyQ80 without and with UIMs when expressed pan-neuronally only in adults. PolyQ80 was noticeably less toxic than versions with UIMs (figure 3.2E). Flies expressing Q80-UIM3 or UIM1-2-Q80 lived longer than those expressing polyQ80 with all UIMs (figure 3.2E).
There were tissue-dependent variations in toxicity from the above constructs. For example, Q80-UIM3 was markedly more toxic in fly eyes compared to the polyQ80 alone or with the addition of other UIMs (figure 3.2D), whereas polyQ80 with all three UIMs was the most toxic species in adult neurons (figure 3.2E). These are intriguing finding that require future attention;
also, they are not dissimilar from the neuronal-glial differences we observed in figure 3.1, figure 3.1 - figure supplement 3S.1, and figure 3.2B. These outcomes highlight the utility of the new fly models that we have generated to understand tissue-selective toxicity in vivo in the future. Collectively, data in figure 3.2 indicate a regulatory role for UIMs in the pathogenicity of the expanded polyQ of ataxin-3, and led us to explore them in the full-length protein.

3.3.3 UIMs of full-length ataxin-3 regulate toxicity and aggregation

To investigate a regulatory role for the UIMs in the pathogenicity of full-length ataxin-3, we generated an additional transgenic line that expresses human ataxin-3 with mutated UIMs (figure 3A). This line was created through site-directed mutagenesis of the plasmid used to generate flies in figure 3.2A. The mutations that we selected were previously shown to disable the ability of ataxin-3 UIMs to bind poly-ubiquitin [212,225,244].

Expression of pathogenic ataxin-3 with mutated UIMs was consistently and markedly less toxic than its counterpart with intact UIMs in all of the tissues tested (figure 3.3B, C). Motility assays also showed that adults expressing UIM-mutated ataxin-3 retained mobility longer than adults expressing pathogenic ataxin-3 with intact UIMs (figure 3.3D). We conclude that mutating its UIMs renders ataxin-3 significantly less pathogenic.

Next, we examined if reduced toxicity from UIM mutations reflected changes in ataxin-3 levels or aggregation. Mutating the UIMs does not cause loss of ataxin-3 protein (figure 3.4A). We used three antibodies for this assay: a rabbit monoclonal antibody that recognizes the C-terminal HA tag; a mouse monoclonal antibody that recognizes an epitope upstream of UIM1-2 (1H9); and a rabbit polyclonal antibody generated against full-length ataxin-3 (MJD; [211]). While there is variance in the exact species recognized by each antibody, overall we conclude that mutating the UIMs does not deplete ataxin-3 protein. Differences in species likely stem from
epitopes recognized by each antibody, SDS solubility of different ataxin-3 species, and epitope exposure on PVDF membrane.

Prior work showed that nuclear presence of polyQ disease-causing proteins is a critical regulator of pathogenicity [3,160,221,281]. Based on sub-cellular fractionation, there was no significant difference in distribution between ataxin-3 with intact or mutated UIMs (figure 3.4B). However, according to two different assays, mutating the UIMs rendered pathogenic ataxin-3 less
aggregation-prone. Centrifugation-based protocols and filter-trap assays both showed higher aggregation of pathogenic ataxin-3 with intact UIMs (figure 3.4C, D). PolyQ protein aggregation is a hallmark of this family of diseases [2, 3, 63, 206, 273, 274]. In our Drosophila studies, aggregation of pathogenic ataxin-3 and other polyQ proteins precedes toxicity; also, the level of aggregation mirrors their extent of pathogenicity [174, 179, 215, 216, 222, 223, 277]. Based on data from figures 3.3 and 3.4, we conclude that the UIMs of pathogenic ataxin-3 enhance its aggregation and toxicity in Drosophila.

Figure 3.4 Mutating the UIMs of pathogenic ataxin-3 decreases its aggregation. (A) Western blots from simple lysates of adults expressing transgenes pan-neuronally. Red arrows: likely proteolytic products of ataxin-3 that we observe sometimes. (B) Cytoplasmic/nuclear fractionation of lysates from adults expressing transgenes pan-neuronally. Histograms: quantification of images from the left and other, independent biological replicates. Means +/- SD. P value: two-tailed Student’s t-test. Italicized P value denotes lack of statistical significance. The entire ataxin-3 signal in each lane was used for calculations, from the main band to the top. (C) Soluble/pellet centrifugation of lysates from adults expressing the noted transgenes. Histograms are from the left and other, independent biological repeats. Shown are means +/- SD. P value: two-tailed, Student’s t-test. The entire ataxin-3 signal in each lane was used for calculations, from the main band to the top. (D) Filter-trap assay of lysates from adult flies expressing transgenes pan-neuronally. Each image is from an independent biological repeat. Histograms: quantification of images from left. Means +/- SD. P value: two-tailed, Student’s t-test. Black arrows in panels: main
ataxin-3 band. Blue arrows in panels: ubiquitinated ataxin-3. SDS-soluble and -resistant labels are based on our lab’s prior work with ataxin-3.

### 3.3.4 UIMs mediate ataxin-3 interaction with Hsc70-4

UIMs exist in proteins with various functions and interact with poly-ubiquitin [272,282-285]. The UIMs of ataxin-3 indeed bind poly-ubiquitin [212,225,244]. Still, we wondered whether proteins other than ubiquitin also interact with ataxin-3 through UIMs. After all, at least one other ubiquitin-binding domain on ataxin-3, UbS2, interacts with ubiquitin and a non-ubiquitin protein, Rad23 [223,266,286]. To find proteins that interact with ataxin-3 through its UIMs, we homogenized flies expressing ataxin-3 pan-neuronally and depleted resulting lysates of poly-ubiquitin by using Tandem Ubiquitin-Binding Entities (TUBEs; Methods). Afterwards, we immunopurified (IP-ed) ataxin-3 from the supernatants and resolved all proteins on an SDS-PAGE gel (figure 3.5A). We observed a band above ataxin-3 that was present with intact UIMs, but not as clearly visible when they were mutated. Mass spectrometry from the dissected regions yielded heat shock protein cognate 4 (Hsc70-4) only in the presence of ataxin-3 with intact UIMs (figure 3.5A).

We confirmed mass spectrometry data with co-IPs. Full-length, pathogenic ataxin-3 co-IP-ed endogenous Hsc70-4 in a UIM-dependent manner: whereas pathogenic ataxin-3 with intact UIMs co-IP-ed Hsc70-4, we did not observe the same with UIM-mutated, pathogenic ataxin-3 (figure 3.5B). The interaction of ataxin-3 with Hsc70-4 could potentially rely more heavily on UIM3. We inferred this from IPs with isolated polyQ peptides. PolyQ80-UIM3 co-IP-ed endogenous Hsc70-4 more readily than polyQ80 alone or polyQ80 with UIMs 1 and 2 (figure 3.5C). In panel 5D we conducted co-IPs with modified conditions to approximate more the levels of the HA-tagged polyQ forms. All three polyQ80 versions co-IP-ed Hsc70-4, but UIM3 appeared
more important than the two other UIMs in facilitating an interaction. Collectively, these results highlight the UIMs of ataxin-3 as an interaction site for Hsc70-4 \textit{in vivo}.

3.3.5 \textbf{Hsc70-4 enhances aggregation and toxicity of pathogenic ataxin-3}

Hsc70-4 belongs to the heat shock protein 70 superfamily, whose closest human orthologue is the constitutively expressed, heat shock protein family A member 8 (HSPA8 [287]). Alongside other HSP70 members, its primary role is that of an ATP-dependent chaperone for unfolded proteins in protein quality control [288-297]. HSPA8 is a part of the ubiquitin-proteasome degradation system and is also involved in chaperone-mediated autophagy [293,294,298]. Chaperones are presumed to alleviate toxicity from polyQ proteins by refolding or degrading them. Our studies above, however, counterintuitively hint at the possibility that Hsc70-4 might do the opposite with ataxin-3, enhancing its polyQ toxicity.

Hsc70-4 is a different member of the HSP70 superfamily than fly Hsc70Cb and human HSPA1L, both of which were previously reported to suppress polyQ-dependent degeneration [299-302]. Also, there is recent evidence that HSP70 members have pleiotropic effects on misfolded proteins [303]. Therefore, we knocked down Hsc70-4 in fly eyes to examine the possibility that it exacerbates SCA3. We selected fly eyes because widespread targeting of Hsc70-4 in the fly was developmentally lethal (larval, pupal and pharate adult lethality), whereas expression in fly eyes was not problematic (clearly present pseudopupil and good overall morphology). We utilized three independent RNAi lines for Hsc70-4. Expression of each line in fly eyes led to reduced Hsc70-4 levels (figure 3.6A). Levels of Hsc70-4 in fly eyes were likely even lower than what was captured by blots; we expressed RNAi in eyes but conducted blotting using whole heads that also contained tissues where Hsc70-4 was not targeted.
Figure 3.5 Pathogenic ataxin-3 interacts with Hsc70-4 in a UIM-dependent manner.

(A) SDS-PAGE gel stained for all proteins with Sypro Ruby. Boxed areas were excised and proteins were identified with LC/MS/MS. Two versions of immunopurifications (IPs) were conducted. Middle panel shows results from IP-ed ataxin-3 from whole fly lysates. Right panel shows results from IP-ed ataxin-3 after lysates were first incubated with tandem ubiquitin binding entities (TUBEs; see the main text and Materials and methods). (B–D) Western blots from anti-HA IPs from flies. Results are representative of independent biological replicates conducted three times. NS: non-specific.

We next determined the effect of knocking down Hsc70-4 on the phenotype caused by pathogenic ataxin-3 by scoring fly eyes, as exemplified in figure 3.6B; a higher number denotes a worse phenotype. Hsc70-4 knockdown consistently improved eye phenotype, denoted by the persistence of the pseudopupil – an indicator of underlying structure integrity – and fewer cases with depigmentation (figure 3.6C). We also observed biochemical changes in ataxin-3 when
Hsc70-4 was knocked down. Both SDS-resistant ataxin-3 species and aggregates isolated by filter-trap assays were reduced in the presence of Hsc70-4 RNAi (figure 3.6D, E). Since aggregation of ataxin-3 is a key determinant of its toxicity in flies [216,222,223,277], suppression of eye phenotypes by Hsc70-4 knockdown is likely a consequence of reduced SCA3 protein aggregation. Lastly, we tested the effect of Hsc70-4 knockdown on eye phenotypes from ataxin-3 with mutated UIMs. Knocking down Hsc70-4 did not have a detectable impact on eye phenotype caused by pathogenic ataxin-3(UIM*) (figure 3.6F). Collectively, these data highlight Hsc70-4 as an enhancer of pathogenic ataxin-3 toxicity in a manner dependent on UIMs.

Figure 3.6 Hsc70-4 knockdown improves ataxin-3 toxicity in fly eyes.
(A) Western blots from dissected fly heads. Histograms are from blots on top and other, independent biological repeats. Shown are means +/- SD. Statistics: one-way ANOVA with Dunnett’s correction. Lysates were from crosses used for studies in panels (B), (C) and other, similar experimental setups. (B) Representative images and summary of the scoring system designed to evaluate the effect of Hsc70-4 knockdown on the phenotype caused by pathogenic ataxin-3 in fly eyes. 1: normal-looking eyes that have a...
clearly defined pseudopupil and good overall appearance; 2: eyes where the pseudopupil is not as clearly pronounced compared with normal eyes and with some minor unevenness/irregularity in peripheral bristle arrangement; 3: eyes where the pseudopupil is not visible; 4: eyes with depigmentation. Quantitative outcomes are shown in panels (C) intact, pathogenic ataxin-3 and (F) UIM-mutated, pathogenic ataxin-3. (C) Statistics: Kruskal-Wallis tests comparing Hsc70-4 RNAi to Ctrl from independent biological replicates when pathogenic ataxin-3 is expressed in fly eyes. Shown are means +/- SEM. (D) Western blots from dissected fly heads. Quantifications are from panels above and independent biological repeats. Shown are means +/- SD. P values: one-way ANOVA with Dunnett's correction. Black arrow: unmodified ataxin head. Qu.

3.3.6 Hsc70-4 knockdown suppresses polyQ toxicity more generally

We were intrigued by the suppressive effect of Hsc70-4 knockdown on pathogenic ataxin-3. Because we also observed Hsc70-4 co-IP with the isolated polyQ (figure 3.5C, D) we tested whether its knockdown impacts polyQ80 eye phenotypes. Expression of polyQ80 led to pseudopupil loss and depigmentation. Knockdown of Hsc70-4 improved depigmentation, even though it did not recover the pseudopupil, indicative of ameliorated pathology (figure 3.7A). Importantly, improved eye phenotype was accompanied by reduced polyQ80 aggregates (figure 3.7B). (We note here that we have not observed a single, non-aggregated band of polyQ80 by Western blots; we only observe aggregated species.)

We confirmed a protective effect from Hsc70-4 knockdown in yet another polyQ model [304]. This model contains polyQ78 and the C-terminus of an ataxin-3 isoform that lacks UIM3; UIMs 1 and 2 are also absent. Because expression of this polyQ species does not always yield a clear external eye phenotype, we examined internal structures. Normally, the ommatidia (eye units) are organized in a fan-like pattern, which was absent in the presence of polyQ78 (ATXN3TR-Q78; figure 3.7C). When Hsc70-4 was knocked down, the ommatidia were visible in a clear sign of toxicity suppression. Altogether, these results suggest the possibility of a wider role for Hsc70-4 as an enhancer of toxicity from different polyQ proteins.
3.3.7 UIMs are important for inter-ataxin-3 interactions

So far, we observed that Hsc70-4 affects aggregation and toxicity of pathogenic ataxin-3 and that it interacts with the SCA3 protein through UIMs (figures 3.5, 3.6). Since our data also implicated Hsc70-4 in the toxicity of an isolated polyQ that does not contain UIMs (figure 3.7) we wondered whether the UIMs of ataxin-3 have additional properties that might impact SCA3 protein aggregation.

Ataxin-3 proteins bind each other [243,305-307]. Therefore, we investigated whether the UIMs play a part in inter-ataxin-3 associations. We examined whether HA-tagged, pathogenic ataxin-3 with intact or mutated UIMs co-IPs V5-tagged pathogenic ataxin-3 with intact UIMs in Drosophila. We found that the UIMs were important, but not necessary, for these interactions; a
reduced amount of full-length, V5-tagged pathogenic ataxin-3 co-IP-ed with UIM* ataxin-3 compared to ataxin-3 with intact UIMs (figure 3.8A). Additional co-IPS using truncated polyQ fragments and ataxin-3 with a wild-type polyQ indicated that UIM3 is more important for this interaction than UIMs 1 and 2 (figure 3.8B) and suggested that the UIMs are involved not only in the interaction of disease-causing ataxin-3, but also in the unexpanded version. It is not unexpected that the UIMs are not essential for ataxin-3 interactions; in vitro experiments showed that the N-terminal, Josephin domain facilitates inter-ataxin-3 associations [305-307]. Thus, the UIMs of ataxin-3 enable inter-ataxin-3 interactions, which might additionally contribute to aggregation and toxicity in SCA3.

Figure 3.8 UIMs are important for inter-ataxin-3 interactions. (A, B) Western blots from IPs of ataxin-3 with expanded (A) or wild-type (B) polyQ repeats expressed in fly eyes. All transgenes (HA- or V5-tagged) were expressed at the same time and in the same tissues. Gray arrow: V5-positive signal absent in IP lanes.

3.4 Discussion

In this study, we systematically explored the role of protein context in SCA3 in an intact organism. We found that the UIMs of ataxin-3 are key players in pathogenicity through an unexpected role from Hsc70-4, which exacerbates SCA3. We propose that Hsc70-4 increases the toxicity of pathogenic ataxin-3 by interacting with its UIMs and worsening polyQ aggregation (figure 3.9A). Additionally, since the UIMs facilitate inter-ataxin-3 binding, misfolded ataxin-3 proteins may be brought into close proximity that can further increase their chances of interaction, aggregation and toxicity (figure 3.9A); as mentioned earlier, in our Drosophila studies of SCA3, ataxin-3 aggregation is consistently linked to its extent of pathogenicity. Our work provides the first, comprehensive view of the role of non-polyQ domains on ataxin-3 toxicity (figure 3.9); it furthermore highlights an understudied and impactful property of a member of the normally beneficial HSP70 superfamily.

We previously reported on the role of three ataxin-3 domains on its polyQ toxicity: the catalytic domain, UbS2 and the VBM (figure 3.9B). Based on our work in Drosophila, the catalytic domain is important for ataxin-3’s ability to induce production of DnaJ-1/Hsp40, whose upregulation suppresses polyQ toxicity [216,223]. The capacity of ataxin-3 to upregulate DnaJ-1 also requires its binding to the proteasome-associated protein, Rad23 at UbS2 [216,223]. Through upregulated DnaJ-1, ataxin-3 suppresses its own polyQ-dependent toxicity [215,216,223]. Therefore, it is not surprising that mutating either the catalytic site or UbS2 on pathogenic ataxin-3 renders the protein more toxic. The general suppressive contribution of each domain to ataxin-3 polyQ toxicity, as measured by longevity, is similar.
Unlike the catalytic site and UbS2, the other domains investigated here – VBM and UIMs – act as enhancers of ataxin-3 polyQ toxicity. Elegant work by the Wanker and Bonini labs identified a region immediately preceding the polyQ of ataxin-3 as the binding site for VCP ([308]; figure 3.9B). The stoichiometry of this interaction was suggested to be 4 ataxin-3 proteins for 1 VCP hexamer [308]. These results and our findings that VCP binding enhances ataxin-3 aggregation and toxicity [222] led us to propose that the interaction of multiple ataxin-3 proteins with a single VCP hexamer increases their chances of interaction and accelerates aggregation ([222]; figure 3.9B). The other domains that enhance the toxicity of ataxin-3 are the UIMs. Their relative impact on ataxin-3 polyQ toxicity might be stronger than the VBM (figure 3.9 - figure supplement 3S.4). Our results strongly implicate Hsc70-4 as a driving force behind the role of the UIMs as enhancers of ataxin-3 pathogenicity.

Figure 3.9 Proposed model.
(A) Ataxin-3 UIMs enable its interaction with Hsc70-4 and other ataxin-3 proteins, both of which enhance aggregation and toxicity. UIM3 (thicker lines) appears to be a stronger contributor to these interactions than the other UIMs (thinner lines). (B) Summary of results from the current study and our prior publications on the effect of ataxin-3 domains on its polyQ toxicity. Ub: ubiquitin.
It is intriguing that knockdown of Hsc70-4 reduces toxicity from pathogenic ataxin-3. As mentioned above, heat shock proteins generally work to improve toxicity from aggregated proteins [296,299,309-314]. Hsc70-4 may perform in a context-, protein- and disease-dependent manner. For example, studies with its mammalian orthologue, HSPA8, indicate that it can suppress aggregation and toxicity of the HD-causing protein, huntingtin, harboring an expanded polyQ [315-317]. However, investigations with the disease-causing protein, Tau, suggest this chaperone as a potential enhancer of tauopathies. HSPA8 interacts with Tau and slows down its degradation; also, HSPA8 inhibition can reduce Tau levels in brain [318-320]. These data present the possibility of HSPA8 activities that can lead to enhanced toxicity from some proteins, similar to what we observed with ataxin-3.

How might Hsc70-4 enhance toxicity from pathogenic ataxin-3? Perhaps, ATP hydrolysis by Hsc70-4 is directly inhibited by binding to the UIMs of pathogenic ataxin-3, abrogating its refolding functions. It may also be that under normal circumstances ataxin-3 forms a functional complex with Hsc70-4, collaborating in protein quality control. Just like Hsc70-4/HSPA8, wild-type ataxin-3 is implicated in proteostasis [206,272]. During normal conditions, a complex comprising wild-type ataxin-3, Hsc70-4 and other proteins could triage various substrates. When the polyQ of ataxin-3 is expanded, in a UIM-dependent manner it may cause Hsc70-4 to adopt a dominant negative role and to enhance pathogenic ataxin-3 aggregation. Our study provides a roadmap towards testing these possibilities and uncovering others.

The model above focuses on the role of the UIMs facilitating an interaction between ataxin-3 and Hsc70-4, because pathogenic ataxin-3 with mutated UIMs does not co-IP Hsc70-4. But, we also observed that the isolated polyQ80 co-IP-ed some Hsc70-4 and that the chaperone’s knockdown reduced its toxicity. Two earlier studies also reported suppression of truncated polyQ
toxicity by the knockdown of Hsc70-4 in flies [300,313]; as far as we know, no additional findings have supplemented these data since their original publication. Binding of Hsc70-4 to an isolated polyQ is not surprising. PolyQ peptides aggregate and heat shock proteins interact with misfolded proteins [309,310,314]. However, it is compelling that reducing levels of Hsc70-4 improves, instead of worsening, truncated polyQ toxicity. Other proteins involved with refolding, such as the HSP40 members Dnaj and mrj, and the HSP70 members Hsc70Cb and HSPA1L reduce polyQ protein aggregation and toxicity [179,216,223,299,300,312,321]. But, these proteins are not the same as Hsc70-4/HSPA8, which may function differently depending on protein identity and cellular circumstances. In fact, HSPA8 is not the only member of the HSP70 superfamily to enhance the aggregation of a misfolded disease protein. Recent work in cultured mammalian cells indicated that HSPA1L – previously reported as a polyQ suppressor [299,301,302] – can in fact increase the aggregation of a non-polyQ misfolded protein, SOD1; the mechanism behind this effect is not entirely clear, but it appears to involve HSPA1L binding partners [303]. Consequently, it is not outside of reasonable biological expectations that not all HSP70 members, such as Hsc70-4, function similarly; they may trigger different outcomes on misfolded proteins depending on their binding partners, the identity of the misfolded protein, the cellular environment and physiological conditions.

How might Hsc70-4 worsen polyQ fragment toxicity? The interaction of Hsc70-4 with severely truncated polyQ peptides could inhibit its ability to refold them, especially if additional proteins or binding partners are necessary to steer Hsc70-4 towards refolding, as suggested by studies of other HSP70 members [303]. This point pertains to isolated polyQ fragments. In the case of full-length ataxin-3, the UIMs’ interaction with Hsc70-4 appears stronger and likely supersedes binding through the polyQ, leading to the models of aggregation postulated above. We
draw this conclusion based on our results that mutating the UIMs of full-length, pathogenic ataxin-3 seems to abrogate its interaction with endogenous Hsc70-4. Additionally, knockdown of Hsc70-4 does not significantly change eye phenotype scores when the UIMs of full-length ataxin-3 are mutated. Nevertheless, we cannot fully discount the possibility that Hsc70-4 activities on the SCA3 protein may also depend on polyQ-based interactions, once UIM-based binding is established. At this point, the primary takeaway is that Hsc70-4 can enhance polyQ toxicity, in the case of ataxin-3 through its UIMs.

In summary, we demonstrated that non-polyQ domains of disease-causing ataxin-3 are key regulators of its toxicity, and that the UIMs are important determinants of the SCA3 protein's pathogenicity. Our studies establish a clear role for protein context in SCA3 and, through Hsc70-4, provide a unique entry point into further examinations and potential therapeutic solutions for this incurable ataxia. Lastly, our findings that Hsc70-4 can enhance polyQ toxicity have broader implications for the general understanding of chaperone biology.

Citation:

CHAPTER 4: TARGETING THE VCP-BINDING MOTIF OF ATAXIN-3 IMPROVES PHENOTYPES IN DROSOPHILA MODELS OF SPINOCEREBELLAR ATAXIA TYPE 3

4.1 Introduction

Spinocerebellar Ataxia Type 3 (SCA3, also known as Machado-Joseph Disease) is the most frequent, dominantly inherited ataxia in the world. Along with Huntington’s disease, it is the predominant member of the family of polyglutamine (polyQ) neurodegenerative disorders that also includes SCAs 1, 2, 6, 7, and 17, Dentatorubral-pallidoluysian atrophy, and Kennedy’s Disease [51-53,67,322-324]. SCA3 is progressive, adult-onset and leads to neurodegeneration in cerebellar pathways, pontine and dentate nuclei, substantia nigra, globus pallidus, cranial motor nerve nuclei, anterior horn cells and peripheral nerves [51,53,69,70,325-340]. The underlying genetic defect in SCA3, like the other polyQ disorders, is the abnormal expansion of a CAG trinucleotide repeat that, once translated, results in polyQ protein aggregation and toxicity [51-53,62,341]. In SCA3, CAG expansion occurs in the ATXN3 gene [51-53,62,341], which encodes the 42 kD protein, ataxin-3 (figure 4.1A). The aberrant lengthening of the polyQ domain is a primary cause of all polyQ diseases, but each one is clinically distinct [3,37,206,243,273,274]. This indicates that determinants of toxicity in these diseases are not only the polyQ repeats, but also the domains and interactions surrounding the expanded repeat, referred to as ‘protein context’.

A body of work in polyQ diseases has highlighted a role for various non-polyQ regions and interactions in polyQ degeneration [2,4,37,209,220,273,342,343]. For ataxin-3, there are several domains that contribute to protein context and pathogenesis. As a deubiquitinating enzyme (DUB), the N-terminal half of ataxin-3 contains a ubiquitin-protease (Josephin) domain housing the catalytic triad that enables the protein to cleave isopeptide bonds [206,212,272]. The Josephin domain also contains two ubiquitin binding sites (UbS) that interact with ubiquitin (UbS1 and
UbS2), or the proteasome-associated protein Rad23 (UbS2)\[206,272,286,344\]. Downstream of the Josephin domain are two ubiquitin-interacting motifs (UIMs 1 and 2), a site that binds the AAA ATPase known as valosin-containing protein (VCP, or p97; this site is termed the VCP-binding motif or ‘VBM’), and the polyQ domain \[206,212,243,272\]. The C-terminus of ataxin-3 following the polyQ domain commonly contains an additional UIM, although an isoform also exists that does not contain the third UIM \[206,272\].

We recently showed that the VBM is a significant contributor to ataxin-3 pathogenicity \[222,223\]. The VBM comprises the arginine-rich sequence ‘RKRR’ which, when mutated into the amino acid sequence ‘HNHH’, no longer binds VCP \[308\]. VCP is a ubiquitous, homo-hexameric AAA ATPase that is bound directly at its N-terminus by ataxin-3 through the VBM \[308,345\] (figure 4.1A). VCP regulates the proteasomal degradation of various proteins \[345-348\]; however, mutations that disable binding of ataxin-3 to VCP do not impact ataxin-3 protein levels \[216,222,223,266\] or its subcellular distribution \[222\].

Our previous work exploring the role of the ataxin-3-VCP interaction utilized *Drosophila melanogaster* models of SCA3 expressing full-length, human ataxin-3 with and without the mutation that prevents VCP-binding. Mutating the pathogenic ataxin-3 VBM improved fly motility and longevity compared to counterparts expressing pathogenic ataxin-3 Q80 with an intact VBM \[222\]. Biochemically, mutating ataxin-3’s VBM or reducing the levels of VCP through RNA-interference did not reduce SCA3 protein levels, but decreased pathogenic ataxin-3 aggregation \[222,343\] – a critical observation given that, in our assays with pathogenic ataxin-3, its aggregation precedes toxicity \[222,223,277,343\]. These and other findings led us to conclude that pathogenic ataxin-3 with a mutated VBM is less aggregation-prone and less toxic. They also described an important role for VCP in determining ataxin-3 toxicity, suggesting
A model where multiple ataxin-3 proteins bind individual VCPs in a single hexamer [222]. These interactions may bring pathogenic ataxin-3 proteins into closer proximity, increasing their chances of interaction and aggregation – i.e., VCP seeds the aggregation of ataxin-3, exacerbating its toxicity.

These prior findings prompted the question that we tackle in this study: might we reduce the toxicity of pathogenic ataxin-3 by disrupting its interaction with VCP through a ‘decoy’ approach targeting the VBM? We utilized a truncated protein that consists of the VCP N-terminus (amino acids 1-199), denoted at ‘N-VCP’, that binds ataxin-3 [308], but cannot hexamerize [349-352] (figure 4.1A). As VCP is a critical protein, whereas ataxin-3 is dispensable in mice [206,217-220], we reasoned that it would be more practical to target ataxin-3 rather than VCP so as not to impede VCP functions while still hindering the interaction of interest. As shown in the results, we find that N-VCP ameliorates the toxicity of pathogenic ataxin-3 at all stages and in all tissues tested in the fruit fly. The protective effect of N-VCP is specific to pathogenic ataxin-3 and is dependent on its VBM. Also, the beneficial effect of the N-VCP truncated protein is enhanced by increasing its levels. This work expands on our understanding of this critical protein-protein interaction in SCA3 and highlights the disruption of ataxin-3’s interaction with VCP as a potential therapeutic strategy for this neurodegenerative disease.

4.2 Materials and Methods

4.2.1 Antibodies

Anti-Ataxin-3: mouse monoclonal 1H9, 1:500-1000; Millipore, rabbit polyclonal, 1:15000 [63]. Anti-HA: rabbit monoclonal, 1:500-1000; Cell Signaling Technology. Anti-Myc: mouse monoclonal 9E10, 1:500-1000; Santa Cruz Biotechnology. Anti-VCP: rabbit monoclonal, 1:500-
1:1000; Cell Signaling Technology. Peroxidase-conjugated secondary antibodies: goat anti-mouse and goat anti-rabbit, 1:5000; Jackson ImmunoResearch.

4.2.2 Construct Design

Human ataxin-3 cDNAs were based on sequences from previous publications and include either 77 or 80Q [1,2,12,216,222,223,225,265,267,286,343,344,353]. We used the company Genscript (genscript.com) to synthesize the VCP N-terminus cDNA as well as full-length fly VCP. An N-terminal Myc tag-encoding sequence was added immediately preceding the VCP start codon. Transgenes were sub-cloned into pWалиum-10.moe. Transgenic fly lines were generated via phiC31 integration into either attP2 (ataxin-3) on chromosome 3 or attP40 (N-VCP, VCP) on chromosome 2 [174,216,222,223,242,277,343,353-355]. Each insertion was confirmed and validated by PCR, genomic sequencing, and Western blotting, using procedures described in previous work [174,216,222,223,252,277,343].

4.2.3 Drosophila husbandry

Common fly lines were obtained from the Bloomington Drosophila Stock Center: GMR-Gal4 (#8121), the isogenic host strain attP2 (#36303), and the isogenic host strain attP40 (#36304). FlyORF was used to obtain an additional fly VCP line (TER94, #F001765) used in supplemental figure 4.4. The following stocks were gifts: sqh-Gal4 (Dr. Daniel P. Kiehart, Duke University), elav-Gal4-GS (Dr. R. J. Wessells, Wayne State University), elav-Gal4 and repo-Gal4 (Dr. Daniel F. Eberl, University of Iowa). Unless otherwise stated in the figure legends and text, all flies were heterozygous for driver and transgene(s). Whenever ataxin-3 and VCP or N-VCP were co-expressed, their expression was driven by the same Gal4 driver, in the same tissues, and at the same developmental and/or adult time points. All crosses were conducted at 25°C in diurnal
environments with 12h light/dark cycles and on conventional cornmeal or RU486-containing media, and all resulting offspring were maintained in the same conditions [251,277,343].

4.2.4 *Drosophila* examinations

Longevity experiments were performed with adults collected on the day of eclosion and deaths were recorded daily. Flies were observed from the embryo stage through eclosion and deaths at each developmental stage were recorded daily, as summarized in figures. Fly eye scoring was conducted using numerical scales where a higher score indicates a worse phenotype. Breakdowns of the phenotypes for each scale are shown in respective figures and legends. Motility tests were conducted using the negative geotaxis assay, where fifteen flies per vial were tapped to the bottom of the vial and the total number of flies reaching the top was recorded at 5, 15 and 30s and expressed as percent of total flies. Flies that reached the top were scored only once.

4.2.5 Western blotting and quantification

Depending on the experiment, 3 or 5 flies, or 10 dissected adult heads (unless otherwise specified) per group were homogenized in boiling fly lysis buffer (50mM Tris pH 6.8, 2% SDS, 10% glycerol, 100mM dithiothreitol), sonicated, boiled for 10 min, and then centrifuged at 13300xg at room temperature for 10 min. Western blots were developed with either PXi 4 (Syngene) or ChemiDoc (Bio-Rad) and quantified with GeneSys (Syngene) or ImageLab (Bio-Rad), respectively. To conduct direct blue stains, the PVDF membranes were submerged for 10 min in 0.008% Direct Blue 71 (Sigma-Aldrich) in 40% ethanol and 10% acetic acid, rinsed in 40% ethanol/10% acetic acid, air dried, and imaged.

4.2.6 Filter-trap assay

For each group, 10 dissected fly heads were homogenized in 200μL NETN buffer (50mM Tris, pH 7.5, 150mM NaCl, 0.5% Nonidet P-40) supplemented with protease inhibitor cocktail
(PI; S-8820, Sigma-Aldrich). Following homogenization, lysates were diluted with 200μL PBS containing 0.5% SDS, sonicated briefly, then centrifuged at 4500xg for 1 minute at room temperature. 100μL of supernatant was then diluted further with 400μL PBS. Depending on the experimental setup, either 20 or 30μL of each sample were filter-vacuumed using a Bio-Dot apparatus (Bio-Rad) through a 0.45μm nitrocellulose membrane (Schleicher and Schuell) that was pre-incubated and rinsed with 0.1% SDS in PBS. Following filter-vacuuming of the samples, the membrane was rinsed twice with 0.1% SDS in PBS, incubated with primary and secondary antibodies, and analyzed by Western blotting.

4.2.7 Co-immunopurifications

40-80 dissected fly heads or 20 whole flies, depending on the experiment, were lysed in NETN/PBS (50%/50%)+PI, tumbled at 4°C for 30 min, then centrifuged at 4°C for 5 min at 5000xg. The supernatant was incubated with bead bound antibody (either Myc- or HA-tagged; Fisher Scientific) for 2 hours. Beads were then rinsed 1X-3X with NETN/PBS+PI (depending on experiment) followed by elution of bead-bound complexes by boiling in 2% SDS buffer with 100μM DTT.

4.2.8 Statistical Analyses

Statistical tests used are noted in the figure legends. Prism 8 (GraphPad) was used for log-rank tests with Holm-Bonferroni adjustments, Repeated-measures and ordinary one-way ANOVA with Tukey’s multiple comparisons test, Mann-Whitney tests, and Kruskal-Wallis tests. Additional data collection and organization, as well as student’s t-tests, were performed in Excel (Microsoft), or Numbers (Apple). P-values were calculated by the software used for analyses and indicated in the corresponding figures and legends, alongside the number of biological replicates.
4.3 Results

4.3.1 N-VCP is non-toxic to Drosophila and interacts with pathogenic ataxin-3

The N-terminus of VCP is sufficient and necessary to bind ataxin-3 at its VBM in vitro [308].

The N-VCP truncated protein that we utilize here consists of amino acids 1-199 of VCP (figure 4.1A). An N-terminal Myc-tag was added to N-VCP and expression in Drosophila was confirmed...
with a ubiquitous Gal4 driver (spaghetti squash, sqh, figure 4.1B; additional examples are in figures 4.3, 4.4, 4.6, 4.8). Here, as in previous studies, we take advantage of *Drosophila* genetics and the Gal4-UAS binary system of expression [174,179,216,222,223,252,266,277,343,356]. Each Gal4 driver expresses ataxin-3 and N-VCP in a specific tissue either individually or together. In longevity assays, N-VCP expression is non-toxic whether expressed ubiquitously (sqh-Gal4) or pan-neuronally (elav-Gal4; figure 4.1C,D). This was also validated in motility assays, where the pan-neuronal expression of N-VCP did not impact fly motility compared to controls that contained the driver but did not express N-VCP (Supplemental figure 4S.1). Co-immunopurification assays (co-IPs) conducted in flies pan-neuronally co-expressing N-VCP and full-length, human pathogenic (80Q) ataxin-3 confirmed that N-VCP interacts with ataxin-3 (figure 4.1E), setting the stage for an investigation into N-VCP’s impact on pathogenic ataxin-3 toxicity in the fly.

### 4.3.2 Presence of N-VCP is ameliorative in SCA3 models of different toxicity

The *Drosophila* lines modeling SCA3 were designed in two stages. While both express full-length, human ataxin-3 protein with patient range repeats of 77-80Q, one of the lines is designated as ‘stronger’ (Q80) and the other as ‘weaker’ (Q77), based on their comparative toxicity to flies. The difference in lethality is due to the ‘stronger’ line containing an ataxin-3 transgene with an optimized Kozak sequence that results in higher ataxin-3 protein levels [277]. However, it is important to note that beyond the differences in toxicity, the ‘stronger’ and ‘weaker’ transgenes are inserted into the same chromosomal site in *Drosophila* (attP2 on chromosome 3) and are on the same genetic background (w1118; figure 4.2A). Expression was tested in various tissues to obtain a comprehensive understanding of the impact of N-VCP on ataxin-3-related toxicity: ubiquitous expression because ataxin-3 is expressed everywhere; pan-neuronal expression because
SCA3 is primarily a neuronal disease; adult-only, pan-neuronal expression because SCA3 is adult-onset; and expression in fly eyes to observe more changes across phenotypes over time.

<table>
<thead>
<tr>
<th>Fly Line</th>
<th>Description</th>
<th>Expression</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>No pathogenic ATXN3</td>
<td>Genetic background used to generate transgenic flies</td>
<td>Not applicable</td>
<td>Normal development and adult fly longevity</td>
</tr>
<tr>
<td>ATXN3-Weaker (Q77)</td>
<td>Pathogenic ataxin-3 inserted at chromosomal site attP2. Expresses at lower protein levels than “stronger” line</td>
<td>Ubiquitous</td>
<td>Marked developmental death; a handful of adult flies eclosed and were all dead by 40 days; figure 2</td>
</tr>
<tr>
<td>ATXN3-Stronger (Q80)</td>
<td>Pathogenic ataxin-3 inserted at chromosomal site attP2; optimized Kozak sequence. Expresses at higher protein levels than “weaker” line</td>
<td>Ubiquitous</td>
<td>Marked developmental lethality. No adult flies eclosed; figure 3</td>
</tr>
</tbody>
</table>

Figure 4.2 N-VCP reduces toxicity from pathogenic ataxin-3 with 77Q.

(A) Summary of the pathogenic ataxin-3 and control fly lines along with outcomes of their ubiquitous expression in the absence of N-VCP. Driver: sqh-Gal4. (B) Observations of developmental lethality in flies ubiquitously expressing ‘weaker’ pathogenic ataxin-3 (Q77) with or without co-expression of N-VCP. The fly life cycle is outlined above the data and color coordinated with each developmental stage shown in the graph. Means +/- SD. p-values comparing the differences in death at each developmental stage between groups with or without N-VCP were calculated with student’s t-tests and are shown below the graph. (C) Survival analyses from ubiquitously expressed transgenes in flies throughout development and adulthood. p-values: log-rank tests. Black font p-value: comparison between “Ctrl” and “Atxn3(Q77) + N-VCP” groups. Red font p-value: comparison between “Atxn3(Q77) + N-VCP” and “Atxn3(Q77) without N-VCP” groups. (D) Survival analyses from pan-neuronally expressed transgenes in flies throughout development and adulthood. p-value: log-rank test. (E) Survival analyses from pan-neuronally expressed transgenes in
adult flies only. We utilized an RU486-dependent elav-GS-Gal4 driver to initiate expression of ataxin-3 and N-VCP transgenes in adults. Flies developed and eclosed in media without RU486 and were introduced to media with RU486 on the day of eclosion and were maintained on media with RU486 for their entire lives. p-value: log-rank test.

We began our investigations by examining the impact of N-VCP expression on the 'weaker' SCA3 line. As summarized in figure 4.2B, ubiquitous co-expression of pathogenic ataxin-3 and N-VCP led to reduced lethality at various developmental stages: a significantly smaller proportion of developing flies died in pupal, pharate and eclosing stages in the presence of N-VCP than in its absence when pathogenic ataxin-3 was co-expressed in all tissues. Concomitantly, more adult flies successfully eclosed from the pupal case and were tracked for longevity in the presence of N-VCP than in its absence (figure 4.2C).

Perhaps counterintuitively at first, most 'weaker' SCA3 adults died earlier with co-expression of N-VCP (figure 4.2C and supplemental figure 4S.2). We believe this is due to the fact that only a small portion of SCA3 flies eclose as adults in the absence of N-VCP and that they represent the strongest among that population of developing flies. While only the presumed healthiest flies expressing the ‘weaker’ ataxin-3 alone emerged as adults – ~11% of the total developing flies – ~70% of flies co-expressing ataxin-3 and N-VCP emerged as adults. Among the latter, the longest-surviving adults lived markedly longer than flies without N-VCP (Supplemental figure 4S.2), nearly as long as controls that do not express pathogenic ataxin-3 (figure 4.2C). Expression of the ‘weaker’ SCA3 line pan-neuronally at all times (figure 4.2D) or only in adults (figure 4.2E) also resulted in less severe toxicity in the presence of N-VCP. Collectively, these findings indicate that N-VCP suppresses lethality in the ‘weaker’ SCA3 line.

Encouraged by the reduction in lethality in the presence of N-VCP with the ‘weaker’ SCA3 model, we next tested if these results persist in the ‘stronger’ line. Similar to what we observed with the ‘weaker’ model, ubiquitous N-VCP co-expression alongside the ‘stronger’ SCA3
transgene significantly decreased the proportion of developmental deaths and was critical in allowing pupae to enter the pharate adult stage (figure 4.3A). Among flies that ubiquitously expressed pathogenic ataxin-3 Q80 alone, over 95% died as pupae, with only a small portion of the remaining developing flies advancing to the pharate adult stage (figure 4.3A). Conversely, flies co-expressing N-VCP and ataxin-3 Q80 commonly reached the pharate adult stage prior to death (~79%), and a small percentage initiated the eclosing process from the pupal case, but died before emerging successfully (figure 4.3A).

Figure 4.3. N-VCP reduces toxicity from pathogenic ataxin-3 with 80Q.
(A) Observations of developmental lethality in flies ubiquitously expressing the ‘stronger’ pathogenic ataxin-3 (Q80) with or without N-VCP. Means ±/− SD. p-values below graph: student’s t-tests. (B) Survival analyses from adult-only pan-neuronally expressed transgenes. p-value: log-rank test. (C) Western blots of lysates from whole flies pan-neuronally expressing the ‘stronger’ pathogenic ataxin-3 alone or with Myc-tagged N-VCP. Black arrow: main, unmodified ataxin-3 band. Blue arrows: ubiquitinated species of ataxin-3. (D) Co-immunopurification using whole fly lysates expressing the noted transgenes.

To allow for more longitudinal studies in adult flies and to mimic the adult-onset characteristic of SCA3, an inducible pan-neuronal driver was once again employed to express the ‘stronger’ transgene line with and without N-VCP in adult flies only, with expression starting on
day 1 after emergence from the pupal case. N-VCP co-expression significantly increased fly longevity (figure 4.3B), mirroring what we observed with the ‘weaker’ SCA3 model (figure 4.2E).

As will be detailed further below, reduced toxicity from N-VCP was not due to lack of pathogenic ataxin-3 in the presence of N-VCP. Pathogenic ataxin-3 is still abundantly present when co-expressed with N-VCP (figure 4.3C). Additionally, co-expression of N-VCP alongside pathogenic ataxin-3 leads to reduced levels of endogenous VCP co-precipitating with the SCA3 protein (figure 4.3D; please see figure 4.8C for additional data and quantitative information).

We also performed motility studies in flies with the inducible pan-neuronal driver (Supplemental figure 4S.3). Although early in life flies expressing the ‘stronger’ pathogenic ataxin-3 with N-VCP showed a slight improvement compared to those expressing ataxin-3 alone, they did not reach the motility level of control flies that did not express pathogenic ataxin-3; also, the improvement was no longer apparent by the third week (Supplemental figure 4S.3). Collectively these results indicate that the presence of N-VCP is effective in reducing SCA3 toxicity in *Drosophila*.

As an additional counterpoint to the protective role from N-VCP that we observed so far, we tested whether increased expression of full-length VCP has the opposite effect, as predicted by our prior work [222]. We generated flies that contain either one or two genetic copies of fly VCP and compared lethalities when co-expressed with either the ‘weaker’ or ‘stronger’ SCA3 model. As shown in supplemental figure 4.4, VCP over-expression led to worse developmental outcomes that were exacerbated by increased genetic copies of VCP in both SCA3 models, providing additional support for our model that VCP impacts ataxin-3 toxicity.
4.3.3 The effects of N-VCP are specific to ataxin-3 with an intact VBM

We have, so far, observed a protective effect from N-VCP on ataxin-3-dependent toxicity (figures 4.1-4.3). Our model posits that this effect centers on the VBM of ataxin-3. To ensure that indeed the VBM of ataxin-3 is required for the protective role of N-VCP, we examined its ability to impact toxicity caused by pathogenic ataxin-3 with a mutated VBM, which does not interact with VCP in vitro, in mammalian cells or in flies [222]. This VBM-mutated line of SCA3 is isogenic to the SCA3 ‘weaker’ line introduced above [222] and was expressed in flies with or without co-expression of N-VCP.

In this modified ataxin-3 line, the VBM sequence ‘RKRR’ was mutated into ‘HNHH’. Flies expressing pathogenic ataxin-3 with mutated VBM show reduced degenerative phenotypes and this version of ataxin-3 is less aggregation-prone than the SCA3 protein with a functional VBM [222]. Through co-IPs, we confirmed that N-VCP requires a functional VBM on pathogenic ataxin-3 to interact with it specifically (Figure 4.4A). Survival analysis from these flies showed that mutating the VBM of ataxin-3 eliminates the toxicity-mitigating effect of N-VCP pan-neuronally. In fact, N-VCP exacerbates lethality when it is co-expressed ubiquitously with pathogenic ataxin-3 that it can no longer bind (figure 4.4B and 4.4C).

To further clarify that the benefits of N-VCP co-expression are specific to ataxin-3, we conducted experiments examining the potential impact of N-VCP on a pathogenic protein linked to a different polyQ disorder, SCA6. α1ACT is a polyQ-containing transcription factor that is generated from the bicistronic mRNA of the CACNA1A gene, which causes SCA6 [3,176]. We selected this protein because it is not reported to have any interaction with VCP, a statement that we would not be as confident in making with proteins related to other polyQ diseases, like SCA7 and SCA1 [357-359]. The interaction of VCP with other polyQ disease proteins, including SCA1
and SCA7, has been confirmed by experimental evidence, as indicated in previous work and the STRING database [357,359,360].

Figure 4. The ameliorative effects of N-VCP are specific to ataxin-3 and require its intact VBM.
(A) Co-immunopurification assays of whole fly lysates expressing one of two forms of ataxin-3: pathogenic ataxin-3 Q77 or pathogenic ataxin-3 Q80 with a mutated VBM (`RKRR' to `HNHH') without or with N-VCP. (B) Survival analyses of flies pan-neuronally expressing VBM-mutated, pathogenic ataxin-3 with or without N-VCP. Statistics: log-rank test. (C) Survival analyses of flies ubiquitously expressing pathogenic, VBM-mutated ataxin-3 with or without VCP, p-value: log-rank test. (D) Summary of transgenes and outcomes of their expression. (E) Survival analyses of flies pan-neuronally expressing α1ACT(Q70) with or without VCP. Statistics: log-rank test. “ns”: non-statistically significant.

We expressed human α1ACT with a pathogenic 70Q repeat in flies with and without N-VCP, either ubiquitously or pan-neuronally, throughout development and in adults. Ubiquitously expressed α1ACT(Q70) by itself was toxic: most developing flies died as pharate adults, and the handful that emerged as adults died by day 16 (figure 4.4D). In this SCA6 model, ubiquitous co-expression of N-VCP did not improve the phenotype: all flies died as pharate adults, and no adults emerged (figure 4.4D). Pan-neuronal expression of α1ACT(Q70) was less toxic than ubiquitous expression (figure 4.4E), as also shown before [174,179]. Co-expression of N-VCP did not provide a detectable benefit compared to flies expressing α1ACT(Q70) alone (figure 4.4D, E). We
conclude that the reduced toxicity of pathogenic ataxin-3 observed in the presence of N-VCP is specific to ataxin-3 with a functional VBM.

4.3.4 N-VCP alleviates SCA3 phenotype in fly eyes

To expand on our observations, we shifted our focus to a SCA3 model specific to the fly eye. We previously utilized similar models to observe the phenotypic deterioration that occurs with pathogenic ataxin-3 expression, as well as to perform screens of various molecules and genetic modifications that may ameliorate SCA3 [343,361]. Observation of the eye allows for the detection of more subtle changes among groups of flies that might go undetected in other expression patterns. We thought it reasonable to use the fly eye to further examine the impact of N-VCP in flies that co-express human, pathogenic ataxin-3. In models expressing the ‘stronger’ ataxin-3 in fly eyes (GMR-Gal4), the eyes appear normal (wild-type) at eclosion and worsen continually over time [343].

We developed the following scoring system to track the worsening of eye phenotypes for this study: score 1) normal (wild-type-looking) eye; score 2) loss of the pseudopupil; score 3) early signs of color variegation throughout the eye in addition to pseudopupil loss; score 4) depigmentation of a portion of the eye in addition to color variegation and pseudopupil loss (figure 4.5A). Eyes were scored weekly for six weeks following eclosion using the scale described above, with a higher score indicating a worse phenotype. As shown in figure 4.5B and 4.5C, the presence of N-VCP significantly improved eye phenotypes beginning at day 14 and continuing for the remainder of the study. Figure 4.5B shows the average score for each group at each time point, while figure 4.5C expands those averages to display the proportion of each score in each group at a given time point to provide further visualization of the worsening phenotype over time. Thus, just as with other tissues, N-VCP suppresses toxicity from pathogenic ataxin-3 in fly eyes.
Figure 4.5 N-VCP improves pathogenic ataxin-3 toxicity in fly eyes.

(A) Scoring scale and representative images. Distinguishing features from each score category are highlighted. 1) Normal (wild-type-looking) eye; 2) Loss of the pseudopupil; 3) Early signs of color variegation in addition to pseudopupil loss; 4) Depigmentation of a portion of the eye in addition to color variegation and pseudopupil loss. (B) Average eye score in each group at each time point. Statistics: Mann-Whitney tests comparing eye scores at each time point between those expressing pathogenic ataxin-3 alone or with co-expression of N-VCP. Shown are means −/+ SEM. p-value: **** = 0.0001. “ns”: non-statistically significant. (C) Representation of the dispersion of eye scores at each time point for each group from (B).

A strong molecular indicator of toxicity and disease progression in polyQ disease models is the level of disease protein aggregation [2,3,63,206,274]. In our studies using Drosophila SCA models, we found that aggregation of pathogenic ataxin-3 and other polyQ proteins precedes toxicity [174,179,215,216,222,223,277,343]. We also observed that the level of aggregation of these
proteins correlates with the extent of their pathogenicity [174,179,215,216,222,223,277,343]. Next, we sought to examine this relationship in fly eyes.

To assess ataxin-3 aggregation, we performed Western blot analyses using lysates from dissected fly heads expressing the 'stronger' pathogenic ataxin-3 alone or alongside Myc-tagged N-VCP (figure 4.6A). The presence of N-VCP seemed to increase the amount of SDS-soluble ataxin-3 (figure 4.6A, green outlines). Lanes with lysates from flies expressing pathogenic ataxin-3 alone have lower amounts of SDS-soluble ataxin-3 in comparison to those also expressing N-VCP (figure 4.6A, green outlines); additionally, the level of SDS-resistant ataxin-3 appears higher in the absence of N-VCP (figure 4.6A, orange outlines), indicative of increased levels of aggregated species that migrate more slowly through SDS-PAGE gels. Since these blots suggested that the presence of N-VCP reduces aggregated species of pathogenic ataxin-3, we sought to confirm these results through the utilization of filter-trap assays. These assays take advantage of a
porous nitrocellulose membrane that, when a sample is passed through it via suction, captures higher-order aggregated protein species [213,255,256,277,343]. Lysates from flies expressing pathogenic ataxin-3 alone or with N-VCP showed that the presence of the truncated protein significantly reduced the amount of aggregated ataxin-3 trapped on the membrane (figure 4.6B). From these data we conclude that N-VCP leads to reduced levels of aggregated, pathogenic ataxin-3.

4.3.5 Dose-dependent effects of N-VCP

Thus far, we observed improvement in developmental progression, longevity, eye phenotypes, and pathogenic ataxin-3 aggregation of SCA3 models because of co-expression of the N-terminal VCP truncated protein (figures 4.1-4.6). We also observed a significant worsening of developmental outcomes with overexpression of full-length VCP in those same SCA3 models (Supplemental figure 4S.4). As we already confirmed improvement from a single copy of N-VCP, and an increased copy number of full-length VCP intensified SCA3 toxicity (Supplemental figure 4S.4), we next asked the question: can an additional copy of N-VCP further ameliorate SCA3 toxicity in flies?

We returned to the fly eye as our model with *Drosophila* containing either zero, one, or two copies of the N-VCP transgene alongside the ‘stronger’ pathogenic ataxin-3. Just as in figure 4.5, eyes were scored on a weekly basis for six weeks following eclosion; however, the scoring scale for this set of studies was expanded to account for what we thought might be more subtle differences among flies expressing one versus two copies of N-VCP. This new scale is as follows: score 1) normal (wild-type-looking) eye with a clear pseudopupil; score 2) weaker pseudopupil that has begun to fade and lose its clear shape; score 3) undetectable pseudopupil; score 4) color variegation at the edge of the eye in addition to pseudopupil loss; score 5) widespread
depigmentation throughout the eye in addition to color variegation and pseudopupil loss (figure 4.7A). Multiple statistical comparisons were made among the groups to detect differences among the varying copy numbers of the N-VCP transgene.

**Figure 4.7 Dose-dependent N-VCP alleviation of pathogenic ataxin-3 toxicity in fly eyes.**

(A) Expanded scale of representative images for eye scoring system. Distinguishing features from each score category are highlighted. Scoring system is as follows: 1) Normal (wild-type-looking) eye with a clearly defined pseudopupil; 2) Weaker pseudopupil that has begun to fade; 3) Undetectable pseudopupil; 4) Color variegation at the edge of the eye in addition to pseudopupil loss; 5) Widespread depigmentation in addition to color variegation and pseudopupil loss. (B) Average eye score for each group at each time point. Shown are means +/- SEM. Statistics: Kruskal-Wallis and Dunn's multiple comparison tests comparing all three groups. p-value: **** = 0.0001. “ns”: non-statistically significant. (C) Representation of the distribution of eye scores for each group at each time point.
Beginning on day 7, we observed significant improvement in eye phenotype in the presence of both one (validating the observations of figure 4.5) and two copies of N-VCP compared to expression of the ‘stronger’ pathogenic ataxin-3 alone (figure 4.7B). Additionally, starting on day 14, the data revealed a significant improvement comparing one copy to two copies of N-VCP (figure 4.7B). Just as in figure 4.5, figure 4.7B shows the average score for each group at each time point, while figure 4.7C expands those averages to display the proportion of each score in each group at a given time point. The differences among the three groups persisted for the remainder of the observation timeline with a single copy of N-VCP improving eye phenotypes over ataxin-3 alone, and two copies of N-VCP showing improvement over both of the other groups. This indicates that the response from N-VCP on ataxin-3-dependent phenotypes is dose-dependent.

4.3.6 N-VCP reduces pathogenic ataxin-3 aggregation in a dose-dependent manner

As with a single copy of the N-VCP transgene, we again assessed whether phenotypic improvement observed with two copies of N-VCP coincides with further reduction in ataxin-3 aggregation. Analyses of Western blots from dissected fly heads from each line showed that there is a significant increase in the amount of N-VCP from one copy of the transgene to two, and that this increase in N-VCP does not result in significant reduction of total ataxin-3 protein levels (figure 4.8A).

We then focused on SDS-soluble and -resistant ataxin-3. As outlined in figure 4.8A, the amount of SDS-soluble ataxin-3 appears to increase in the presence of N-VCP (figure 4.8A, green outlines) coincident with a decrease in SDS-resistant species (figure 4.8A, orange outlines). This apparent decrease in aggregated ataxin-3 was validated with filter-trap assays, which revealed a steady decrease in higher-order ataxin-3 species from zero to one and zero to two copies of the N-VCP transgene (figure 4.8B). There was also a trend in reduced levels of filter-trapped pathogenic
ataxin-3 when comparing one vs. two copies of N-VCP (figure 4.8B). This trend did not reach statistical significance with ANOVA (P=0.086) when comparing one copy vs. two copies, but did reach significance with a student's t-test (P=0.012), overall supporting the notion that higher levels of N-VCP lead to increased reduction of higher-order ataxin-3 species.

Lastly, we examined whether N-VCP impacts ataxin-3 binding to endogenous VCP. Through co-IPs for HA-tagged ataxin-3 using dissected fly heads, we observed that N-VCP significantly reduced the amount of endogenous VCP that co-IPed with pathogenic ataxin-3 compared to ataxin-3 alone (figure 4.8C). We conclude that N-VCP reduces the interaction of ataxin-3 with endogenous VCP and leads to lower levels of aggregated, pathogenic ataxin-3.
4.4 Discussion

Domains outside of the polyQ repeat play important roles in the pathogenicity of polyQ disease proteins. For the SCA3 protein, ataxin-3, we and others reported that the polyQ-adjacent VBM is important for its interaction with the hexameric protein, VCP [222,308,362-364]. We also found that VCP exacerbates the toxicity of human, full-length, pathogenic ataxin-3 in fly models of SCA3 [222]. Based on those data [222]; on studies by others that investigated the interaction of ataxin-3 with VCP [308]; and on studies that investigated ataxin-3 aggregation in vitro — a two-step process where the N-termini of ataxin-3 proteins interact and facilitate subsequent polyQ length-dependent fibrilization [222,305-307] — a model emerged where VCP hexamers binding to multiple pathogenic ataxin-3 proteins raise their local concentration and likelihood of interacting and aggregating (figure 4.9, top). This model was supported by our prior data that inhibiting the binding of pathogenic ataxin-3 to VCP, or knocking down endogenous VCP, decreases pathogenic ataxin-3 aggregation and toxicity in Drosophila [222]. The VBM of ataxin-3 was reported to contain a nuclear-localization signal [365]; however, based on our earlier research, the VBM does not impact ataxin-3’s sub-cellular localization in Drosophila [222]. Thus, it is unlikely that interventions involving the VBM impact the sub-cellular localization of pathogenic ataxin-3. Collectively, these studies identified the VBM of pathogenic ataxin-3 as a potential targeting site to combat SCA3.

Here, we presented evidence that engaging the VBM of pathogenic ataxin-3 with a truncated protein — the N-terminus of VCP lacking the domains necessary to homo-hexamerise — ameliorates ataxin-3-dependent toxicity in flies. Phenotypic improvement coincided with reduced aggregation of pathogenic ataxin-3 and diminished binding of the SCA3 protein to VCP. Our observations from this work further inform the working model in figure 4.9. We propose that VCP-
dependent aggregation of pathogenic ataxin-3 can be decelerated through the introduction of the N-VCP truncated protein that displaces the VCP-ataxin-3 interaction (figure 4.9, bottom). These results strengthen the rationale to pursue the VBM of ataxin-3 as a potential target for SCA3 intervention.

Ataxin-3 and VCP have been proposed to cooperate functionally in endoplasmic reticulum-associated protein degradation (ERAD) [362-364,366]. The retro-translocation to the cytosol of substrates produced in the ER lumen is a critical step in this branch of proteasome-associated protein quality control. Ataxin-3’s DUB activity has been implicated in regulating the flow of
ERAD substrates in a VCP binding-dependent manner [362-364,366]. Other protein partners of VCP possess a similar VBM and may compete with ataxin-3 to bind VCP, further regulating the retro-translocation process [366]. These studies suggested that one of ataxin-3’s normal functions is to modulate ERAD by altering VCP’s ability to facilitate protein retro-translocation [362-364,366]. However, the physiological significance of the ataxin-3-VCP interaction in vivo is not entirely clear, since Atxn3 knockout mice appear normal [217-219], suggesting either that ataxin-3 is not fully or always required for ERAD-dependent processes, or that another DUB is able to conduct its functions when it is absent. Further uncovering the functional nature of ataxin-3’s interaction with VCP, and whether that interaction is a part of a unique action, is important in understanding ataxin-3’s physiological functions.

The RXXR sequence (where ‘X’ denotes any amino acid) within a predicted α-helix that comprises the VBM is not exclusive to ataxin-3. The ubiquitin ligase E4b (Ube4b), ubiquitin fusion degradation protein 2a (Ufd2a), hydroxymethylglutaryl reductase degradation protein (Hrd1) (also known as Synoviolin 1), and the ER-resident ubiquitin ligase M, 78 000 glycoprotein (gp78) all possess a version of this sequence that can bind directly to the N-terminus of VCP [366]; Drosophila has well conserved orthologues of these genes and their peptides also contain the RXXR sequence [367]. Each of these proteins is involved in ERAD [366]. Their function can be influenced by the mutual exclusivity of their interaction with VCP in direct competition with other VCP-interacting proteins [366]. While these VBM-containing proteins may be involved in similar pathways, their distinct interactions, competitions, and variability in cofactor binding allows each of them to have unique roles in ERAD [366]. Because of these interactions, one might have considered that the N-VCP approach we described here could have led to deleterious effects. While N-VCP suppressed toxicity from pathogenic ataxin-3, it enhanced the toxicity of pathogenic
ataxin-3 with mutated VBM and of the SCA6 protein (figure 4.4). These observations suggest the possibility that N-VCP disrupts some cellular processes during misfolded protein stress, perhaps by interacting with other VBM-containing proteins; the truncated protein, however, does not appear to be detrimental under normal conditions in the fly (figure 4.1 and supplemental figure 4S.1). Since the overarching future aim of this investigation is to inhibit direct binding of ataxin-3 to VCP through a highly targeted, small molecule design, that approach would presumably circumvent the above possibility.

The ultimate goal for studying SCA3 is to devise therapeutic interventions for it. As the SCA3 therapeutics field forges ahead, interventions focused on biomarkers; mechanistic targets that may upregulate specific pathways, such as autophagy; oligonucleotide-based targeting of the ATXN3 mRNA; and others are being considered to mitigate ataxin-3-based toxicity [53,216,223,234,277,343,368-371]. To this list we add the ataxin-3-VCP interaction and the VBM as another potential therapeutic entry point. N-VCP is too large to be a deliverable; however, the basic concept that we presented here can be utilized to design or discover compounds that disrupt the interaction of ataxin-3 with VCP. Since there is benefit from the approach we described here, studies that further evaluate this disruption at a structural level and weaponize it against SCA3 will likely prove beneficial.

To conclude, targeting the VBM of pathogenic ataxin-3 brings phenotypic benefits in fly models of SCA3 and provides further evidence of the importance of protein-protein interactions in the etiology of SCA3 and of other, similar diseases.
Citation:

5 CHAPTER 5: *DROSOPHILA* AS A MODEL OF UNCONVENTIONAL TRANSLATION IN SPINOCEREBELLAR ATAXIA TYPE 3

5.1 Introduction

Long, tandem repeating nucleotide sequences and resulting amino acid tracts are associated with various human genetic disorders[239,372,373]. When expanded, tandem repeats can detrimentally influence protein folding, function, and aggregation resulting in repeat-expansion disorders[239,372,373]. In addition to anomalies stemming from changes at the protein level, expanded nucleotide repeats are also implicated in altering gene structure and causing RNA-mediated toxic effects[372,373]. One mechanism of RNA-based toxicity stems from repeat-associated, non-AUG-initiated (RAN) translation[232,239-241].

RAN translation allows for translation initiation and elongation through a repeat strand in the absence of an AUG initiation codon and in all three reading frames of a specific transcript[232,241]. This process enables the production of multiple homopolymeric or multi-amino acid repeat-containing proteins[232,241]. RAN translation was discovered by pioneering work in Spinocerebellar Ataxia Type 8 (SCA8) after the mutation of the ATG initiation codon in the CAG-repeat containing, ATXN8 gene did not block expression of the polyglutamine (polyQ) repeat in transfected cells[239]. Upon further investigation, homopolymeric polyQ, polyalanine (polyA), and polyserine (polyS) proteins were found to be expressed in all three reading frames in the absence of an AUG or near-cognate codons[239]. This type of translation not only opens the door to the possibility of translation of long tandem repeats into potentially toxic fragments, but also the translation of repeats in alternate reading frames[232,241].

Although RAN translation was initially described in the context of CAG-repeat expansions in SCA8, it has since been shown to occur with expansions of CAG, CUG, GGGCC, and CGG...
repeats associated with several repeat-expansion diseases[232,239-241,374-380]. These repeats drive RAN translation from within 5’ untranslated regions (UTRs), introns, and protein-coding, open reading frames[232,239-241,374-380]. The exact mechanism through which RAN translation occurs remains to be clarified; however, as evidence of RAN translation expands among various diseases several common themes have emerged. First, RAN translation is repeat-length dependent, with the likelihood of RAN-derived protein accumulation increasing with longer repeat tracts[232,241]. Second, long repeat sequences form secondary structures, including hairpin loops, which appear to be a requirement for RAN translation to take place – non-hairpin-forming repeats (such as those with CAA in place of CAG repeats) do not produce RAN proteins[232,241]. Third, the secondary structure of repeat-expanded RNA downstream of the start codon may recruit initiation factors and ribosomal subunits to internal ribosomal entry sites (IRESs) regardless of the presence of initiation or near-cognate initiation codons[232,241].

There are various repeat-containing disease proteins where RAN translation has yet to be identified as a factor in pathogenesis. One such example is another CAG-repeat expansion disease, SCA3, also known as Machado-Joseph Disease. SCA3 is the most prevalent dominant ataxia worldwide and, along with Huntington’s disease, is among the most common of the family of nine polyQ repeat-expansion diseases[2,51-53,67,206,322-324]. It is caused by the expansion of the CAG repeat of the ATXN3 gene that encodes an expanded polyQ tract in ataxin-3, a deubiquitinase implicated in protein quality control and DNA repair[51-53,62,341]. Normally, these repeats exist in a range of 12-42, but expand to ~60-87 repeats in patients[2,206]. This long CAG repeat, similar to that of diseases like SCA8 and Huntington’s disease where RAN proteins have already been detected, makes a SCA3 an intriguing candidate for another potential instance of RAN translation.
While RAN translation is a possible detrimental factor in long-repeat expansions, mRNA toxicity was reported as a potential contributor in polyQ diseases, and more specifically SCA3 [228,229,231,233,381]. The contribution of mRNA-based toxicity in SCA3 was tested by altering the CAG repeat sequence of ATXN3 to have an alternating CAGCAA repeat, which significantly reduced toxicity in Drosophila melanogaster[228]. Expression of an untranslated version of the expanded CAG repeat also resulted in neuronal degeneration[228]. Additional studies suggested that the expanded ATXN3 gene is prone to frameshift mutations[230,382-385] and that these frameshifts result in the production of harmful polyA-containing proteins that have been observed in SCA3 patient lymphoblasts and pontine neurons[382]. These studies were reported before the discovery of RAN translation; their conclusions may have been unrecognized evidence of the presence and influence of RAN proteins. The long CAG repeat, reliance on homomeric CAG repeat secondary hairpin structure for toxicity, and the presence of polyA proteins in SCA3 patient tissue point to the potential of RAN translation in SCA3.

In this study, we set out to search for evidence of the occurrence of RAN translation in SCA3 in vivo. We took advantage of the genetic versatility of Drosophila melanogaster models of SCA3 to express versions of human ataxin-3 transgenes with the native homomeric CAG or interrupted alternating CAGCAA repeats that are, respectively, more and less likely to form the secondary structure necessary to enable RAN translation. Although we were able to identify the expression of polyA proteins when these constructs were expressed in cultured mammalian cells, we found no evidence of RAN proteins in any of our fly models of SCA3. We also returned to the idea of mRNA-based toxicity by expressing ATXN3 constructs that cannot be translated and found that the presence of a homomeric CAG repeat was, albeit to a small degree, more toxic than the interrupted repeat. Overall, this study points to a version of SCA3 pathogenesis that experiences
little to no input from RAN translation and apparent but limited contributions from mRNA toxicity in Drosophila, and underscores limitations with the Drosophila model system in studying RAN translation in this specific disease.

5.2 Materials and Methods

5.2.1 Construct Design

The ATXN3-Q80(CAGCAA) cDNA was based on the human ATXN3 sequences used in previous publications[212,216,222,223,225,265,277,286,343,344,353,386]. ATXN3-Q80(CAG) cDNA and the Met-Null cDNA counterparts to both ATXN3-Q80 constructs were designed from the ATXN3-Q80(CAGCAA) sequence and synthesized using the company Genscript (genscript.com). The transgenes were subcloned into pWalium-10.moe and transgenic fly lines were generated with the phiC31 integrase system into attP2 on chromosome 3 at Duke University Model Systems[174,216,222,223,242,277,343,353-355,386]. All insertions were confirmed and validated with PCR, genomic sequencing and then Western blotting using procedures described in this paper and previous work[174,216,222,223,252,277,343,386].

5.2.2 Antibodies


5.2.3 Mammalian cells and assays

Ataxin-3 constructs containing a pure CAG repeat or an interrupted CAGCAA repeat were sub-cloned into pcDNA 3.1(+1) for expression in a mammalian cellular environment. M-17 or HEK-293T cells were purchased from ATCC, confirmed for lack of mycoplasma contamination.
and cultured under conventional conditions at 37°C and 5% CO2 in DMEM supplemented with 10% FBS and 5% penicillin-streptomycin (Thermo Fisher). Cells were transfected as indicated in figure 5.1 using Lipofectamine LTX (Thermo Fisher). 24 h later, cells were harvested for protein extraction and Western blotting. Cells were scraped with boiling cell lysis buffer (50 mM Tris pH 6.8, 2% SDS, 10% glycerol, 100 mM dithiothreitol), sonicated, boiled for 10 min, centrifuged at top speed at room temperature for 10 min and loaded onto SDS-PAGE gels (4%–20% gradient gels, Bio-Rad).

5.2.4 Drosophila stocks and husbandry

The control fly line, isogenic host strain attP2 (#36303), and the MYC positive control line, a MYC-tagged ataxin-3 (#33610), were obtained from the Bloomington Drosophila Stock Center. Stocks obtained as gifts are as follows: sqh-Gal4 (Dr. Daniel P. Kiehart, Duke University), elav-Gal4-GS (Dr. R. J. Wessells, Wayne State University), elav-Gal4 and repo-Gal4 (Dr. Daniel F. Eberl, University of Iowa). The fly line used as the V5-positive control was a V5-tagged ataxin-3 Q80 line that we used in a previous publication[343].

Stocks and crosses were maintained at 25°C and approximately 60% humidity in diurnal incubators with 12-hour light/dark cycles. Flies were kept on a conventional cornmeal diet when being kept as stocks as well as throughout crosses and experimentation. The one exception to the standard diet was glutamine-substituted (L-Glutamine; Millipore Sigma) food that is described in the Results section. All flies were heterozygous for both driver and transgene, unless otherwise noted.

5.2.5 Quantitative Real-time Polymerase Chain Reaction

Five intact adults or 10 flies collected during development, depending on the experiment, had total RNA extracted using TRIzol (Life Technologies). Extracted RNA was then treated with
TURBO DNase (Ambion) and reverse transcription was carried out using a high-capacity cDNA reverse transcription kit (ABI). Finally, mRNA levels were quantified with the StepOne Real-Time PCR system using a Fast SYBR Green Master Mix (ABI). Primers used for original ATXN3-Q80 constructs:

\[
\begin{align*}
\text{ATXN3 F: } & 5'-\text{GAATGGCAGAAGGAGGAGGTTACTA- } 3'; \\
\text{ATXN3 R: } & 5'-\text{GACCCGTCAAGAGAGAATTCAAGT- } 3'; \\
\text{rp49 F: } & 5'-\text{AGATCGTGAAGAGCGCACAAG- } 3'; \\
\text{rp49 R: } & 5'-\text{CACCAGGAACTTCTTTGAATCCGG- } 3'
\end{align*}
\]

Primers used for Met-Null constructs:

\[
\begin{align*}
\text{Met-Null + ATXN3 F: } & 5'-\text{GGCGATGCTTAGAGCGAAGA- } 3'; \\
\text{Met-Null + ATXN3 R: } & 5'-\text{AATCGAGACCGAGGAGGAGG- } 3'; \\
\text{rp49 F: } & 5'-\text{AGATCGTGAAGAGCGCACAAG- } 3'; \\
\text{rp49 R: } & 5'-\text{CACCAGGAACTTCTTTGAATCCGG- } 3'
\end{align*}
\]

5.2.6 Drosophila examinations

For longevity experiments, adults were collected on the day of eclosion and monitored daily to record deaths. Flies were switched into fresh food every other day. Males and females were tracked separately. Number of adults tracked in each experiment is noted in figures and legends. For developmental death tracking, flies were observed from the embryo stage through eclosion or death, and deaths at each developmental stage (outlined in figures) were recorded daily.

5.2.7 Western blotting and quantification

Western blots were performed with either 3 whole flies or 5-10 flies collected during development per group depending on the experiment and driver being used. Samples were homogenized in boiling fly lysis buffer (50mM Tris pH 6.8, 2% SDS, 10% glycerol, 100mM
dithioreitol (DTT)), briefly sonicated, boiled for 10 minutes and centrifuged for 8 minutes at 13300xg at room temperature. The PXi 4 (Syngene) or ChemiDoc (Bio-Rad) were used to develop the Western blots, which were then quantified with GeneSys (Syngene) or ImageLab (Bio-Rad), respectively. Direct blue stains of total protein were done by submerging the PVDF membranes for 10 minutes in 0.008% Direct Blue 71 (Sigma-Aldrich) in 40% ethanol and 10% acetic acid and then rinsed with a solution of just 40% ethanol/10% acetic acid, before being air dried and imaged.

5.2.8 Filter-trap assays

Three intact adults or 5 flies collected during development, depending on the driver being used, per group were homogenized in 200µL of NETN buffer (50mM Tris, pH 7.5, 150mM NaCl, 0.5% Nonidet P-40) that was supplemented with a protease inhibitor (PI; S-8820, Sigma-Aldrich). The resulting lysates were then diluted with 200µL 0.5% SDS in PBS. Diluted lysates were sonicated, centrifuged at 4500xg for 1 minute at room temperature and then diluted further by combining 100µL lysate with 400µL PBS. Thirty-five µL of this final lysate of each sample were added to a Bio-Dot apparatus (Bio-Rad) and filter-vacuumed through a 0.45um nitrocellulose membrane (Schleicher and Schuell) that had been pre-incubated with 0.1% SDS in PBS. After samples were filter-vacuumed through the membrane, the membrane was rinsed twice with 0.1% SDS in PBS, incubated in primary and then secondary antibody, and analyzed via Western blotting.

5.2.9 Co-immunopurification assays

Ten intact adults or 15 flies collected during development, depending on the driver being used, were homogenized in 400µL of RIPA (50mM Tris, 150mM NaCl, 0.1% SDS, 0.5% deoxycholic acid, 1% NP40, pH7.4)+PI buffer, sonicated and then mixed with an additional 400µL of RIPA+PI. Samples were then tumbled at 4°C for 30 minutes. While tumbling, bead-bound antibodies (either MYC- or V5-tagged; Thermo Fisher Scientific) were prepped with three rinses
of RIPA+PI buffer. Following tumbling, 40µL of sample was combined with 10µL of 6X SDS, boiled for 10 minutes and then saved as an input. The remaining sample was combined with the prepped bead-bound antibodies and tumbled at 4°C for 2 hours. The beads were then rinsed 3 times with RIPA+PI with a 5-minute tumble at 4°C during each rinse. Beads then underwent 5 additional rinses with RIPA+PI without additional tumbling. Finally, bead-bound complexes were eluted by combining with 30µL of lysis buffer (50mM Tris pH 6.8, 2% SDS, 10% glycerol, 100mM DTT) and boiling for 5 minutes.

5.2.10 Statistics

Statistical tests used for each experiment are stated in their respective figure legends. Prism 8 (GraphPad) was used for the Kaplan-Meier log-rank tests for survival analysis, the Welch’s t-tests for RQ comparisons from qRT-PCRs, the unpaired t-test for western blot analysis, and RM two-way ANOVA with Geisser-Greenhouse correction and Tukey’s multiple comparisons test for developmental death stage analysis. Additional data collection, organization, and student’s t-tests were done in Excel (Microsoft) or Numbers (Apple). P-values were calculated by the software used for analysis and are outlined in the corresponding figures and legends, along with the number of biological replicates.

5.3 Results

5.3.1 Development of new constructs to examine RAN in SCA3

We generated two plasmids for mammalian cell transfection that contain the full, human ataxin-3 sequence with either a pure homomeric CAG repeat (ATXN3-Q80(CAG)) or an interrupted, alternating CAGCAA repeat (ATXN3-Q80(CAGCAA)). Both produce the same pathogenic 80Q tract in the translated ataxin-3 protein; however, the pure CAG repeat is susceptible to the formation of abnormal/slippery hairpin loops and could lead to RAN translation,
while the CAGCAA mRNA is less likely to produce hairpin loops or RAN peptides and precludes the production of proteins in alternative reading frames[226-233,241]. Constructs were HA-tagged in the normal, sense frame (encoding ATXN3 protein). MYC and V5 tags were added in the +1 and +2 sense reading frames respectively (Figure 5.1A). Our expectation was that these tags would allow us to biochemically detect the presence of any possible RAN translated polyserine (polyS) or polyalanine (polyA) proteins from the sense strand.

**Figure 5.1 Designing ataxin-3 constructs for studying RAN translation**
(A) Diagrammatic representation of construct design with an uninterrupted CAG triplet repeat or alternating CAGCAA repeats and HA, MYC, and V5 tags in the ataxin-3, +1 and +2 reading frames, respectively. Also shown are the expected translated sequences of full-length proteins from each sense reading frame. These proteins are expected to contain either polyglutamine, polyserine, or polyalanine repeats and their associated tags for identification. (B) Western blot from lysates of two different mammalian cell lines, HEK-293T or M-17, transfected with an empty pcDNA vector or a pcDNA-ATXN3-80(CAG) or pcDNAATXN3-80(CAGCAA) vector. Transfection and lysis were performed independently for each cell line and then loaded together in one gel as shown.

The ATXN3-Q80(CAGCAA) fly line has been employed in previous studies in our lab and no specific protein signal was observed in either the MYC- or V5-tagged reading...
Upon transfection of cultured, mammalian HEK-293T cells with either ATXN3-Q80(CAG) or ATXN3-Q80(CAGCAA), we observed no specific protein signal from alternate reading frames with ATXN3-Q80(CAGCAA) and its alternating repeat. However, transfection of ATXN3-Q80(CAG) possessing homomeric CAG repeats, resulted in the detection of polyA proteins translated from the +2 V5-tagged frame (Figure 5.1B). As SCA3 is a primarily neuronal disease, transfections were also conducted in a human neuroblastoma (M17) cell line producing similar results (Figure 5.1B). These results provide evidence that RAN translation is possible with the expanded ATXN3 transcript. Our next step was to determine whether this same phenomenon occurs in vivo.

5.3.2 Homomeric and alternating pathogenic ataxin-3 polyQ repeats are differentially toxic in Drosophila melanogaster

To investigate SCA3 RAN translation in vivo we again used Drosophila melanogaster[215,216,222,223,266,277,343,386]. Utilizing the versatility of Drosophila genetics along with the Gal4-UAS binary system of expression, we can express either the homomeric or the alternating polyQ repeat of pathogenic ataxin-3 transgenes individually in specific tissues[174,179,216,222,223,252,266,277,343,356,386]. Expression of each construct was validated via quantitative real-time polymerase chain reaction (qRT-PCR). In flies pan-neuronally expressing the transgenes individually, ATXN3-Q80(CAG) was expressed at lower levels than ATXN3-Q80(CAGCAA) (Figure 5.2A). However, no significant differences in mRNA levels were detected between the two when expressed individually in glial cells (Figure 5.2B). Western blot analyses of samples obtained by lysing adult flies expressing either ATXN3-Q80(CAG) or ATXN3-Q80(CAGCAA) pan-neuronally (Figure 5.2C) or in developing flies expressing either
version only in glia (Figure 5.2D) showed no significant difference in ataxin-3 protein level between the two transgenes.

Figure 5.2 Homomeric CAG repeats are more toxic than alternating CAGCAA in fly models of SCA3
(A) qRT-PCR of one-day-old female adult flies pan-neuronally expressing ATXN3-80(CAGCAA) or ATXN3-Q80(CAG). Means +/- SEM. p-value from Welch’s t-test. (B) qRT-PCR of developing flies expressing ATXN3-80(CAGCAA) or ATXN3-Q80(CAG) in glia. Means +/- SEM. “ns”: non-statistically significant; Welch’s t-test. (C) Western blot of one-day-old female adult flies pan-neuronally expressing ATXN3-80(CAGCAA) or ATXN3-Q80(CAG) with whole-lane quantification of total ataxin-3 protein levels. Means +/- SD. “ns”: non-statistically significant; unpaired t-test. (D) Western blot of developing flies expressing ATXN3-80(CAGCAA) or ATXN3-Q80(CAG) in glia. Means +/- SD. “ns”: non-statistically significant; unpaired t-test. (E) Survival analysis from male and female adult flies pan-neuronally expressing ATXN3-Q80(CAG), ATXN3-Q80(CAGCAA), or an empty vector insertion into the attP2 site (control). No adult male flies expressing ATXN3-Q80(CAG) eclosed as adults. “*”: p<0.0332; “**”: p<0.0021; “***”: p<0.0002; “****”: p<0.0001. p-values from log-rank tests. (F) Survival analysis from male and female adult flies expressing ATXN3-Q80(CAG), ATXN3-Q80(CAGCAA), or an empty vector insertion into the attP2 site (control) in glia. No adult flies eclosed in lines expressing either ataxin-3 transgene in glia.
We next determined if the presence of homomeric CAG repeats, and thus the potential for hairpin loop formation and RAN translation, resulted in differences in disease progression compared to the alternating, CAGCAA repeats. We performed longevity experiments with flies expressing either construct pan-neuronally. Survival analyses from males and females revealed that ATXN3-Q80(CAG) led to significantly decreased longevity compared to those expressing the alternating, ATXN3-Q80(CAGCAA) in either sex (Figure 5.2E). In males, particularly, flies expressing the alternating repeat lived as long as 46 days, while none of their homomeric repeat-expressing counterparts were able to emerge as adults, dying as pharate adults instead (Figure 5.2E; Supplemental figure S5.1).

Expression of ataxin-3 in glia, regardless of the transgene used, was developmentally lethal and no adults emerged for longevity (Figure 5.2F). The experiments conducted in figures 5.2B and 5.2D were conducted with flies collected during development from lines expressing either construct in glia. Although we were unable to study differences in longevity between our two lines in glia, the high levels of toxicity in these cells compared to neurons points to an interesting potential role for glia in SCA3 pathogenesis and is an active field of investigation, including a recent study that identified impaired oligodendrocyte maturation as a contributor to pathogenesis in a transgenic mouse model of SCA3[387].

In addition to pan-neuronal and glial expression, we performed survival studies in flies expressing either of the transgenes ubiquitously. This is an expression model that we suspected would be developmentally lethal and we thus focused on tracking the stages of development when flies died, similar to what we did in previous work[386]. Analysis yielded a similar pattern of results as our other expression models: we observed that ATXN3-Q80(CAG)-expressing flies died
earlier during development than ATXN3-Q80(CAGCAA)-expressing ones (Supplemental Figure S5.1).

These data indicate that although the sequence of each ATXN3 transgene encodes the same pathogenic 80Q ataxin-3 protein with comparable protein levels, the presence of homomeric CAG is more toxic than the alternating repeat. This opens the possibility that factors outside of the full ataxin-3 protein alone may influence SCA3 pathogenesis. To determine if RAN translation is one of those factors in SCA3, we next examined the presence of proteins from alternative reading frames that could appear due to RAN translation.

5.3.3 Lack of evidence of RAN peptides in Drosophila from pathogenic ataxin-3

To detect the presence of potential RAN proteins, we took advantage of the alternate frame tagging of our ATXN3 constructs outlined in Figure 5.1. These tags allow us to probe for the polyS (+1 MYC-tagged frame) and polyA (+2 V5-tagged frame) proteins that we anticipate would be translated if RAN translation were occurring – particularly the polyA proteins that we observed in cultured mammalian cells expressing ATXN3-Q80(CAG). We conducted these probes using several biochemical assays in both pan-neuronal and glial expression.

Simple Western blot analysis of whole fly lysates from lines pan-neuronally expressing either ATXN3-Q80(CAGCAA) or ATXN3-Q80(CAG) did not reveal any detectable, translated protein in either the +1 MYC-tagged or +2 V5-tagged frames (Figure 5.3A). As aggregation of pathogenic ataxin-3 precedes toxicity in our Drosophila models of SCA3[222,223,277,343,386], we next wondered whether RAN species were trapped in larger aggregates, along with full-length ataxin-3 and thus may not have migrated properly into a standard SDS-PAGE gel. Thus, we conducted filter-trap assays designed to capture higher-order aggregated
species. There were no detectable MYC- or V5-tagged proteins in these assays (Figure 5.3B).

Figure 5.3 No evidence of RAN translation in pan-neuronal Drosophila models of SCA3
(A) Western blot of one-day-old adult female flies pan-neuronally expressing ATXN3-80(CAGCAA), ATXN3-Q80(CAG), or a negative control lacking any ataxin-3 transgene to examine MYC- or V5-tagged translated proteins from alternate reading frames. No specific signal was detected in the +1 or +2 reading frames. n ≥ 3 each. (B) Filter-trap assays of lysates from one-day-old adult female flies pan-neuronally expressing ATXN3-80(CAGCAA), ATXN3-Q80(CAG), or a negative control lacking any ataxin-3 transgene and blotting for either V5- or MYC-tagged proteins. V5- and MYC-blotted filter traps were conducted independently. No signal above background was detected in either case. n ≥ 3 each. (C, D) Co-immunopurification with anti-MYC (C) or anti-V5 (D) antibody-tagged beads of lysates from one-day-old female flies pan-neuronally expressing ATXN3-Q80(CAG) or ATXN3-Q80(CAGCAA). No alternate ataxin-3 frame MYC- or V5-tagged proteins were detected. n ≥ 3. Asterisk (*) denotes a non-specific band observed with the V5 antibody under some conditions. For all applicable panels, MYC-positive controls were from whole flies ubiquitously expressing a MYC-tagged ataxin-3 acquired from the Bloomington Drosophila Stock Center; V5-positive controls were from dissected heads of flies expressing a V5-tagged ataxin-3 in the fly eyes that was used in a previous publication.
Figure 5.4 No evidence of RAN translation in the glia of Drosophila models of SCA3

(A) Western blot of developing flies expressing ATXN3-80(CAGCAA), ATXN3-Q80(CAG), or a negative control lacking ATXN3 in glia. No specific signal was detected in the +1 or +2 reading frames. n ≥ 3 each. (B) Filter-trap assays of lysates from developing flies expressing ATXN3-80(CAGCAA), ATXN3-Q80(CAG), or a negative control in glia. No signal above background was detected in either reading frame. n ≥ 3 each. (C, D) Co-immunopurification with anti-MYC (C) or anti-V5 (D) antibody-tagged beads from lysates of developing flies expressing ATXN3-Q80(CAG) or ATXN3-Q80(CAGCAA) in glia. No alternate ataxin-3 frame MYC- or V5–tagged proteins were co-immunopurified. n ≥ 3. Asterisk (*) denotes a non-specific band observed by the V5 antibody under some conditions. For all applicable panels, MYC-positive controls were from whole flies ubiquitously expressing MYC-ataxin-3; V5-positive controls were from the heads of flies expressing a V5-ataxin-3 in fly eyes.
Finally, we conducted co-immunopurification (co-IP) assays targeting MYC- and V5-tagged proteins, in case the levels of these proteins were too low for capture by simple lysis of flies or through filter-traps. MYC and V5 beads were used in separate co-IP experiments. We again did not observe specific signal in either our MYC (Figure 5.3C) or V5 (Figure 5.3D) co-IPs.

These experiments were repeated in glial tissue. Due to the developmental lethality of glial expression of either ataxin-3 transgene in *Drosophila*, flies were collected during development for biochemical analysis, at either pupal or pharate adult stages. Western blot analyses recapitulated findings from pan-neuronal expression for all assays conducted (Figure 5.4). We did not detect any proteins translated in alternative, sense frames, suggesting that RAN translation is not a factor in SCA3 pathogenesis in *Drosophila*. The question still remains: why do we observe significantly higher toxicity from pathogenic ataxin-3 transgenes with homomeric CAG repeats compared to those with alternating CAGCAA tracts? To attempt to answer these questions, we turned to the previously theorized concept of contributions to ataxin-3 toxicity from its RNA[228,229,231,233,381].

5.3.4 **New transgenes to investigate mRNA-based toxicity in Drosophila models of SCA3**

mRNA-based toxicity can occur in genes with long repeat sequences that allow for the formation of RNA foci; these foci can recruit mRNA away from their native functions, recruit mRNA-binding proteins, and can become toxic[381]. To investigate mRNA toxicity in fly models of SCA3 ATXN3, we examined the impact of ATXN3 mRNA with homomeric CAG repeats versus CAGCAA repeats, while preventing the translation of ataxin-3 protein. We utilized the pathogenic ataxin-3 transgenes from the previous assays (ATXN3-Q80(CAG) and (CAGCAA)) and mutated all ATG (start) codons into TGA (stop) codons. These methionine-less constructs are referred to as Met-Null(CAG) and Met-Null(CAGCAA) throughout the rest of this work. They
differ from their homomeric and alternating repeat, protein expressing, counterparts only at the ‘start’ to ‘stop’ mutations (Figure 5.5A). Western blots from flies expressing these transgenes in all neurons or in all glia did not yield any specific signal in the ataxin-3 frame (Figure 5.5B, C).

**Figure 5.5 ATXN3 constructs to investigate mRNA-toxicity**

(A) Each start codon was mutated into a stop codon (Stop) in ATXN3. Mutations were identical for the Met-Null(CAG) and Met-Null(CAGCAA) constructs, changing every ‘ATG’ into ‘TGA’. (B, C) Western blots of one-day-old adult female and male flies expressing Met-Null(CAGCAA), Met-Null(CAG), or a negative control lacking ATXN3 in all neurons (B) or glia (C). n ≥ 3 each. No ataxin-3 protein was detected. n ≥ 3 each. For B, C, MYC-positive controls were from whole flies ubiquitously expressing MYC-ataxin-3; V5-positive controls were from dissected heads of flies expressing V5-ataxin-3 in fly eyes.

**5.3.5 Minimal survival difference in flies expressing alternating or homomeric, Met-Null CAG/A repeats**

We examined if differences in longevity between alternating and homomeric ataxin-3 CAG repeats observed in figure 5.2 resulted in part from toxicity at the mRNA level. We expressed Met-Null(CAG) or Met-Null(CAGCAA) pan-neuronally in flies and examined their survival (Figure 5.6A). Analyses of these survival data revealed significant sex-specific differences in longevity in flies expressing both homomeric and alternating Met-Null constructs, with females outliving males in both cases (Figure 5.6A). In addition to outliving their male counterparts, females flies pan-neuronally expressing Met-Null(CAGCAA) lived significantly longer than those expressing Met-Null(CAG) (Figure 5.6A). However, this result was reversed in male flies expressing either of the
Met-Null constructs: males pan-neuronally expressing Met-Null(CAG) lived slightly, but significantly, longer than ones expressing Met-Null(CAGCAA) (Figure 5.6A). qRT-PCR of these same flies showed no difference in pan-neuronal expression between the homomeric and alternating CAG repeat constructs (Figure 5.6B).

Figure 5.6 Homomeric CAG mRNA can be slightly more toxic than alternating CAGCAA in flies expressing Met-Null ATXN3
(A) Survival analyses from adult male and female flies pan-neuronally expressing Met-Null(CAG) or Met-Null(CAGCAA). “*: p<0.0332; ***: p<0.0021; *****: p<0.0001. p-values from log-rank tests. (B) qRT-PCR of one-day-old adult female flies pan-neuronally expressing Met-Null(CAGCAA) or Met-Null(CAG). Means +/- SEM. “ns”: non-statistically significant; Welch’s t-test. (C) Survival analyses from adult male and female flies expressing Met-Null(CAG) or Met-Null(CAGCAA) in glia. “ns”: non-statistically significant; “*: p<0.0332; ***: p<0.0021; *****: p<0.0001. p-values from log-rank tests. (D) qRT-PCR of one-day-old female flies expressing Met-Null(CAGCAA) or Met-Null(CAG) in glia. Means +/- SEM. “ns”: non-statistically significant; Welch’s t-test.
We repeated the same assays in glia. Male flies expressing homomeric Met-Null(CAG) ATXN3 lived slightly shorter lives than those expressing alternating repeats, but this difference was not significant in females (Figure 5.6C). qRT-PCR confirmed that there were no significant differences in expression between the two transgenes in glia (Figure 5.6D). We conclude that mRNA toxicity is unlikely to be significant in SCA3 pathogenesis.

5.3.6 Alternative mechanisms of homomeric CAG toxicity

Recently, a hypothesis emerged that translation of pure, expanded CAG repeats in polyQ disorders depletes available glutaminy charging tRNA\textsuperscript{Gln-CUG}, which pairs exclusively with CAG codon\cite{388}. This depletion would result in a transcriptome that is prone to mistranslation and could account for pathological differences observed between pure and interrupted CAG transcripts\cite{388}. This notion provided a potential mechanism to help explain the differences we observe in our own experiments between both our translated protein and Met-Null versions of ATXN3.

Our initial probe into this hypothesis focused on the availability of glutamine. In order for tRNA\textsuperscript{Gln-CUG} to be charged and contribute to the translation of long CAG repeats, there must also be an appropriate level of available glutamine to charge tRNA\textsuperscript{Gln-CUG}. Perhaps, glutamine is a bottleneck, decreasing in levels as CAG repeats get longer and causing toxicity as a result of depletion. We supplemented our normal fly food with 0.00, 12.5, or 50.0 grams of glutamine per liter of food to determine if increased dietary consumption of glutamine could overcome potential reductions in this amino acid. Concentrations for glutamine supplementation were based on previous studies of amino acid supplementation in flies\cite{389,390}.

We tested the effect of increasing amounts of available glutamine in the fly diet using ubiquitous expression of ATXN3-Q80(CAG) and ATXN3-Q80(CAGCAA) that resulted in high toxicity (Supplemental figure S5.1). By tracking developmental deaths, we were able to rapidly
assess if there is any benefit from increasing amounts of glutamine. Additional dietary glutamine did not provide any detectable benefit to developing flies expressing ATXN3-Q80(CAG) or ATXN3-80(CAGCAA) (Supplemental figure S5.3). These data suggest that glutamine itself is not a limiting factor in our models.

While glutamine supplementation does not appear to be the answer in overcoming the hypothesized tRNA depletion that results from expanded CAG repeat translation, there are additional components to the translation machinery that could still be tested in this model. These components, including the tRNA charging machinery and the tRNA$_{Gln-CUG}$ itself, are outside the scope of this work, but comprise viable candidates for future investigations.

5.4 Discussion

Here, we took aim at identifying the presence and potential influence of RAN translation in Drosophila models of SCA3. As with previously identified RAN translation-associated diseases, such as SCA8 and Huntington’s disease, SCA3 is a CAG repeat-expansion disorder with a long repeat stretch. Based on prior reports on ATXN3 mRNA and that of other CAG repeat disorders[228,229,231,233,381], SCA3 toxicity may be influenced by the ability of ATXN3 mRNA to form secondary structures and by the presence of polyA species that have been identified in SCA3 patient tissue[382]. These studies pointed to the possibility of RAN translation and made SCA3 a candidate for exploratory research to add it to the growing list of repeat-expansion diseases influenced by RAN translation.

We utilized transgenic Drosophila lines to express two versions of pathogenic ATXN3 in the tissues primarily affected in SCA3, neurons and glia. The two versions of ATXN3 had either an uninterrupted CAG or an alternating CAGCAA repeat in their polyQ-encoding tract and represented ATXN3 mRNA capable (pure CAG) or incapable (alternating CAGCAA) of forming
secondary structures for RAN translation. Although we observed differences in toxicity between the two transgenes, we did not find biochemical evidence of RAN-translated proteins in *Drosophila*. This lack of evidence does not preclude the possibility of RAN translation occurring in the sense, ataxin-3 HA-tagged reading frame, or RAN products in the antisense strand. Detection of these possibilities is outside the capability of our current model systems; however, our work in mammalian cells, and the presence of polyA proteins in SCA3 patient tissue[382], supports our focus on aberrant translation of the polyS and polyA reading frames as a reasonable measurement of the occurrence of RAN translation in flies.

The absence of evidence of RAN translation in our SCA3 model system led us to consider mRNA playing a possible role in ATXN3-based toxicity, as previously reported by the Bonini lab[228]. We tested this non-mutually exclusive alternative to RAN translation by using redesigned, pathogenic ATXN3 transgenes whose potential start codons were mutated to stop codons. Expression of these ‘Met-Null’ transgenes yielded minor differences in toxicity that were not consistent between sexes and between glia and neurons. These results provide some support for the possibility of mRNA-based toxicity in SCA3, as indicated previously[228], with the caveat that evidence in our study points to a highly subdued role for pure CAG mRNA toxicity in SCA3.

Differences in extent of mRNA toxicity between our work and prior findings[228] could be explained by the potentially higher level of expression reached by prior transgenic lines (P element-based transgenesis that can lead to multiple insertions) compared to the single copy, safe-harbor integration of transgenes through the phiC31 integrase system that we employed. Also, earlier observations were reached by using isolated, hyperexpanded CAG expansions[228], instead of the patient-range used here. It is possible that mRNA toxicity does not feature prominently in
the context of patient-range expansions, or that mRNA toxicity from isolated CAG repeats is stronger than when those repeats are in the context of ATXN3 sequence.

Various mechanisms may explain the contribution in toxicity by CAG repeat-containing mRNA. The uninterrupted CAG mRNA secondary structure could be more prone to aggregation than CAGCAA mRNA and help accelerate aggregation of ataxin-3 protein [226,227,241]. Homomeric mRNA could also be more prone to stalling at the ribosome, similar to other repeat-expansion diseases, triggering dysregulation of essential processes [391]. In addition, these prolonged homomeric repeats could create shortages in the tRNA charging machinery that may further injure the cellular environment. Together with toxicity stemming from ataxin-3 protein itself, these mRNA-based mechanism might hasten SCA3.

Pathogenesis from RAN translation and RNA-based toxicity are being increasingly recognized as key contributors to various diseases. Although we suspect SCA3 to be a likely member of the RAN translation family, we did not find evidence of it in Drosophila and found minimal evidence of exclusively RNA-based toxicity. We conclude that the Drosophila model system is not well suited for studying this type of RNA-based toxicity. Our study introduces to the field additional tools that can be utilized towards understanding SCA3 and other, similar disorders.

Citation:

CHAPTER 6: FUTURE DIRECTIONS

Through the development of novel Drosophila genetics and tools, the work of this dissertation took aim at understanding the domains and protein-protein interactions of the polyQ-expanded ataxin-3 protein that play a role in its toxicity in SCA3 pathogenesis. While I set out to with the intention of establishing as comprehensive understanding of ataxin-3 protein context as possible, this work is far from complete. On top of the countless uncategorized protein-protein interactions waiting to be found and studied, there is an additional, established domain of the ataxin-3 protein that talented members of the Todi lab have already commenced investigations into, the first ubiquitin binding site of ataxin-3’s catalytic domain (UbS1). As it relates to my work, the major questions that remain are centered around investigations into the protein-interactions with ataxin-3’s UIMs (Chapter 3) as well as the potential role for mRNA-based toxicity in SCA3 (Chapter 5).

The first questions that need to be answered are: what is the nature of ataxin-3’s interaction with the chaperone protein Hsc70-4 and how does this interaction result in the increased aggregation and toxicity of polyQ-expanded ataxin-3 toxicity? Initial investigations into these questions focused on determining if there is a direct interaction between the two proteins and whether this interaction disrupts the functional capabilities of either protein but, thus far, neither have yielded promising results. Instead, my focus shifted to the possibility that the interaction between ataxin-3 and Hsc70-4 was part of a larger functional complex that gets disrupted upon polyQ expansion. Utilizing the same fly eye scoring system that was used to clarify the impact of Hsc70-4 on ataxin-3 toxicity (figure 3.6), I conducted a larger screen of genes involved in the chaperone machinery to find those whose knockdown causes a similar presentation of eye toxicity
to that of Hsc70-4. Genes of interest from this screen could provide a potential link between ataxin-3 and Hsc70-4 and would warrant further investigation.

From this initial screen, I was able to identify two fly genes in the network of chaperone components that could potentially interact with both Hsc70-4 and ataxin-3 and whose knockdown impacted ataxin-3 toxicity in a manner similar to Hsc70-4 knockdown, CHIP and Droj2. While the E3 ubiquitin ligase CHIP (human ortholog, STUB1) has a well reported functional and proposed pathological interaction with ataxin-3[371,392], the relationship between Droj2 and ataxin-3 is less studied. As a J-protein (human orthologs DNAJA1 and DNAJA4), Droj2 and its orthologs has published interactions with heat shock proteins in the chaperone folding system[393,394], including Hsc70-4[357], making its impact on ataxin-3 toxicity particularly interesting as a potential link between the two proteins. To further study this trio of interactions, it will be necessary to lean on the power of the Drosophila genetics system to develop new lines that express biochemically tagged versions of either Hsc70-4 or Droj2. These new Drosophila lines would allow for co-expression and co-immunoprecipitation of each protein in order to determine if there is any association between each set of proteins and, more importantly, if the same complexes that allow Hsc70-4 to enhance ataxin-3 toxicity also contain Droj2. If an association between all three proteins is observed, it is possible that Droj2 could be a link between ataxin-3 and Hsc70-4 that helps us better understand this chaperone’s aberrant function in the presence of expanded polyglutamine tracts. Any associations could then be further validated as direct interactions in vitro using bacterially produced pure proteins and pull-down assays.

The utility of the RNAi eye screen in flies in identifying interesting ataxin-3 toxicity modifying proteins in the chaperone machinery inspired an additional future project that focuses on searching more broadly for proteins that interact with polyQ-expanded ataxin-3 and modulate
its toxicity. A larger, more general screen would not only provide an alternative approach for any potential negative results with Droj2 and the Hsc70-4, but it also creates a rapid and reproducible way to study the ataxin-3 functional interactome in both an expanded and wild-type state. Additionally, these same screens could be conducted with other polyQ proteins or truncated pure polyQ protein, as in figure 3.7, to determine if any effects observed are specific to the context of a given polyQ protein or polyQ disorders as a whole. Figure 6.1 shows a consolidated example of the target identification power that these screens can offer as you can compare the effects of knockdown over time to a respective control as well as to each other. Genes with interesting score trends can be further analyzed for significance and investigated for other hallmarks of disease like aggregation. Overall, the methodology established in these screens has broad applications in multiple disease models and will be useful in developing new avenues of research for the lab moving forward.

The final project that I will discuss in this work is the future of the mRNA-based toxicity studies started in Chapter 5. In that chapter, I identified a minimal role for mRNA-based toxicity that did not seem to fully explain the differences in longevity and accumulation observed between proteins with either homomeric CAG or alternating CAGCAA repeats. There are two theories to help explain this that are discussed in the chapter that I believe warrant future investigation. The first is that the secondary structure of the pure CAG mRNA could be more prone to aggregation which could then help seed the aggregation of ataxin-3 at the protein level more efficiently than CAGCAA mRNA. This theory could be explored using fly genetics to co-express either Met-Null(CAG) or Met-Null(CAGCAA) with ATXN3-Q80(CAGCAA) or ATXN3-Q80(CAG). If pure CAG mRNA is in fact acting as a seed for ataxin-3 protein aggregation, it is expected that co-expressing it with ATXN3-Q80(CAGCAA) would synergistically enhance toxicity in survival
analysis beyond the toxicity of the protein or mRNA individually. It is also expected that these effects would be mirrored, to a lesser extent perhaps, with ATXN3-Q80(CAG) co-expression models and reduced in models co-expressing either ataxin-3 protein with the Met-Null(CAGCAA) mRNA.

The other mRNA-based toxicity theory is that the presence of a prolonged and uninterrupted CAG repeat sequence creates a shortage in the tRNA charging machinery necessary to translate glutamines[388]. While the first test into this theory substituting excess glutamine into fly food (supplemental figure S5.3) did not result in any improvements in longevity, there are other elements to the glutamine charging process that could still be acting as bottlenecks to translation. These elements include the specific CTG glutamine tRNA that corresponds to the CAG codon and the glutamine synthetase that aids in the charging of CTG tRNA with a glutamine for translation. New constructs for each of these proteins could be designed for transgenic expression in flies also expressing pathogenic ataxin-3. Longevity studies could then be employed to determine if the over-expression of either has an effect in overcoming any potential translation dysfunction caused by an expanded repeat CAG repeat sequence.

The importance of protein context is relevant far beyond the study of polyglutamine disease pathogenesis. Searching outside the known primary contributors to disease progression provides a more comprehensive view of all that goes into toxicity and should open the door for the discovery of therapeutic entry points in many diseases that currently lack any form of treatment. This dissertation, the novel Drosophila tools that were developed in completing it, and the focus on the entire context of a disease protein represent an approach that I believe to be valuable across the family of polyglutamine diseases and beyond. Although the work presented here covers many of
ataxin-3’s most important domains, there is still a great deal of work to be done in what I believe to be a very exciting avenue of future research.

Figure 6.1 Heat map visualization of effect of gene knockdowns in Drosophila eyes co-expressing pathogenic ataxin-3
Displayed are visualizations of the Mean Rank comparisons from Dunn’s multiple comparison test between each gene knockdown and their respective controls. A) Knockdowns at the attp40 chromosome 2 insertion site with females and males separated. B) Knockdowns at the attp2 chromosome 3 insertion site with females and males separated. Legend displayed with green representing improved eye phenotype (based on score) compared to control upon gene knockdown, and red representing worsening eye phenotype (based on score) compared to control upon gene knockdown. Brighter color is associated with larger Mean Rank change in either direction and thus a more pronounced change in eye phenotype.
APPENDIX

Figure S2. 1
Anti-HA antibody shows similar difference in protein levels compared to the MJD antibody. Whole flies lysates from independent crosses. Ataxin-3 isoforms 1 or 2 were driven in all fly neurons throughout development and in adults.

Figure S2. 2
Another set of cytoplasmic/nuclear fractionation preparations similar to, but independent of, figure 1G. Results from these blots were quantified and included in histograms in figure 1G.
<table>
<thead>
<tr>
<th>sqh-Gal4 &gt;</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neither Ataxin-3 isoform</td>
<td>Flies develop normally and emerge successfully as adults; normal adult longevity</td>
</tr>
<tr>
<td>Ataxin-3 Isoform 1</td>
<td>Vast majority die as pharate adults. Some pupal lethality. No adults eclose successfully.</td>
</tr>
<tr>
<td>Ataxin-3 Isoform 1 + Ataxin-3 Isoform 2</td>
<td>Vast majority die as pupae. Some pharate adult lethality. No adults eclose successfully.</td>
</tr>
</tbody>
</table>

Figure S2. 3
Worsening of phenotype when both isoforms 1 and 2 of pathogenic ataxin-3 are expressed in all fly tissues and at all stages. Driver was sqh-Gal4. All flies were heterozygous for driver and constructs.

Figure S2. 4
Difference in ataxin-3 levels in pupae. Left: blots from whole pupal lysates where sqh-Gal4 is driving the noted pathogenic ataxin-3 variant. Five pupae per lysate were homogenized in hot SDS lysis buffer, sonicated, boiled, centrifuged and loaded on SDS-PAGE gel. Each lane is from an independent repeat. Right: quantification of signal from the blot on the left. P values are from student’s t-test comparing middle and right lanes separately to left lane.
Figure S2. 5
Isoform 1 protein levels after longer-term autophagy or proteasome inhibition. HEK-293T cells were transfected with isoform 1, then treated with 20 µM MG132 or 100 µM Chloroquine for the indicated times. Blot from whole cell lysis. Each lane is independent.

Figure S2. 6
qRT-PCR results from the same experimental set as in figure 7, using HEK-293T cells expressing the constructs noted in the X axis. N=3 independent biological replicates. No significant differences based on Student’s t-test comparing each group individually to “Iso 2”.

<table>
<thead>
<tr>
<th>Treatment Time (Hours)</th>
<th>MG132</th>
<th>Chloroquine</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>24</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>24</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>0</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>24</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>0</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>24</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>0</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>24</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Anti-Ataxin-3 (MJD)

Direct Blue (Loading)
Figure S3. 1 Ataxin-3 domains impact the toxicity of its polyQ-expanded variant in glia.

(A) Longevity curves from adults expressing the noted versions of ataxin-3 in all fly glia, throughout development and in adults. These are the same lines as in Figure 1. P values are from log-rank tests with Holm-Bonferroni adjustment. P values in italics are not statistically significant. There is no statistical difference for 100% death between flies expressing ‘Intact’ versus ‘Inactive’, even though 50% lethality is different. This is because each vial with flies is treated as one individual and for statistical purposes the day when all flies are dead is considered, not the rate of death over the course of time for each vial. (B) Negative geotaxis assay comparing motility of flies not expressing any pathogenic ataxin-3 versus flies expressing pathogenic ataxin-3 with mutated VBM in all glial cells. Driver, as in panel (A), was repo-Gal4. No statistical differences are observed at any time point between the groups based on two-tailed Student’s t-tests. We selected to conduct this assay for the VCP-binding mutations since in longevity curves we noticed no significant differences in the longevity of flies not expressing pathogenic ataxin-3 compared to ones expressing the SCA3 protein with the VCP-binding mutation. Note: The ‘Intact’ line shown here is a different transgenic line than the one we generated for the polyQ studies concerning the UIMs that are shown in Figure 3.2 and beyond. Those lines (Figure 3.2 and on) have a ‘CAGCAA’ repeat for reasons summarized in main text, compared to a pure ‘CAG’ repeat for all lines shown in Figure 3.1 and in this figure, which were generated for prior work. The lines generated for Figure 3.2 also have an optimized Kozak sequence. As we showed earlier[277], the ‘CAGCAA’ repeat-harboring ‘Intact’ line is at higher protein levels and is more toxic than the pure ‘CAG’ repeat shown here. For all the work in this study, isogenic sets are compared only to one another, that is, none of the lines shown here and in Figure 3.1 is directly compared to lines in Figure 3.2 and beyond, since these lines were conceptualized as isogenic sets from the beginning. However, we remind the reader that all ataxin-3 lines are inserted into the same chromosomal site of Drosophila, attP2, in the same orientation and as a single copy number.
Figure S3. 2 Levels of polyQ proteins, related to Figure 3.2D.
Western blots from lysates of whole, dissected fly heads; 25 heads per group. We caution here that the comparison of overall polyQ proteins is not precise, especially concerning polyQ80-UIM3, which leads to markedly severe eye degeneration. The reason for generating these truncated lines was to conduct an initial exploration of the role of UIMs on the isolated polyQ of ataxin-3. The main focus of this work is full-length ataxin-3 and the role of UIMs in its pathogenicity, explored in subsequent figures. Future work will dissect how individual UIMs impact various properties of the isolated polyQ80 in a tissue- and age-dependent manner.
Figure S3. 3 Uncropped blot from panel 3.6E.

Figure S3. 4 Relative importance of UIMs on fly longevity.

(A) Longevity curves of adult flies expressing the denoted transgenes pan-neuronally. Continuous curves are the same as in Figure 3.1C; X-axis is extended to 86 days to show the full longevity of control flies. Blue dots are ‘approximated longevity’ of flies expressing pathogenic ataxin-3 with mutated UIMs, based on data from Figure 3.3C. The UIM* flies are isogenic to the polyQ truncation transgenes and the full-length ataxin-3 transgene used in Figures 3.2–3.8; therefore, the UIM* flies cannot be directly compared to the lines used for Figure 3.1, since the Kozak sequence is different; also, the polyQ is encoded by a ‘CAGCAA’ repeat for flies used in Figures 3.2–3.8, compared to a pure ‘CAG’ for flies in Figure 3.1. Lines in Figure 3.1 are all isogenic to each-other. Calculations for the rough estimation are on the table on the right. (B) Another measure of improvement in fly longevity when the VBM or the UIMs are mutated, compared to their respective, isogenic ‘Intact’ controls. Columns D and G show percent improvement in longevity for flies expressing VBM-mutated ataxin-3 compared to their own isogenic ‘Intact’ (column D; data from Figure 3.1) and flies expressing UIM-mutated ataxin-3 compared to their own isogenic ‘Intact’ (column G; data from Figure 3.1). Green font indicates that UIM* flies showed more improvement than VBM* ones.
Motility assays validate non-toxic presence of N-VCP in Drosophila.
Negative geotaxis assays conducted in flies pan-neuronally expressing N-VCP. Driver: elav-Gal4. Tests were performed weekly starting on day 7 and continuing through day 42. N-VCP-expressing flies were compared to flies with the same genetic background and pan-neuronal driver but lacking N-VCP. Means +/- SD. No statistical significance was found under any conditions; two-tailed student’s t-tests. N>70 flies per genotype.
Figure S4.2 Daily alive adult numbers from Figure 4.2C. Another graphical representation of the survival analyses shown in figure 4.2C. Whereas figure 4.2C shows the ‘probability of survival’, these curves show the ‘daily alive percentages’ of flies in each group. As discussed in ‘Results’, flies denoted by green color likely represent the effect of N-VCP more accurately as improvements in longevity begin during development. As the zoomed-in portion of the longevity curve shows, the strongest flies expressing N-VCP (green) outlive those expressing pathogenic ataxin-3 alone (red), approaching ‘Control’ flies, denoted by black color in the upper portion of the graph.

A point to address when comparing the results from the ubiquitous driver (this graph and figure 4.2C) to those with the pan-neuronal drivers (figure 4.2D,E): a handful of flies that co-express pathogenic ataxin-3 alongside N-VCP in all tissues live markedly longer than those that express the same proteins only in neurons. This outcome may indicate non-neural aspects of lethality from ataxin-3, which deserve future attention. Indeed, we recently reported that pathogenic ataxin-3 causes different extents of lethality in neuronal versus glial cells in Drosophila[343].
Figure S4. 3 Motility assays with the ‘stronger’ pathogenic ataxin-3.
Negative geotaxis performed with flies expressing the ‘stronger’ pathogenic ataxin-3, without or with N-VCP. Means +/- SD.
Statistics: one-tailed student’s t-tests comparing “red” and “green” histograms to each other. P value not shown if not <0.05. N>70 flies per genotype.
Figure S4. VCP increases pathogenic ataxin-3 toxicity in a dose-dependent manner.
(A, B) Longitudinal observation of developmental stage death among flies ubiquitously expressing the noted version of pathogenic ataxin-3 with zero, one, or two copies of exogenous full-length VCP. Driver: sqh-Gal4. Means ± standard deviations. P-values: one-tailed student’s t-tests, shown below graphs.
Figure S5. Ubiquitous expression longevity data and individual comparisons from survival analysis in figure 2
(A) Ubiquitous expression of ATXN3-Q80(CAG) and ATXN3-Q80(CAGCAA) individually was developmentally lethal in flies. Observations of developmental deaths in both lines based on the fly life cycle (shown at top of figure) showed ATXN3-Q80(CAG) flies dying primarily at earlier developmental stages than flies ubiquitously expressing ATXN3-Q80(CAGCAA). These observations were confirmed with survival analysis in tissue-specific expression models shown in figure 2. (B) Individual comparisons of all pan-neuronal expressing fly lines shown in figure 2E. "*: p<0.0332; "**": p<0.0021; "***": p<0.0002; "****": p<0.0001. p-values from log-rank test. No adults eclosed in the glial expression analysis in figure 2F, thus no individual comparisons were made.
Figure S5. 2 Individual comparisons from survival analysis in figure 6
(A) Individual comparisons of all pan-neuronal expressing fly lines shown in figure 6A. "*": p<0.0332; "**": p<0.0021; "***": p<0.0002; "****": p<0.0001. p-values from log-rank test. (B) Individual comparisons of all glial expressing fly lines shown in figure 6C. "ns": non-statistically significant; "*": p<0.0332; "**": p<0.0021; "***": p<0.0002; "****": p<0.0001. p-values from log-rank test.
Figure S5. 3 Developmental deaths of ATXN3-Q80-expressing flies with glutamine supplementation

(A) Developmental death tracking based on the percentage of fly deaths at each developmental stage in flies ubiquitously expressing ATXN3-Q80(CAG) in food with 0.00, 12.5, or 50.0 g/L glutamine supplementation. Driver was sqh-Gal4. Each liter of food made accounted for 100 fly vials, with each vial then receiving an additional 0.000, 0.125, or 0.500 g/L of glutamine respectively. Means of percentages at each stage +/- SEM. “ns”: non-statistically significant. p-values of all comparisons shown to the right of graph from RM two-way ANOVA with Geisser-Greenhouse correction and Tukey’s multiple comparison test. n ≥ 3 vials each. (B) Developmental death tracking based on percentage of fly deaths at each developmental stage in flies ubiquitously expressing ATXN3-Q80(CAGCAA) in food with 0.00, 12.5, or 50.0 g/L glutamine supplementation in food. Means of percentages at each stage +/- SEM. “ns”: non-statistically significant. p-values of all comparisons shown to the right of graph and from RM two-way ANOVA with Geisser-Greenhouse correction and Tukey’s multiple comparison test. n ≥ 3 vials each.
REFERENCES


dentatorubral and pallidoluysian atrophy (DRPLA). Nat Genet 1994, 8, 177-182, doi:10.1038/ng1094-177.


142. Pluciennik, A.; Liu, Y.; Molotsky, E.; Marsh, G.B.; Ranxhi, B.; Arnold, F.J.; St-Cyr, S.; Davidson, B.; Pourshafie, N.; Lieberman, A.P.; et al. Deubiquitinase USP7 contributes to...


179. Todi, Tsou, doi:10.1007/s00401
178. spinocerebellar alpha(1A) Ishiguro, doi:10.1007/s12311
177. IRES doi:10.1016/j.cell.2013.05.059.
175. Ishikawa, doi:10.1007/s12311
174. Fujigasaki, doi:10.1038/scitranslmed.aaf5660.


191. Lines, M.A.; Hébert, M.; McTaggart, K.E.; Flynn, S.J.; Tennant, M.T.; MacDonald, I.M. Electrophysiologic and phenotypic features of an autosomal cone-rod dystrophy caused by


218. Zeng, L.; Tallaksen-Greene, S.J.; Wang, B.; Albin, R.L.; Paulson, H.L. The deubiquitinating enzyme ataxin-3 does not modulate disease progression in a knock-in mouse


Spinocerebellar Ataxia Type 3 (SCA3) is a member of the family of polyglutamine (polyQ) neurodegenerative disorders that includes Huntington's Disease and several other SCAs. SCA3, the most common dominant ataxia in the world, is caused by polyQ tract expansion in the protein, ataxin-3. How SCA3 occurs and how to treat it remain unresolved issues. The primary culprit of toxicity in all polyQ diseases is the glutamine repeat: its abnormal expansion leads to neuronal dysfunction and death. With that said, there is indisputable evidence that the way polyQ-dependent toxicity presents—areas impacted, cellular processes perturbed—is predicated in large part on regions outside of the polyQ tract, i.e., protein context. Defining the role of non-polyQ regions of ataxin-3 in the toxicity of its expanded glutamine tract will increase our understanding of the biology of disease in SCA3. Functions including ubiquitin binding, proteasome shuttling, and aggregation depend heavily on these non-polyQ regions. The work highlighted in this dissertation expands on the current understanding of protein context in SCA3 by providing a detailed investigation into the importance of various ataxin-3 domains and interactions to disease
pathogenesis. It leans on the genetic versatility of the *Drosophila melanogaster* model organism to develop novel methods of studying the ataxin-3 interactome and its impact on polyQ expanded ataxin-3 toxicity. All this in hopes that it will help uncover potential therapeutic targets for this currently incurable disorder.
AUTOBIOGRAPHICAL STATEMENT

EDUCATION
Wayne State University
MI
Ph.D. Candidate in Pharmacology
Concentration: Molecular Neuropharmacology

University of Michigan
Ann Arbor, MI
B.S., Cellular and Molecular Biology

PUBLICATIONS


Full bibliography with undergraduate publications available at: