
Hunting for huntingtin associated factors: Identification and characterization of
Huntingtin expanded polyglutamine aggregate-associated factors and their impact on
Huntington disease model cellular toxicity.

by

Maggie P. Wear

Dissertation submitted to the Faculty of the
Molecular and Cellular Biology Graduate Program
Uniformed Services University of the Health Sciences
In partial fulfillment of the requirements for the degree of
Doctor of Philosophy 2016



FINAL EXAMINATION/PRIVATE DEFENSE FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN THE MOLECULAR AND CELL BIOLOGY GRADUATE PROGRAM

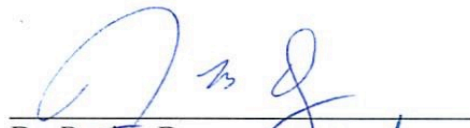
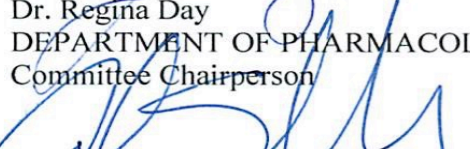
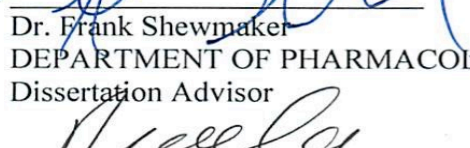
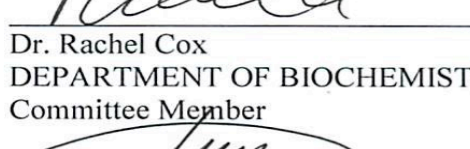

Name of Student: Margaret Wear

Date of Examination: February 17, 2016

Time: 1:00pm

Place: C2015

DECISION OF EXAMINATION COMMITTEE MEMBERS:

	PASS	FAIL
 Dr. Regina Day DEPARTMENT OF PHARMACOLOGY Committee Chairperson	<u>X</u>	—
 Dr. Frank Shewmaker DEPARTMENT OF PHARMACOLOGY Dissertation Advisor	<u>X</u>	—
 Dr. Rachel Cox DEPARTMENT OF BIOCHEMISTRY Committee Member	<u>X</u>	—
 Dr. John Wu DEPARTMENT OF OBSETETRICS AND GYNECOLOGY Committee Member	<u>X</u>	—
 Dr. Barrington Burnett DEPARTMENT OF ANATOMY, PHYSIOLOGY AND GENETICS Committee Member	<u>X</u>	—

ACKNOWLEDGMENTS

This work would not have been possible but for the grace of God. I have to express my gratitude to the Shewmaker lab – Frank Shewmaker and Dmitry Kryndushkin – for mentorship and assistance with all of the research in this dissertation. I also need to thank my committee members – Dr. Regina Day, Dr. Frank Shewmaker, Dr. Barrington Burnett, Dr. Rachel Cox, and Dr. John Wu – particularly for their assistance with formatting my research project, writing, and throughout the dissertation process.

I would like to thank Barrington Burnett and Christopher Grunseich for the PC12 Htt-polyQ cell lines and assistance working with them. Dr. Cara Olsen for her assistance with the statistical analysis. Drs. Aviva Symes, Dennis McDaniel, Jeff Stinson, and Rachel Cox provided invaluable assistance with fluorescence and confocal microscopy. The Snow lab – Dr. Andy Snow, Sasha Larsen, Dr. Gil Katz, and Dr. Jeff Stinson – provided vital support with training, performing, and interpreting the flow cytometry experiments. I must express my gratitude to my labmate Zach for his patience and invaluable conversations about science throughout this process. A special thanks to Dr. Burnett for stepping in and providing support and guidance throughout the writing process. I may not have been your student, but I don't think I could have done this without your help.

Personally I would like to thank my friends in science, Tracey, Jason, and Zach for their support throughout this process, scientific and otherwise. I would also like to thank my friends and classmates at USUHS, especially Dawn Patti for being my running

partner and helping me keep my stress under control. All of my friends deserve a prize for listening to me vent and keeping me happy, often with alcohol and chocolate – Thank you Katy, Kristen, Dallas, Grizzle, Bayley, Michelle, Erica, and all the rest. Thanks to my friends in fitness who have made going to the gym and wonderful social event. Michele, Danyelle, Krysta, and Bayley – you guys rock! Finally, a big thanks to my family, Jim, Judith, Jamie, Mary, Amy, and Claire for their support, even if they don't always understand what I am doing or why, they still support me with food and love. I could not have done this without each and every one of you. All together you allowed me to maintain some semblance of sanity throughout this process.

DEDICATION

This is dedicated to my grandfather, Vincent Paris, whose greatest wish was to have a doctor in the family. I wish you could be here for this day!

COPYRIGHT STATEMENT

The author hereby certifies that the use of any copyrighted material in the dissertation manuscript entitled: **Hunting for huntingtin associated factors:** Identification and characterization of huntingtin expanded polyglutamine aggregate-associated factors and their impact on Huntington disease model cellular toxicity is appropriately acknowledged and, beyond brief excerpts, is with the permission of the copyright owner.

[Signature] Maggie P. Wear

Maggie P. Wear

[May 20, 2016]

ABSTRACT

Hunting for huntingtin associated factors: Identification and characterization of huntingtin expanded polyglutamine aggregate-associated factors and their impact on Huntington disease model cellular toxicity.

Maggie P Wear, PhD 2016

Thesis directed by: Frank Shewmaker, PhD, Assistant Professor, Department of Pharmacology

Gene mutations resulting in the formation of insoluble protein aggregates, or amyloid-like structures, have been correlated with a number of pathological conditions. Huntington's disease (HD) is one of the most prevalent neurodegenerative diseases, characterized by movement, memory, behavioral, and cognitive difficulties, which become more severe as the disease progresses. Protein aggregation in HD is caused by expansion of the CAG repeat tract, also known as a polyglutamine (polyQ) expansion, in exon one of the Huntingtin protein (Htt). Research has shown that Huntingtin polyQ (Htt-polyQ) mutant expansions beyond 35 repeats result in protein aggregation. These Htt-polyQ aggregates form amyloid-like inclusions characterized by cross-beta structure and insolubility; and can result in loss of Huntingtin protein function or sequestration of essential cellular proteins that tightly interact with the aggregate. Thus, proteins interacting with Htt-polyQ aggregates may mediate disease pathogenesis by conferring *de novo* cytotoxicity or by contributing to a general loss of function. Determining how and

why specific proteins interact with polyQ aggregates is important to understanding disease mechanisms. We hypothesized that polyQ aggregates recruit and/or sequester proteins with common biophysical properties like size, charge, or conserved domains. To examine this, a mass spectrometry-based approach for non-targeted identification of intracellular amyloid-forming and amyloid-associated proteins was developed. We demonstrated that long intrinsically disordered (ID) domains – defined as domains of ≥ 100 amino acids that lack a defined structure – are a common biophysical property shared by proteins associated with Htt-polyQ aggregates. Deletional analysis was used to show that ID domains were required for association of two specific proteins, Sgt2, and FUS, with the Htt-polyQ aggregates. This research has allowed the unbiased identification of proteins associated with polyQ aggregates as well as the identification of ID domains as a feature required for the association of specific proteins with these aggregates. Recent studies show that careful analysis of ID domain conformational ensembles can reveal drug target sites within these unstructured regions. Thus, the development of drugs targeting ID domains may inhibit cellular protein interaction with Htt-polyQ aggregates, potentially reducing the neuron death associated with Huntington's disease.

TABLE OF CONTENTS

LIST OF TABLES	xi
LIST OF FIGURES	xii
LIST OF APPENDICES	xiii
CHAPTER 1: Introduction and Background	1
Huntington Disease	1
Protein Aggregation in Disease	3
Mechanisms of polyQ toxicity	6
Chapter 2: Proteins with intrinsically disordered domains are preferentially recruited to polyglutamine aggregates	12
Introduction	12
Results	14
Isolation & Analysis of Polyglutamine Aggregates in Yeast	14
Molecular functions of HttQ103-GFP aggregate-associated proteins in <i>S. cerevisiae</i> ...	16
Yeast proteins recruited into polyQ aggregates share common biophysical properties	19
Biochemical confirmation of protein recruitment to polyQ aggregates in yeast	22
Analysis of Htt polyglutamine aggregate-associated proteins in PC-12 cells	24
Molecular functions of HttQ74 aggregate-associated proteins from PC-12 cells	24
Biophysical properties of proteins that associate with polyQ aggregates in PC-12 cells	27
Neurodegenerative disease-linked proteins are recruited into polyQ aggregates	27
Confirmation of recruitment of identified proteins to polyQ aggregates by immunocytochemistry	30
Intrinsically-disordered domains play a role in the recruitment of proteins to polyQ aggregates	30
Discussion	31
Compositional Analysis of Polyglutamine Aggregates using TAPI	31
RNA binding proteins are disproportionately recruited to polyglutamine aggregates	34
Mitochondrial proteins are found in polyglutamine aggregates	35
Intrinsically disordered domains facilitate protein recruitment to Htt-polyQ aggregates	36
Why are neurodegenerative disease-associated proteins recruited to polyglutamine aggregates?	37
Mechanisms of Toxicity	38
CHAPTER 3: Discussion and overall conclusions	43
Mechanism of Htt-polyQ Toxicity	43
ID Domains in Htt-polyQ Aggregates	43
CORRELATION BETWEEN TAPI FINDINGS WITH OTHER STUDIES OF Htt-polyQ-ASSOCIATED PROTEINS	45
ID domains as Huntington Disease Therapeutic Targets	51
Conclusion	54

CHAPTER 4: Materials and methods	55
Cell Lines and Maintenance	55
Technique for Amyloid Purification and Identification (TAPI).....	55
Mass Spectrometry Analysis	57
Bioinformatic Analysis.....	59
Western Blotting	60
Lysate Partitioning	60
Confocal Microscopy.....	61
Thioflavin-T Analysis.....	61
Determination of overlapping Htt-polyQ associated proteins from multiple studies ...	62
APPENDICES	63
REFERENCES	70

LIST OF TABLES

Table 1: Neurodegenerative disease-associated amyloid-forming proteins.	11
Table 2: Molecular function of proteins associated with Htt-polyQ aggregates as identified by TAPI (N=52) in <i>Saccharomyces cerevisiae</i>	40
Table 3: Molecular function of proteins associated with Htt-polyQ aggregates as identified by TAPI (N=91) in <i>Rattus norvegicus</i>	40
Table 4: Htt-polyQ aggregate-associated proteins found in inclusions and aggregates in various neurodegenerative diseases.....	41
Table 5: Analogous proteins in yeast and rat associate with Htt-PolyQ aggregates.	42

LIST OF FIGURES

Figure 1: Technique for Amyloid Purification and Identification (TAPI).....	9
Figure 2: Polyglutamine-expanded Huntingtin exon 1 forms aggregates in yeast that can be isolated by TAPI.	15
Figure 3: RNA content of TAPI samples and Confirmation of aggregation.	17
Figure 4: Comparison of protein size and abundance.	18
Figure 5: Cellular proteins trapped with Htt polyQ aggregates are disproportionately composed of long intrinsically-disordered (ID) domains. ...	21
Figure 6: Western blotting confirms that TAPI-identified proteins are trapped in large, detergent-resistant Htt-polyQ aggregates.....	23
Figure 7: Proteasomal inhibition does not affect proteins trapped in Htt-polyQ aggregates.	25
Figure 8: Polyglutamine-expanded Huntingtin exon 1 forms aggregates in PC-12 cells that can be isolated by TAPI.	26
Figure 9: Confirmation of polyQ-associated proteins from PC-12 cells identified by TAPI.....	29
Figure 10: The ID domains of Sgt2p and Fus mediate their localization to Htt-polyQ aggregates in yeast cells.....	32
Figure 11: Meta-analysis identifies Htt-polyQ aggregate associated gene candidates.	48
Figure 12: Proteins that alter Htt-polyQ toxicity contain intrinsically disordered (ID) domains.....	51

LIST OF APPENDICES

Appendix 1: Supplemental File S1 - Proteins identified by mass spectrometry following TAPI purification of polyglutamine aggregates from yeast cells.....	63
Appendix 2: Supplemental File S2 - Proteins identified by mass spectrometry following TAPI purification of polyglutamine aggregates from rat cells.....	65
Appendix 3: Supplemental File S3 - PEARL-based algorithm for examining protein sequences for Q/N-rich regions.	66
Appendix 4: Supplemental File S4 - Proteins identified by meta-analysis of Htt-polyQ associated protein studies in yeast.	68

CHAPTER 1: Introduction and Background

HUNTINGTON DISEASE

Upwards of 50 disorders with a disparate symptomatology are correlated with protein aggregates (107). These protein aggregates occur as cytoplasmic, nuclear, or extracellular deposits and have been linked to disease pathogenicity (123; 124; 163; 167; 195). The proteinaceous deposits found in these disorders are primarily composed of a single, disease-linked protein, though they are often associated with other cellular proteins (53; 120; 127; 198; 226; 240). A subset of these disease-related aggregating proteins have been examined in depth, including A β (implicated in Alzheimer's disease), tau (Frontotemporal lobar degeneration, FTL), α -synuclein (Parkinson's disease), prion protein (PrP) (Transmissible spongiform encephalopathies, TSEs), and the polyglutamine (polyQ) proteins huntingtin (htt, implicated in Huntington Disease, HD) and ataxin-3 (Spinocerebellar ataxia 3, SCA3). Although we know that these proteins aggregate in neurodegenerative diseases, the role of the precursor aggregates, the end-stage visible aggregates and the aggregate-associated proteins in neuronal cell death remain unclear.

The intra- or extra-cellular accumulation of a number of neurodegenerative disease proteins is characterized by protein aggregates and selective neuronal death resulting in cognitive and behavioral changes (**Table 1**). One of the most prevalent of these is Huntington Disease. First described by George Huntington in 1872, HD is characterized by spasmodic movements of the limbs and facial muscles, which intensify with disease progression. In the US, HD affects 5-10 in every 100,000 individuals with

100% penetrance when the disease allele is present (191). With a cost of \$4,000 to \$30,000 per patient, per year, in direct medical costs and at least 300,000 patients currently suffering from HD in the US alone, care of these individuals exceeds \$5 billion per year (41; 46). The survival rate for HD patients is 10-20 years from symptom onset and with rising medical costs as the disease progresses, the average cost per patient is over \$250,000 over the course of the disease. These are simply direct medical costs and do not reflect the emotional and other costs to the patient and family. With no cure, treatments for HD exclusively address symptomatology and do not provide prolonged survival for the patients.

The pathology of HD has been linked to the huntingtin protein, specifically the amino-terminal polyglutamine (polyQ) region. Expansion of the glutamine coding CAG nucleotide triplet repeat beyond 40 repeats (referred to as Htt-polyQ) results in protein aggregation and neuronal cell death (167). Although many parts of the brain are affected in HD, the striatum, a subcortical region of the forebrain, experiences the most massive degeneration. The striatum is involved in movement, memory, and decision-making. In HD, the medium spiny neurons of the striatum are lost, resulting in movement difficulties along with cognitive and behavioral changes (35), with loss of up to 30% of overall brain weight over the course of the disease (71; 114). Disease progression is characterized by loss of neurons resulting in deterioration of physical, cognitive, and emotional faculties. Most HD patients ultimately succumb to pneumonia, or commit suicide (41; 166).

The mutated protein in HD, huntingtin, located on chromosome four, has been shown to interact with at least 100 proteins in the cell, but a distinct function for this protein remains elusive (67). The remarkable correlation between the length of the

polyglutamine repeat and disease severity has lead researchers to conclude that the polyQ aggregates are toxic and causative in the pathophysiology of the disease (32; 39; 131; 151; 177; 223; 227). Recent studies have examined the accumulation of aggregates in different types of neurons. Medium spiny projection neurons, those lost in HD, accumulate less polyQ aggregates than large interneurons, which are mostly spared in HD (71; 114). This has led to the speculation that aggregates are protective in these large interneurons and other misfolded Htt-polyQ species such as monomers or oligomers are responsible for the medium spiny neuron death (1; 74; 134). Both molecular chaperones and the ubiquitin proteasome system (UPS) are crucial to refolding or degrading misfolded proteins. Studies have shown that heat shock and UPS proteins are trapped in the insoluble Htt-polyQ inclusion bodies (223) and that these aggregates are ubiquitinated (39; 45). These are just a few of over 100 huntingtin-interacting proteins. A better understanding of other huntingtin-interacting proteins and their molecular pathways may be key to understanding the Htt-polyQ toxicity.

PROTEIN AGGREGATION IN DISEASE

As noted above, many neurodegenerative diseases are characterized by insoluble protein aggregates. These protein deposits were first observed in patients who died of systemic amyloidosis. The aggregates stained with iodine, which is used to detect starch; hence they were named ‘amyloid,’ or ‘starch-like’ (18; 187). Since that time, great advances have been made in characterizing the formation, biophysical properties, and disease linkage of protein aggregates.

There are two theoretical mechanisms of aggregate polymerization: (i) linear (aka isodesmic) and (ii) nucleation-elongation polymerization (57). The linear polymerization theory assumes that the dissociation constant for any monomer addition is identical. Whereas nucleation-elongation includes the formation of a ‘nucleus’ from several monomers; which then serves as a precursor structure increasing the favorability of monomer addition. It is generally accepted that amyloid forms by nucleated-elongation polymerization (1; 25; 29; 76; 88; 98; 108; 145; 173), although there are some that maintain there are many precursor states that lead to fibrillogenesis (11; 82; 91; 128; 172; 245). For any specific protein a number of precursor conformations may exist. These are referred to as prefibrillar oligomers, protofibrils, oligomers, or nuclei, and while the conformations differ, these terms are used somewhat interchangeably. Reports indicate that huntingtin forms precursor aggregates and that these species are key to inclusion body formation (30; 157; 222).

The toxic form of amyloid proteins has been a matter of some debate. Observations in the brain show that while many neurons contain aggregates, the specific types of neurons lost in HD do not contain observable aggregates (252). This has led researchers to speculate that aggregates or inclusions are simply the by-product of abnormal protein accumulation, or perhaps represent a protective event meant to sequester abnormal proteins (4; 169; 175; 188). Using time-lapse microscopy to monitor primary neurons for aggregation, the laboratory of Dr. S. Finkbeiner showed that neurons containing visible aggregates survived better than those without visible aggregates (4). Furthermore, fluorescence resonance energy transfer (FRET) experiments in cells overexpressing polyQ-GFP fusion protein showed that soluble oligomeric species appear

to be the most toxic isoform (201). Compounds which induce aggregate formation have been shown to reduce toxicity in polyQ cell models (13). Interestingly, apoptotic cells in a zebra fish model expressing expanded polyQ huntingtin did not contain aggregates (178). Together these studies strongly suggest that soluble Htt-polyQ oligomers are the toxic mediators of cell death, prior to formation of visible aggregates.

In contrast, a number of other studies suggest that the insoluble polyQ aggregates are, in fact, the toxic species. Examination of polyQ species from HD mouse and cell culture models using agarose gel electrophoresis indicates that oligomers can be detected prior to cell toxicity and that aggregate size correlates with toxicity (233). Increased ubiquitination, correlated with increased aggregate formation, also results in increased cell toxicity (42). It has been observed that aggregate formation leads to cell death in primary neurons (247) and PC12 immortalized neuronal cells, (239) but not in non-neuronal HEK293 cells, suggesting that the cell type-dependent toxicity may be in part the result of the mitotic state of the cells. In this model, cycling cells like HEK293 cells, are able to redistribute the aggregates between mother and daughter cells thereby reducing the toxic effects to any single cell. However, the majority of neuronal cells are mitotically inactive, therefore any aggregates that accumulate cannot be redistributed, resulting in subsequent cell death.

One recent study showed that properly folded polyQ monomers are α -helically folded and non-toxic, but that misfolded monomers or oligomers form β -sheets that correlate with cellular toxicity (142). These misfolded proteins, whether precursors or aggregated species, have a higher propensity to engage in aberrant interactions (193; 194; 203; 236). The huntingtin protein interacts with between 50 and 200 proteins forming a

myriad of heteroprotein complexes; the delicate balance of these complexes is likely altered by the polyQ expansion and subsequent misfolding of huntingtin protein.

Although a diverse collection of proteins can adopt amyloid conformation, the structure of the formed fiber is consistent, as described below (197; 199; 234). Many techniques have been used to determine the biophysical properties of amyloid including X-ray diffraction (53), Solid State-Nuclear Magnetic Resonance (NMR) (3), electron paramagnetic resonance spectroscopy (EPR) (206), circular dichroism (CD), Fourier transform infrared spectroscopy (FTIR), and electron microscopy (EM) (98). Amyloid structure features parallel in-register β -strands running perpendicular to the fibrillar axis. These β -strands become laminated β -sheets resulting in a distinctive cross- β rope-like structure when viewed by electron microscopy (EM) (194; 243). It is this β -rich structure of amyloid that results in strong resistance to degradation, detergents, proteolysis, and mechanical breakage (49). Amyloid exhibits distinct tinctorial properties that have been exploited to monitor amyloid formation in many disease models; specifically staining with thioflavins (Th) and Congo Red (CR) (83; 89; 197; 234).

MECHANISMS OF POLYQ TOXICITY

Although the genetic basis of HD and other polyQ diseases has been clear for many years, the molecular basis of the disease still remains unclear. The underlying Htt-polyQ toxic mechanisms predicted to involve many cellular pathways include transcription, intracellular transport, protein quality control, and mitochondrial function.

Transcriptional dysregulation is proposed to occur when misfolded Htt-polyQ accumulates in the nucleus of the neuron and sequesters the transcription factors

including CREB binding protein (CBP), TATA-binding protein (TBP), and specificity protein 1 (Sp1) (17; 21; 148). Although the mechanisms are very complex, the loss of CBP has been linked to reduced Brain-derived neurotrophic factor (BDNF) mRNA levels in HD neurons (252). Reduced BDNF levels have also been linked to altered axonal transport (252). Htt-polyQ inhibition of anterograde and retrograde axonal transport is polyQ length-dependent (70; 200) and it alters the activity of N-methyl-D-aspartate (NMDA) receptors leading to deregulation of Ca^{2+} homeostasis via an unclear mechanism (201; 202; 249).

Although mostly unexplored, mitochondrial dysfunction has also been correlated with Htt-polyQ. Mitochondria in the HD brain exhibit abnormal energy metabolism including decreased glucose metabolism, increased lactate concentration, and lower membrane potential resulting in decreased calcium necessary for mitochondrial permeability transition overload (121; 154).

One of the best-characterized molecular pathways implicated in polyQ diseases is the ubiquitin proteasome system (UPS). The UPS is responsible for degradation/recycling of misfolded or damaged proteins. Htt-polyQ is ubiquitinated, the first step in UPS recycling, indicating that the UPS recognizes the misfolded state (9; 45; 210). Both proteasome subunits and chaperone proteins are known to associate with huntingtin and have been found in Htt-polyQ aggregates (93; 174). While this indicates that the UPS is recruited to Htt-polyQ, it seems unable to process polyQ aggregates. The loss of function of the UPS results in the accumulation of misfolded proteins – both those associated with Htt-polyQ and other proteins within the cell – leading to neuronal cell death (9; 65; 156).

Association of huntingtin-interacting proteins with misfolded mutant huntingtin – be it monomeric, oligomeric, or in aggregate form – has been shown to sequester other cellular proteins away from their normal functions, leading to a loss-of-function toxicity (24; 40; 74). This phenomenon was specifically elucidated for BDNF in axonal transport described above. Furthermore, Gidalevitz and colleagues showed that co-expression of expanded polyQ with temperature-sensitive proteins in *C. elegans* resulted in the temperature-sensitive proteins being misfolded at a permissive temperature (65). This suggested that the misfolding of expanded polyQ proteins may induce misfolding of non-associated proteins.

It is clear that no single molecular mechanism can fully explain the neuronal death observed in HD. Therefore, identifying misfolded polyQ-associated proteins that mediate aggregation and toxicity will allow for improved understanding of HD pathogenesis. One major obstacle in identifying these proteins is the inherent difficulty in isolating Htt-polyQ aggregates. Our lab has developed a method – Technique for Amyloid Purification and Identification or TAPI – to purify amyloid aggregates in an unbiased way that allows for identification of aggregating proteins as well as proteins that are strongly associated with the amyloid (Figure 1). *We hypothesize that identification of proteins improved understanding of how these proteins contribute to the disease-associated cellular toxicity linked to Huntington disease.*

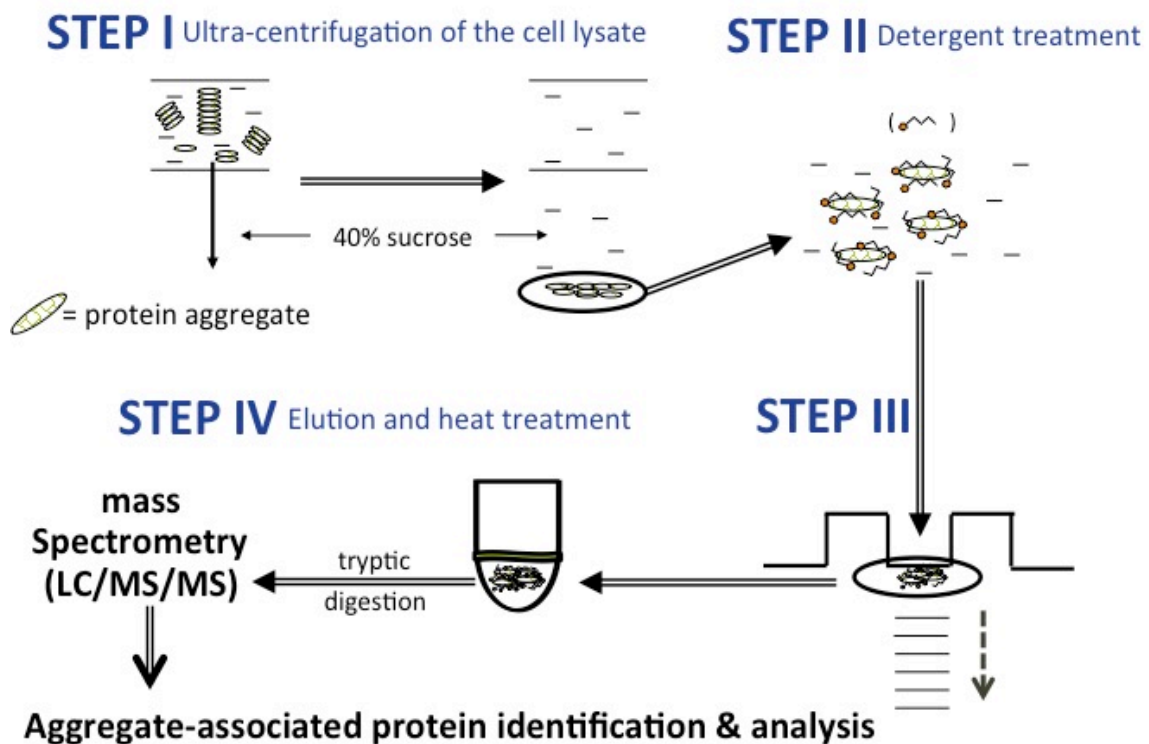


Figure 1: Technique for Amyloid Purification and Identification (TAPI).

The TAPI method utilizes sucrose pad ultra-centrifugation, detergent treatment and gel electrophoresis to purify aggregates from cell lysates in a stringent, unbiased manner. Coupling this purification with mass spectrometry allows for identification of the aggregating protein and proteins trapped in the aggregate.

Identification of mutant htt-associated proteins is important for defining the molecular pathways involved in HD cytotoxicity. Currently methods such as immunoprecipitation or deletion and overexpression screening are limited by protein bias and the requirement for *a priori* knowledge about targeted factors. The TAPI method is unbiased and requires no prior knowledge about the protein of interest (109; 231). Our current knowledge of Htt-polyQ protein associations has implicated a number of cellular pathways, but no conserved biophysical property or protein domain required for a protein to be trapped by Htt-polyQ aggregates has been identified.

Here, using the TAPI method we confirmed previously identified Htt interacting proteins, thus showing that purification of Htt-polyQ aggregates can be utilized to identify Htt-interacting proteins. We also identified and confirmed new Htt-interacting proteins and began exploring their role in HD cellular toxicity. Bioinformatic analysis of TAPI-identified Htt-interacting proteins also revealed disordered protein domains are common to the vast majority of Htt-interacting proteins. Deletion of the disordered domain of selected proteins eliminates Htt-polyQ aggregate interaction, suggesting that this domain may be necessary for cellular protein interaction with mutant huntingtin.

Table 1: Neurodegenerative disease-associated amyloid-forming proteins.

Neurodegenerative Disease	Amyloid-forming protein	Cells/Regions Affected
Alzheimer's Disease	Beta-amyloid	Cortex, hippocampus, basal forebrain, brain stem
Parkinson's disease	Alpha-synuclein	Substantia nigra, cortex, locus ceruleus, raphe, etc.
Transmissible spongiform encephalopathy (e.g. Bovine spongiform encephalopathy)	PrP ^{sc}	Cortex, thalamus, brain stem, cerebellum, other areas
Fatal Familial Insomnia	PrP ^{sc}	Cortex, thalamus, brain stem, cerebellum, other areas
Huntington Disease	Huntingtin	Striatum, basal ganglia, cortex, and other regions
Dentato-rubral and Pallido-Luysian atrophy (DRPLA)	Atrophin-1	Basal ganglia, brain stem, cerebellum, spinal cord
SCA1-7	Ataxin 1-7	Basal ganglia, brain stem, cerebellum, spinal cord
Spinal Bulbar Muscular Atrophy (SBMA)	Androgen Receptor	Basal ganglia, brain stem, cerebellum, spinal cord
Fronto-temporal dementia	Tau	Frontal and temporal cortex, hippocampus
Kuru	PrP ^{sc}	Cortex, thalamus, brain stem, cerebellum, other areas
Gerstmann-Straussler-Scheinker (GSS)	PrP ^{sc}	Cortex, thalamus, brain stem, cerebellum, other areas
Type II Diabetes	Islet amyloid polypeptide (IAPP)	Pancreatic cells
Familial amyloidotic cardiomyopathy (FAC)	Transthyretin	Cardiomyocytes
Sickle-cell disease	Hemoglobin (Hb)	Erythrocytes
Dialysis-related amyloidosis	Beta-2 microglobulin (B2M)	Extracellular affecting joints

Chapter 2: Proteins with intrinsically disordered domains are preferentially recruited to polyglutamine aggregates

INTRODUCTION

Accumulation of intracellular and extracellular protein aggregates is a common feature of multiple age-associated human disorders, particularly neurodegenerative diseases (167; 203). Amyloid is a highly-ordered aggregate that consists of polypeptides arranged in filamentous, beta sheet-rich structures with abundant interlocking hydrogen bonds between sheets (8; 130; 184; 209). Multiple proteins can adopt this architecture (220), and once established, they all share extraordinary resistance to proteolysis, chaotropic agents, detergents, and mechanical breakage (49; 109; 112; 126; 182). Amyloid aggregates – or their oligomeric precursors – are believed to contribute to cellular toxicity through a variety of mechanisms, such as physically disrupting membranes or sequestering essential heterologous proteins (16; 79).

Triplet CAG expansions resulting in polyglutamine (polyQ) tracts are characteristic of at least nine neurodegenerative diseases (143; 152), of which the most common is Huntington disease (HD). Proteins with expanded tracts of polyQ, which are natively unstructured, are predisposed to adopt conformations with amyloid-like properties (23; 176). When polyQ tracts of the Huntingtin protein (Htt) exceed ~35 glutamines, Htt fragments can form intracellular inclusions (39; 69; 176; 186) resulting in impaired Htt function and aberrant protein interactions (40). These inclusions (or their early precursors) may confer a dominant gain-of-function cellular toxicity (86; 170). Other studies suggest that aggregation caused by polyQ expansion can result in loss of

protein function (and thus pathology), either directly via functional loss of the aggregating species (78), or from the sequestration of proteins that tightly interact with the aggregate (155; 243). Thus, the proteins that interact with aggregates may mediate pathological mechanisms, either by conferring new cytotoxicity or by contributing to a general loss of function. For these reasons, determining how and why specific proteins interact with polyQ aggregates is important to understanding – and ultimately combating – disease mechanisms.

Recently, we established a mass spectrometry-based approach for non-targeted identification of intracellular amyloid-forming and amyloid-associated proteins (109; 110). Our method, called TAPI (Technique for Amyloid Purification and Identification), exploits the biophysical characteristics of amyloid, namely detergent resistance and high molecular weight, to isolate amyloid aggregates from cell lysates. Stringent nuclease treatment followed by SDS-gel electrophoresis eliminates non-specific or loosely associated proteins. Thus the TAPI protocol differs from antibody-based pull-down protocols by limiting positive hits to those most tightly associated with amyloid-like aggregates.

In this study we applied the TAPI protocol coupled with mass spectrometry to aggregates formed by polyQ-expanded huntingtin fragments in both yeast (*S. cerevisiae*) and mammalian cells (PC-12, rat neuronal precursor). Previously, various proteins have been shown to interact with Htt or Htt fragments (75; 97; 118; 155; 208; 223), but our approach was designed to identify proteins that are directly trapped within the amyloid-like polyQ aggregates to determine the types of proteins most prone to irreversible inclusion. We hypothesized that these aggregates would recruit and/or sequester proteins

with common biophysical properties. We observed that inclusion into polyQ aggregates was mediated by long intrinsically-disordered (ID) protein domains (≥ 100 amino acids) in two evolutionary divergent cell models. Also, many proteins normally associated with neuronal aggregation in other degenerative diseases (especially amyotrophic lateral sclerosis (ALS)) were disproportionately recruited into polyQ aggregates in mammalian cells. This study expands the emerging connection between ID domains and neurodegenerative disease (213), and demonstrates that long ID domains predispose proteins to be recruited into amyloid-like aggregates.

RESULTS

Isolation & Analysis of Polyglutamine Aggregates in Yeast

To identify aggregate-associated proteins, human HttQ103-GFP and HttQ25-GFP (134) were expressed under control of a galactose-inducible (GAL1) promoter in the yeast *Saccharomyces cerevisiae* (Figure 2). Both expression constructs contain the human huntingtin (Htt) exon 1 fragment with polyQ tracts (103 or 25 glutamines, respectively) fused in frame with green fluorescent protein (GFP). As observed previously, HttQ25-GFP is soluble during expression, whereas HttQ103-GFP forms toxic cytoplasmic SDS-resistant aggregates (134) (Figure 2A) that have amyloid-like tinctorial properties and can be trapped at the top of an SDS acrylamide gel (109) (Figure 2B; Figure 3B). Proteins that are specifically associated with Htt amyloid-like aggregates were isolated using the TAPI method (109; 110), which traps the large detergent-resistant species in acrylamide gel matrix for subsequent extraction and identification. As demonstrated in Figure 2B, the TAPI method isolates amyloid-like aggregates of HttQ103-GFP, whereas the non-

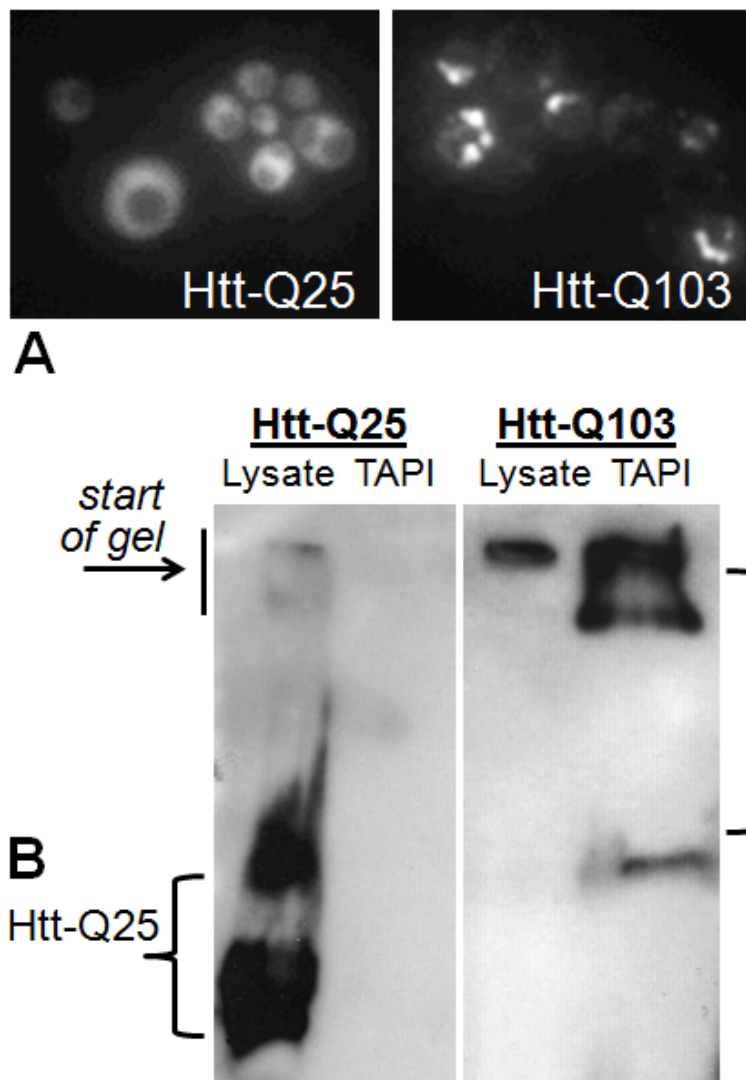


Figure 2: Polyglutamine-expanded Huntingtin exon 1 forms aggregates in yeast that can be isolated by TAPI.

(A) Fluorescence microscopy of yeast expressing GFP-tagged Huntingtin exon 1 (Htt) with a normal (Q25) or an expanded polyglutamine tract (Q103). (B) Western blot of GFP-Htt-Q25 and GFP-Htt-Q103 showing that high molecular weight aggregates can be isolated from Htt-Q103 expressing yeast cells. Lysate = input; TAPI = purified aggregate.

aggregate-forming HttQ25-GFP does not form species large enough, or sufficiently detergent-resistant, for isolation.

Proteins tightly associated with the isolated Htt-polyQ aggregates were identified using tandem mass spectrometry (MS/MS). Qualitative comparison of all identified proteins from the HttQ103-GFP samples – relative to the HttQ25-GFP samples – reveals a subset that is only associated with the large polyQ aggregates (Table 1; Appendix 1). In total, 52 proteins were considered polyQ-associated because they were reproducibly found in the expanded HttQ103 aggregate (Table 1) while absent in the HttQ25 control sample (described in methods). To confirm that our approach is not enriching for abundant, large, or charged proteins, the polyQ aggregate-associated proteins were compared against the entire yeast proteome. No obvious differences in size distribution, abundance ((122); Figure 4.), or charge (the avg. pI of TAPI-identified proteins is 7.1, the same as the approximation for the yeast proteome (180)) were observed between our TAPI-identified proteins and that of the entire yeast proteome.

Molecular functions of HttQ103-GFP aggregate-associated proteins in *S. cerevisiae*

HttQ103-GFP aggregate-associated proteins were examined using gene ontology (GO) and *Saccharomyces* genome database (SGD) to assign their functions and properties (5; 6; 28; 117). Unexpectedly, RNA/DNA-binding (mostly RNA binding; Table 1) proteins make up the largest percentage of HttQ103-GFP aggregate-associated proteins. In fact, more than 1/3rd of the polyQ-associated proteins are specifically characterized as RNA-binding proteins (RBPs). Previous studies suggest that

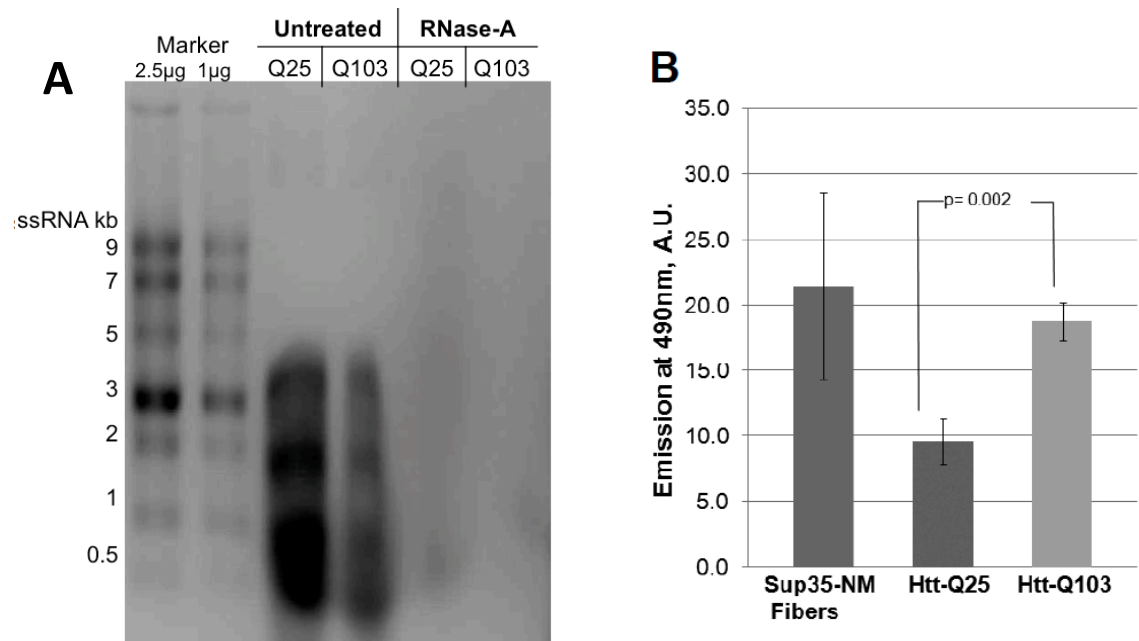


Figure 3: RNA content of TAPI samples and Confirmation of aggregation.

(A) Efficacy of RNase treatment in TAPI procedure. Yeast cell lysates from Htt-Q103-GFP or Htt-Q25-GFP expressing cells were put through the TAPI procedure (see methods and materials Chapter 4) with (lanes 5 and 6) or without (lanes 3 and 4) the addition of 3 mg RNase A enzyme. Samples were run with purified RNA standards (Invitrogen, lanes 1 and 2) at indicated concentrations on a 0.8% agarose gel in RNA-free conditions and (B) Th-T fluorescence of crude aggregates isolated from yeast expressing Htt-Q103-GFP or HttQ25-GFP. The average of fluorescence emission at 490nm over three experiments were compared to control purified prion forming domain (NM) of yeast prion protein Sup35 using a two-tailed t-test to determine statistical significance. Error bars were created using standard deviation.

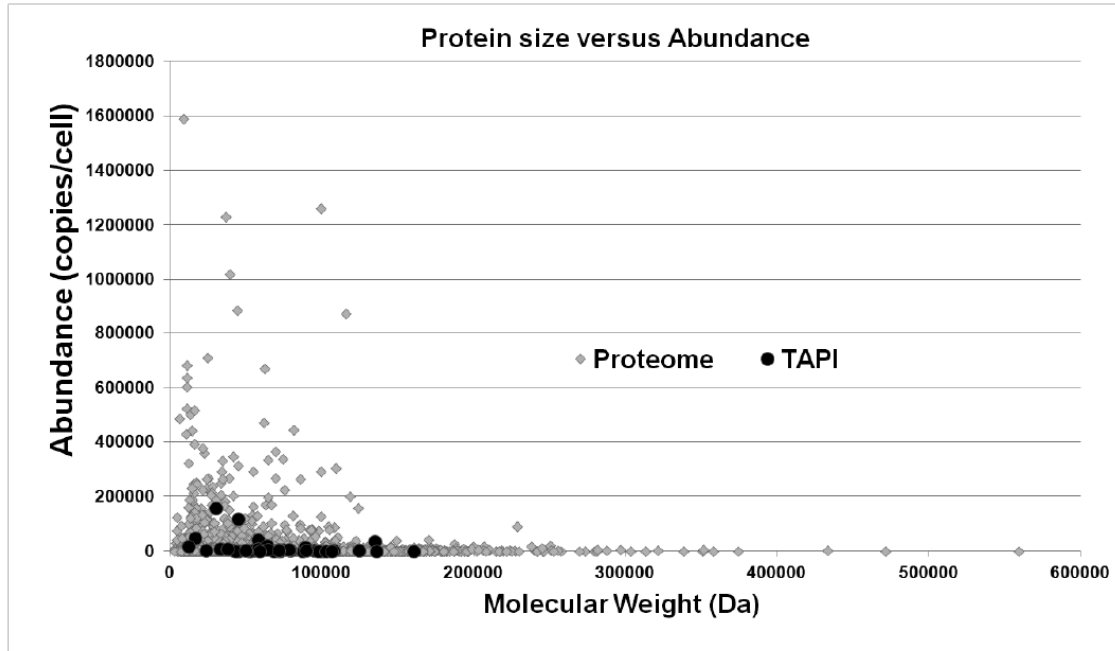


Figure 4: Comparison of protein size and abundance.

The yeast proteome is compared with the 52 proteins identified by TAPI as being tightly associated with Htt-Q103-GFP. The proteins found to co-aggregate with polyQ were not obviously larger or more abundant than the yeast proteome in general (TAPI avg. 1.2×10^4 copies/cell, Proteome avg. 1.3×10^4 copies/cell). TAPI does not enrich for disproportionately large proteins (Avg TAPI protein size 72kDa, proteome avg 53kDa). Protein abundance values accumulated from western blotting (64), GFP expression (144), 2D gel analysis (62), and APEX analysis (122).

RBPs may localize to aggregates because RNA co-aggregates with amyloid-forming proteins (158). However the TAPI method involves extensive RNase treatment (109; 110); while we cannot conclude that RNA is completely absent, the vast majority of RNA is eliminated prior to aggregate isolation (Figure 3A). As most RNA-dependent interactions should be lost, a preponderance of RBPs in the HttQ103-GFP aggregate is likely independent of RNA-mediated interactions.

Yeast proteins recruited into polyQ aggregates share common biophysical properties

PolyQ aggregates have been proposed to induce heterologous protein misfolding via a “cross-seeding” mechanism (60). Previously, Michelitsch and Weissman observed that Q/N-rich domains have an increased propensity to adopt amyloid structure and concluded that 30 glutamine and/or asparagine residues within an 80-amino acid stretch served as a good predictor of amyloid formation (137). Assuming that the polyQ aggregates could “cross-seed” such Q/N-rich proteins, we analyzed the 52 identified proteins and found they are significantly more likely to have Q/N-rich domains than the whole of the yeast proteome (54% vs. 2% (137), respectively; Appendix 1). For comparison, if simply measuring for total glutamine content of the identified proteins, the HttQ103-GFP aggregate-associated proteins had only a two-fold greater total percentage of glutamine content relative to the yeast proteome (respectively, 9.7% vs. 3.8% (99); Appendix 1).

An enrichment of Q/N-rich segments also implies an increase in intrinsically unstructured protein domains. When each of the polyQ aggregate-associated proteins was examined for global intrinsic disorder (described here: (230)), the aggregate-associated proteins showed a higher average total percent intrinsic disorder relative to the average

value for the whole yeast proteome (48% vs. 20%, respectively). However, if the identified proteins are specifically analyzed for containing discrete regions that are intrinsically disordered (using IUPred-L prediction algorithm described in the methods (50)), the polyQ-associated proteins are strongly enriched for the presence of an ID domain (Figure 5). Previous studies have classified ID domains as unstructured regions greater than 20-40 amino acids long (205; 221). For the yeast proteins associated with polyQ aggregates, almost all have an ID domain of at least ≥ 30 amino acids in length (92% vs. 31% of the proteome control, respectively; Appendix 1). However, $\sim 2/3^{\text{rds}}$ of the proteins contain very long ID domains of ≥ 100 amino acids (63% versus 9%; Figure 5), and strikingly of these proteins, $\sim 1/3^{\text{rd}}$ contain no Q/N-rich domain.

ID domains facilitate protein interactions with RNA, so the large cohort of RBPs we found associated with polyQ aggregates could simply be a result of these proteins disproportionately possessing ID domains. The RBP subset of aggregate-associated yeast proteins was compared to all putative and characterized RBPs in the yeast proteome for ID content (Appendix 1). A majority (70%) of the HttQ103-GFP aggregate-associated RBPs contain ID domains of ≥ 100 amino acids, while RBPs in general rarely have such long ID domains ($\sim 17\%$; Appendix 1). Thus, the high frequency of RBPs in the polyQ aggregates might result from the presence of long ID domains in these proteins, rather than a result of some uncharacterized RNA-binding mechanism of polyQ aggregates.

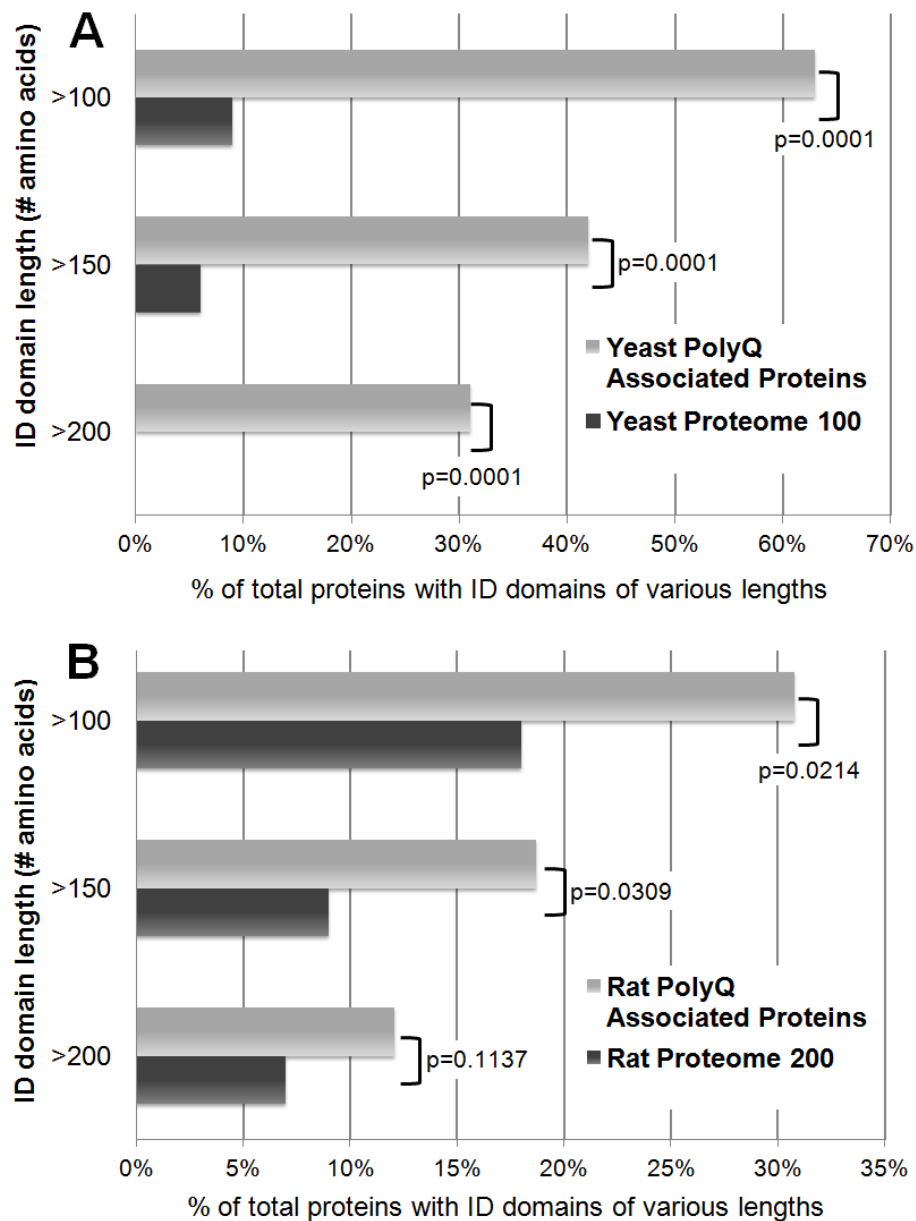


Figure 5: Cellular proteins trapped with Htt polyQ aggregates are disproportionately composed of long intrinsically-disordered (ID) domains.

(A) Comparisons of the percentages of proteins with long ID domains of the 52 yeast proteins in Table 2 (reproducibly found by TAPI to be tightly associated with Htt-Q103-GFP aggregates) versus 100 randomly-selected yeast proteins (Appendix 1). Most of the identified proteins have long ID domains of at least 100 amino acids. (B) Comparisons of the percentages of proteins with long ID domains of the 91 rat proteins in Table 3 (reproducibly found by TAPI to be tightly associated with GFP-Htt-Q74 aggregates) versus 200 randomly-selected rat proteins (Appendix 2). ID domains are defined as regions of 30 or more amino acids with IUPred scores of 0.5 or greater (50; 51). Chi-Square Fisher's Exact test (Graphpad software) was used to determine significance between TAPI-identified proteins and proteome control sets.

Biochemical confirmation of protein recruitment to polyQ aggregates in yeast

Def1p, Ent2p, Sgt2p and Bmh1p are among the proteins we found to be specific to polyQ aggregates in yeast. We also found that mammalian homologs of Ent2p, Sgt2p and Bmh1p (CLINT1, SGTA and YWHAB, respectively) co-aggregate with polyQ in mammalian cells (discussed in detail below). Ent2p, Sgt2p and Bmh1p were previously shown to have effects on protein aggregation (or toxicity) in yeast models (52; 103; 229), while Def1p has not been shown to influence protein aggregation. We selected Def1p, Ent2p, Sgt2p and Bmh1p for biochemical confirmation of the MS results.

The presence of Def1p, Ent2p, Sgt2p and Bmh1p in polyQ aggregates was tested by immunoblotting following a modified version of our TAPI protocol (Figure 6). HttQ103-GFP, expressed in yeast, forms a high molecular weight aggregate that partitions to the pellet fraction and gets stuck at the top of an SDS-PAGE gel (Figure 6A). When HA-tagged Def1p, Ent2p, Sgt2p and Bmh1p are co-expressed with HttQ103-GFP, they show a similar pattern in their respective western blots, but only when expressed with the long polyQ expansion, not with HttQ25-GFP (Figure 6B). Thus, the interactions of all three proteins with polyQ aggregates are sufficiently strong that they co-fractionate and withstand the conditions of SDS-PAGE, resulting in their retention in the large resistant species that cannot migrate into the gel (Figure 6B). The His3 protein was chosen as a negative control as it was never identified in our TAPI samples. Immunoblotting confirms that unlike Def1p, Sgt2p and Bmh1p, HA-tagged His3p is not entangled within polyQ aggregates (Figure 6B), thus recapitulating our MS results biochemically. To ensure that proteins were not independently forming large detergent-resistant aggregates as a consequence of cellular stress caused by HttQ103-GFP, cells

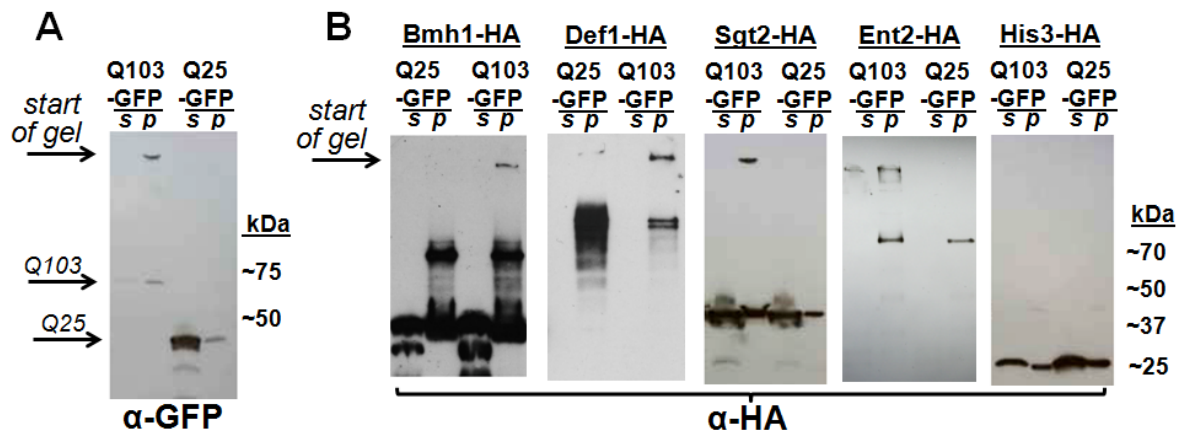


Figure 6: Western blotting confirms that TAPI-identified proteins are trapped in large, detergent-resistant Htt-polyQ aggregates.

(A) Immunoblotting reveals that Htt-Q103-GFP, but not Htt-Q25-GFP, forms large detergent-resistant aggregates that fractionate in the pellet (partial TAPI purification; see methods) and remain at the top of an acrylamide gel under SDS-PAGE conditions. (B) Immunoblotting confirms that HA-tagged Bmh1p, Def1p, Ent2p, and Sgt2p (proteins identified by mass spec) get trapped in the detergent-resistant aggregates that can be seen stuck at the top of the gels in the pellet fractions. As a negative control, HA-tagged His3p (not identified by mass spec) shows no susceptibility to co-aggregation with Htt-Q103-GFP. Note that Def1p was not easily visualized in the supernatant fraction because it is prone to degradation (data not shown). Samples were spun at 45,000 rpm, except Ent2p (10,000 rpm). **S** = supernatant; **P** = pellet fraction.

were treated with two alternative stresses: proteasome inhibitor MG132 and over-expressed human α -synuclein protein, which is toxic in yeast cells(153). Neither condition resulted in Sgt2p getting stuck at the top of an acrylamide gel (Figure 7.)

Analysis of Htt polyglutamine aggregate-associated proteins in PC-12 cells

With the observation that RBPs and proteins with ID or Q/N-rich domains are bound to HttQ103-GFP aggregates in the yeast model, we asked if the same would be true in a mammalian system. To test this, another Huntington disease model was used: mammalian PC-12 cells stably expressing doxycycline-inducible HttQ23-GFP or HttQ74-GFP developed by David Rubinsztein's lab (239). Again, both constructs contain the Htt exon 1 fragment with a polyQ tract (23 or 74) fused in frame with GFP, and only the protein with pathogenic extended polyQ forms SDS-resistant aggregates (Figure 8A). In this model, the polyQ aggregates have an approximately equal distribution in the cytoplasm and nucleus(239), thus may interact with most of the non-secreted cellular proteome. The amyloid-forming HttQ74-GFP, but not HttQ23-GFP, could be successfully purified and detected using the TAPI method, as shown in Figure 8B. MS analysis followed by comparison of all identified proteins showed a subset that was unique to samples with polyQ aggregates. Using the criteria described above, 91 proteins were considered specific to the HttQ74-GFP aggregates (Table 3; Appendix 2).

Molecular functions of HttQ74 aggregate-associated proteins from PC-12 cells

Characterization of the proteins enriched in polyQ aggregates from PC-12 cells reveals a disproportionate number of RBPs, as similarly observed in yeast (Table 3; Appendix 1 & 2). Also, several functional homologs were common to the polyQ aggregates isolated

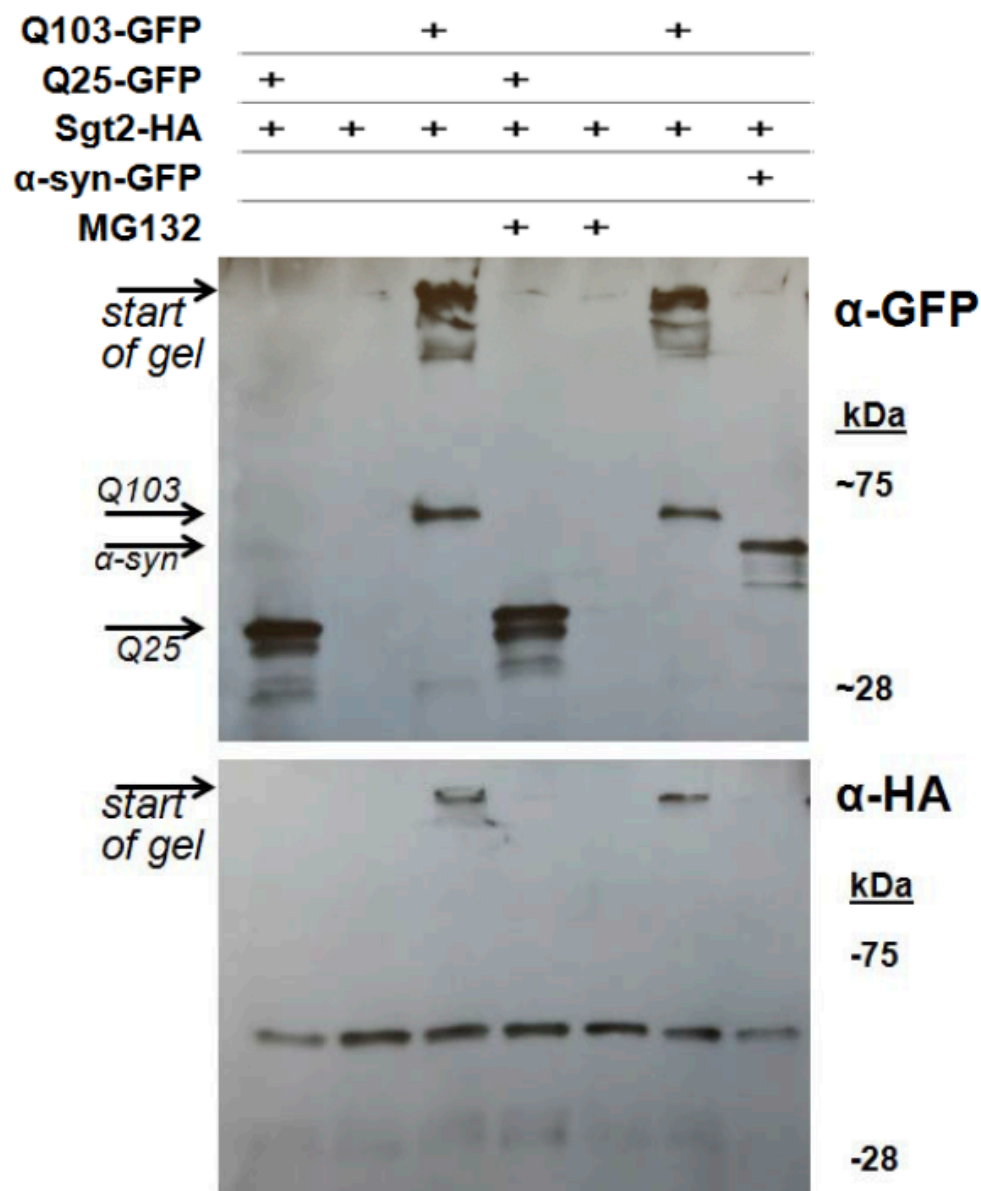


Figure 7: Proteasomal inhibition does not affect proteins trapped in Htt-polyQ aggregates.

Proteasomal inhibition or proteo-toxic stress are not sufficient to cause Sgt2p to be trapped in (or form) detergent-resistant high-molecular weight aggregates. Neither the overnight exposure (~16 hours) of yeast cells to 10 μ M MG-132 nor the overnight expression of human alpha-synuclein altered the electrophoretic migration of yeast Sgt2p. However, the accumulation of Q103-GFP into a large detergent-resistant species was sufficient to affect Sgt2p's migration into an acrylamide gel.

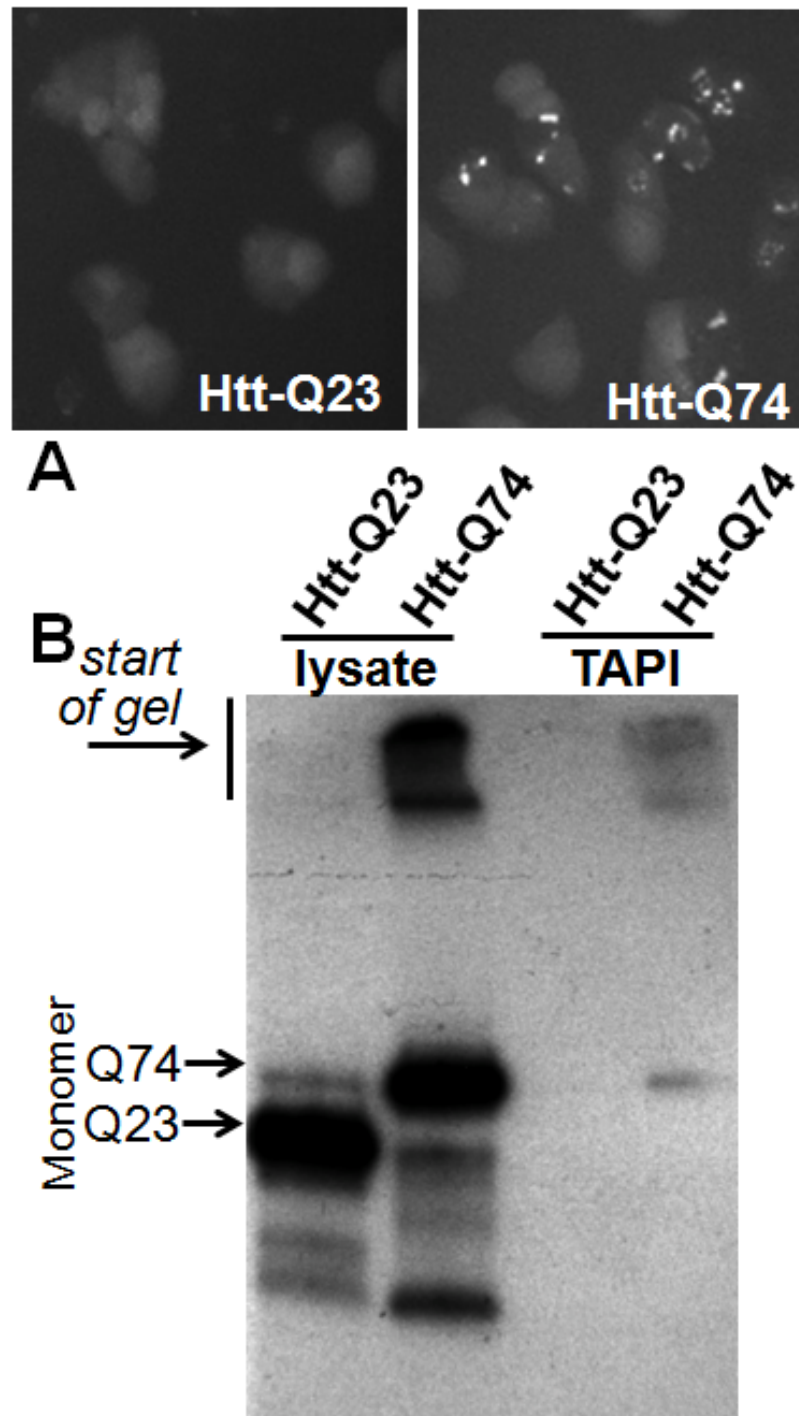


Figure 8: Polyglutamine-expanded Huntingtin exon 1 forms aggregates in PC-12 cells that can be isolated by TAPI.

(A) Fluorescence microscopy of PC-12 cells expressing doxycycline inducible transgene GFP-tagged Huntingtin exon 1 (Htt) with normal (Q23) or expanded polyglutamine tract (Q74). (B) Western blot of GFP-Htt-Q23 and GFP-Htt-Q74 showing high molecular weight aggregates can be isolated from Htt-Q74 expressing PC-12 cells. Lysate = input; TAPI = purified aggregates.

from both the yeast and mammalian cells (Table 5): RNA-binding proteins DDX5 (yeast Dhh1p) and hnRNPA3 (yeast Hrp1p), 14-3-3 proteins YWHAB and SFN (yeast Bmh1p), endocytosis proteins CLINT1 (yeast Ent1/2p) and AAK1 (yeast Akl1p), and chaperone proteins SGTA (yeast GET pathway protein Sgt2p), DNAJA2 and DNAJA4 (yeast Ydj1p and Apj1p).

Biophysical properties of proteins that associate with polyQ aggregates in PC-12 cells

Biophysical characterization reveals the HttQ74-GFP aggregate-associated proteins from PC-12 cells are enriched for Q/N-rich regions relative to the entire rat proteome (7% versus 0.4%), albeit to a much lesser degree than in yeast. Proteins containing long ID domains (≥ 100 amino acids) are significantly increased among the TAPI-identified HttQ74-GFP aggregate-associated proteins (31% vs. 18% for proteome control, respectively, $p=0.021$; Figure 5; Appendix 2). As in yeast, this suggests that cellular proteins with long ID domains may be inherently prone to inclusion in polyQ aggregates.

Neurodegenerative disease-linked proteins are recruited into polyQ aggregates

Among the aggregate-specific proteins in PC-12 cells, we also identified a significant subset of proteins that are neurodegenerative disease-associated (Table 4). Surprisingly, these proteins were not limited to huntingtin-interacting proteins; we identified a cadre of ALS-linked proteins. We hypothesized polyQ aggregates could pull in proteins that are prone to aggregation in other pathological contexts. When we probed the purified polyQ aggregates (purified fraction confirmed in Figure 9A) with antibodies specific to proteins that are known to aggregate in the motor neurons of ALS patients

(and were identified by MS in this study), we corroborated the specific presence of FUS, TDP-43, and UBQLN2 in the Htt-Q74 aggregates (Figure 9B). We hypothesized that other ALS-linked proteins may be similarly recruited into aggregates but escaped detection by MS. Immunoprobng for the ALS-linked HNRPA1 also revealed its presence in the purified polyQ aggregates (Figure 9B). For a control, we probed for the kinase Erk, which was not identified by MS in our samples, and indeed, it could not be found in the polyQ aggregates (Figure 9A). In total, of the HttQ74-GFP aggregate-associated proteins in PC-12 cells, 21% (19/91) have previously been found in the intraneuronal inclusions of various neurodegenerative diseases (Table 4). Of these disease-linked HttQ74-GFP aggregate-associated proteins, many are RBPs (7/19) and more than half (9/19) contain very long ID domains (≥ 100 amino acids) (Table 4; Appendix 2).

Since one fifth of the proteins we identified have been previously linked to pathological aggregates, perhaps many of the remaining proteins represent novel aggregate-associated proteins that have never been specifically probed in various pathological contexts. We selected HSPA8, CLINT1 and SGTA as candidate proteins that could potentially be recruited into pathological aggregates in neurodegenerative disease. We chose CLINT1 and SGTA because intriguingly their homologs (Sgt2p and Ent2p, respectively) were also identified in our yeast model. The Hsp70 protein HSPA8 was selected because Hsp70s have been previously suspected to play critical roles in neurodegenerative disease (14; 133; 138). We confirmed by immuno-blotting the presence of all three proteins in the highly-purified Htt-Q74-GFP aggregates from PC-12 cells (Figure 9B).

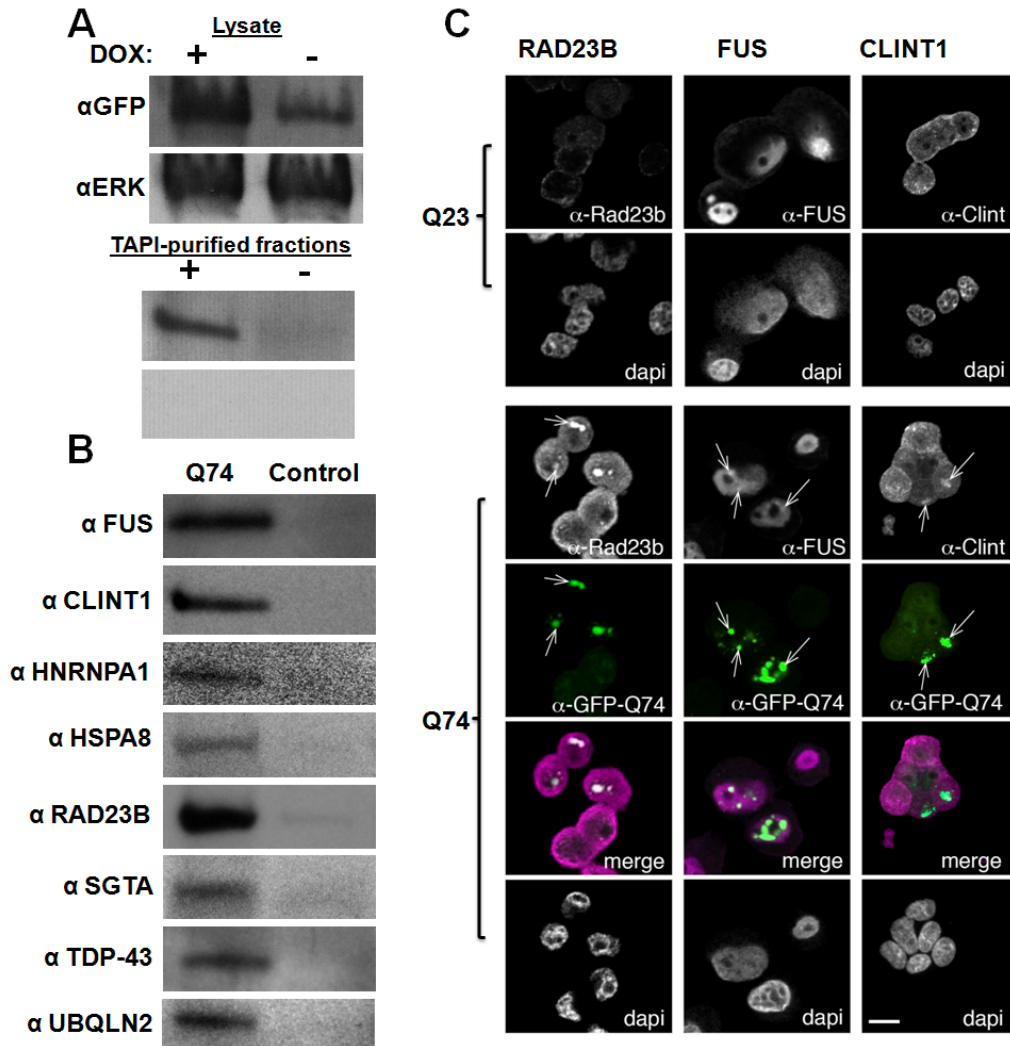


Figure 9: Confirmation of polyQ-associated proteins from PC-12 cells identified by TAPI.

(A) Western blotting shows that the addition of doxycycline to the PC-12 cell model induces the expression of HttQ74-GFP, resulting in aggregates that can be purified by TAPI. The kinase ERK is probed as a negative control; ERK was never identified by mass spectrometry, so is not expected to co-fractionate with polyQ aggregates.

(B) Western blot analysis of TAPI-purified polyQ aggregates from PC-12 cells confirms the presence of several disease-associated proteins only in the Htt-Q74 samples. All proteins migrated near their predicted molecular weights. For control, the TAPI procedure was conducted in parallel on the induced Htt-Q23 cell line (FUS, TDP-43, UBQLN2, HNRNPA1) or the un-induced Htt-Q74 cell line (CLINT1, HSPA8, RAD23B, SGTA). (C) Confocal microscopy shows localization of identified proteins to Htt-Q74 aggregates in PC12 cells. (left) RAD23B, nominally a DNA repair protein, localizes to nuclear Htt-Q74 inclusions but not cytoplasmic inclusions. (middle) FUS, an RNA-binding protein localizes to nuclear and cytoplasmic Htt-Q74 inclusions. (right) CLINT1, a clatherin-interacting protein, is observed in cytoplasmic Htt-Q74 aggregates. Arrows indicate foci with co-localized proteins. Green = GFP; Magenta = CLINT1, FUS or RAD23B in merge.

Confirmation of recruitment of identified proteins to polyQ aggregates by immunocytochemistry

The localization of selected polyQ aggregate-associated proteins was also confirmed by confocal microscopy (Figure 9C). The proteins CLINT1, RAD23B and FUS were selected because they appeared to be particularly strong hits based on the initial analysis of the polyQ mass spectrometry data and Western blot results. Using immunocytochemistry, each was observed to be aberrantly recruited to the major sites of HttQ74-GFP aggregation (Figure 9C). However, despite strong over-expression of HttQ74-GFP and the formation of large intracellular aggregates, the recruited proteins did not appear to completely localize to the aggregates. In fact, only a fraction of each protein's respective total was found at the aggregate.

Intrinsically-disordered domains play a role in the recruitment of proteins to polyQ aggregates

We selected yeast Sgt2p and human FUS to further examine the role of ID domains in localization to Htt-polyQ aggregates. Sgt2p is involved in protein quality control and does not contain a known RNA-binding domain (RBD). Interestingly, we also identified Sgt2p's mammalian homolog, SGTA, among the proteins associated with polyQ aggregates in PC-12 cells. The FUS protein contains a distinct RBD and a long N-terminal ID domain, and when expressed in yeast, exhibits aggregation and toxicity reminiscent of what is observed in diseased motor neurons (61; 95; 111; 196).

To determine the contribution of their respective ID domains toward recruitment to polyQ aggregates, we created expression vectors in which the major ID domain (as determined by IUPred-L) of both FUS and Sgt2p is deleted ($\text{FUS}\Delta\text{ID} = \text{FUS}\Delta^{1-134}$; $\text{Sgt2}\Delta\text{ID} = \text{Sgt2}\Delta^{300-346}$). The full-length versions of FUS and Sgt2p, or their ΔID

counterparts, were co-expressed with either HttQ25-GFP or HttQ103-GFP. We truncated the TAPI protocol to easily evaluate the co-localization of the proteins with polyQ aggregates (lysate partitioning; see methods). When we isolated the Htt-polyQ aggregates by lysate partitioning from the Sgt2-transformed strains (shown in Figure 10A as the resistant species stuck at the top of the gel), we observed an enrichment of full-length Sgt2p in the Htt-polyQ high molecular weight aggregates (Figure 10B, left panel). However, when the major ID domain was deleted, most of the co-localization with the polyQ aggregate is eliminated (Figure 10B, right panel). The exact same pattern was observed for FUS and FUS Δ ID (Figures 10C and 10D, respectively). For comparison, we also used an engineered variant of FUS, which has leucines in place of four conserved phenylalanines (FUS(4F-L): amino acids 305, 341, 359 and 368) in the RNA recognition motif (RRM) domain. FUS(4F-L) was previously shown to be RNA-binding incompetent (38). We observed that FUS(4F-L) was recruited to polyQ aggregates as readily as wild-type FUS (Figure 10E, F), thus suggesting that RNA binding may not play a significant role in facilitating a protein's inclusion into polyQ aggregates.

DISCUSSION

Compositional Analysis of Polyglutamine Aggregates using TAPI

Models of huntingtin exon 1 mimic truncated versions of huntingtin found in intraneuronal aggregates (131), and thus are helpful for studying intracellular aggregation. Here we analyze the protein species that get recruited into amyloid-like aggregates formed by polyQ-expanded Huntingtin exon 1 in both yeast (Q103) and mammalian cells (Q74). The distinctive tinctorial properties and detergent resistance of

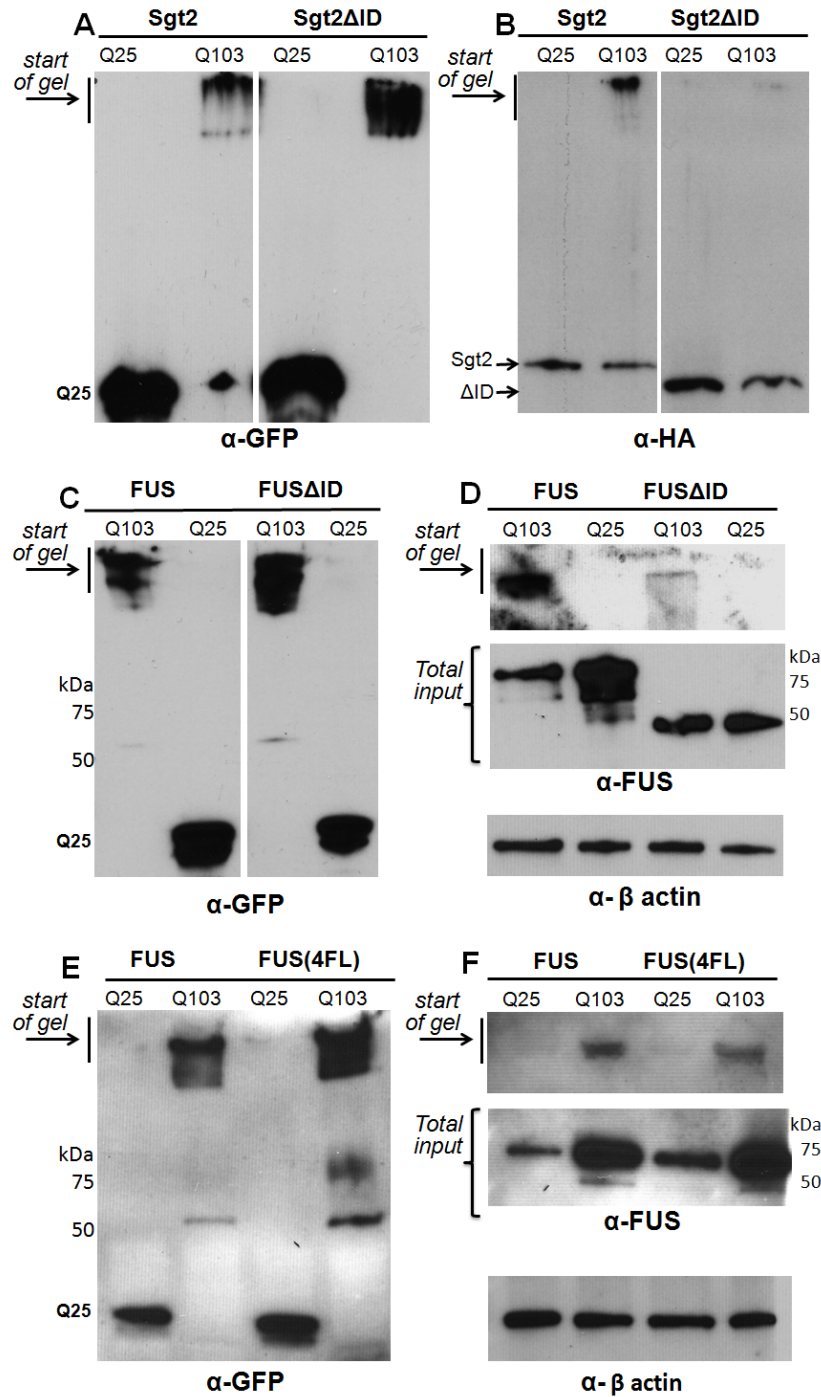


Figure 10: The ID domains of Sgt2p and Fus mediate their localization to Htt-polyQ aggregates in yeast cells.

(A, B) Western blots of lysates from yeast strain W303 expressing HttQ25-GFP or HttQ103-GFP in combination with HA-tagged Sgt2p or Sgt2ΔID (A = αGFP; B = αHA). (C, D) Western blots of cells expressing HttQ25-GFP or HttQ103-GFP in combination with FUS or FUSΔID (C = αGFP; D = FUS & α-β actin). Because FUS is quickly degraded in non-denaturing conditions, input controls using urea lysis of cells were included to show initial protein loads. (E, F) Western blots of FUS or FUS(4FL) in HttQ25-GFP-expressing or HttQ103-GFP-expressing cells.

the aggregates indicates that they are in an amyloid-like state (Figure 3B).

Proteins that specifically associate with polyQ aggregates are hypothesized to play either positive or negative roles in pathogenic processes. Previous attempts to identify the proteins that interact with huntingtin, have employed yeast-two-hybrid, immuno-precipitation and immuno-histochemical screening (47; 75; 80; 97; 155; 160; 171; 208; 228). The overlap in identified proteins from different methods is often low or the identification of hundreds of proteins limits clarity. Also, antibody-based screening is limited to a predetermined set of functioning antibodies and thus may overlook unexpected interactions. Our mass spectrometry-coupled approach, called TAPI, is specific to the most chemically-resistant forms of protein aggregates and eliminates the need to excise individual bands from acrylamide gels (109; 110). The high stringency of TAPI – due to DNase, RNase, and detergent treatment along with SDS-gel electrophoresis of aggregates – minimizes the identification of non-specific and loosely-associated proteins. In sum, our data set, in addition to the work of others, helps identify the important factors that make specific proteins most vulnerable to inclusion into polyQ aggregates.

Our analysis in both yeast and mammalian systems revealed a compact group of proteins enriched in polyQ aggregates. In both cell types, the proteins recruited to aggregates belonged to common functional classes (Tables 1 & 2); RNA-binding proteins and endocytosis-related proteins were disproportionately found in all Htt-polyQ samples. Importantly, the proteins identified by mass spectrometry could be independently confirmed by immuno-blotting. Moreover, our findings are supported by a previous study in Neuro 2A cells, which coupled Sarkosyl treatment with conventional 1D-gel

separation and manual band excision, to identify polyQ-associated proteins (138). Among the twelve proteins identified by Mitsui and coworkers, 8 were also identified by our approach in PC-12 cells (EEF1A1, HSPA8, HSP90AB1, MLF2, PSMC1, RAD23B, UBQLN2, YWHAZ; Table 1), suggesting that across cell types, certain proteins consistently have a high propensity for inclusion into polyQ aggregates. Moreover, functional orthologs are common to both the yeast and mammalian samples (YWHAZ/Bmh1/2p, DDX5/Dhh1p, SGTA/Sgt2p, CLINT1/Ent1/2p, hnRNPA3/Hrp1 and HSPA8/Ssa1p). The recruitment of SGTA/Sgt2p and CLINT1/Ent2p into polyQ aggregates was confirmed in both the yeast and mammalian systems by immunodetection methods. This overlap – not only between TAPI samples, but also across divergent organisms – suggests that these specific proteins (or their properties) may play a role in processes linked to pathological polyQ aggregation. This is supported in the literature where Ssa1p (134), HSP70 (147) and Sgt2p (228) have all been directly connected to polyQ aggregates in various models. As for the clathrin-interacting proteins CLINT1 and Ent2p, their presence suggests that aggregates can have important interactions with vesicle-dependent processes.

RNA binding proteins are disproportionately recruited to polyglutamine aggregates

RNA-binding proteins were highly represented in the Htt-polyQ TAPI data sets from yeast and rat cells. The presence of quality-control proteins, such as chaperones, was not particularly surprising since Htt-polyQ forms toxic intracellular aggregates. However, the enrichment of RNA-binding proteins was unexpected, especially considering the extensive nuclease treatment that is used prior to isolation of the aggregates (Figure 3). RNA-binding proteins have been shown to contribute to the

pathologies of a number of neurodegenerative diseases (34; 158). The aggregation of RNA-binding proteins is transient in normal cellular homeostasis, but may accumulate in neurodegenerative diseases due to pathological alteration of assembly and clearance pathways (158). This aberrant accumulation is frequently tied to interactions mediated by “prion-like” domains (105; 119), which are intrinsically unstructured domains that resemble the domains that enable certain yeast proteins to adopt self-propagating amyloid conformations. It is this intrinsically-disordered property that is likely responsible for the abundance of RNA-binding proteins in polyQ aggregates, because nucleic acid-interacting proteins frequently having intrinsically unstructured regions. Thus, our results suggest that a proteins’s ID domain, not the RNA binding per se, may be the major determinant of inclusion into Htt-polyQ aggregates (Figure 10C, D). This is corroborated by our observation that disrupting the RNA-binding domain of FUS had no effect on its inclusion in polyQ aggregates (Figure 10E, F).

Mitochondrial proteins are found in polyglutamine aggregates

Mitochondrial proteins represent a significant fraction of the polyQ-associated proteins in yeast and rat cells, 8% and 18% respectively. In yeast, this percentage is less than the proteome representation of mitochondrial proteins (~18(235)), but in rat cells this is an over-representation (5-12%(19)). In yeast cells the polyQ aggregates form primarily in the cytoplasm, and in the rat cells the aggregates are equally in the nucleus and cytoplasm. Mitochondrial proteins are mostly synthesized in the cytoplasm and transported as unfolded polypeptides into mitochondria post-translationally. A probable explanation for the presence of mitochondrial proteins in aggregates is that because of their unfolded conformation they are vulnerable to integration into aggregates or because

of their dependence on chaperones may be more sensitive to general problems associated with protein quality control.

Intrinsically disordered domains facilitate protein recruitment to Htt-polyQ aggregates

Why are proteins with ID domains tightly and disproportionately associated with Htt-polyQ aggregates? Not only did we observe an enrichment of ID domain-containing proteins, but previous proteomic studies of polyQ also reveal data sets that are rich in such proteins (97; 160; 244). Ratovitski *et al.* observed that proteins with ID domains of ≥ 30 amino acids were enriched in aggregates formed by Htt-50Q in HEK293 cells. This is consistent with our results, although we found significant enrichment with very long ID domains (≥ 100 aa). Similarly, Raychaudhuri *et al.* performed a bioinformatic analysis of intrinsic disorder in neurodegenerative disease-associated proteins. Their bioinformatics dataset (obtained from Entrez Gene database keyword search) when compared to control datasets indicates an increase in intrinsic disorder (using FoldIndex) for Huntington disease-associated proteins with ID domains up to 100 amino acids in length (161). Intrinsically unstructured regions frequently facilitate molecular interactions or serve as sites of post-translational modifications (205), as well as being prominent features of many nucleic acid-binding proteins and chaperones (85; 204). These domains generally lack hydrophobic residues sufficient for adopting a folded structure in aqueous environment (214) and thus may be more accessible to aggregation simply by virtue of accessibility. We demonstrated that elimination of the major ID domains of two proteins eliminated their co-aggregation with polyQ (Figure 10).

While we do not assert that ID domains are solely responsible for association with Htt-polyQ aggregates, it is clear that ID domain content plays a prominent role in

recruiting secondary proteins to aggregates. In some cases, quality-control proteins could even employ intrinsic disorder within a specific domain to facilitate the functional recognition of misfolded protein aggregates. However, because of the large number of quality-control proteins that we identified it cannot be concluded that there is a single mechanism by which all such proteins are tightly associated with polyQ aggregates; some proteins, such as UBQLN2, likely only have specific affinity for aggregates following ubiquitination.

Why are neurodegenerative disease-associated proteins recruited to polyglutamine aggregates?

Neurodegenerative disease-linked proteins that have been identified in cellular inclusions in their respective diseases represent a fifth (19/91) of the total TAPI-identified Htt-polyQ aggregate-associated proteins (Table 4). Examining this sub-set reveals that nearly all of them contain ID domains. We observed that some ALS-associated proteins were trapped in polyQ aggregates. Three of these proteins – FUS, HNRPA1 and TDP-43 – are RNA-binding proteins that form pathological inclusions in certain forms of ALS and frontotemporal lobar dementia (FTLD) (26; 37; 125). These proteins have ID domains that resemble the domains of yeast prion proteins due to similar amino-acid composition. It has been concluded that these prion-like domains may be primary-aggregating species, but the fact that these ALS-associated proteins are pulled into polyQ aggregates suggests that in some long-lived cells, such as neurons, there could be underlying protein quality control problems with proteins like FUS and TDP-43 getting preferentially recruited into pre-existing primary aggregates (59). For example, it is possible that over decades, intermediate-length polyQ expansions in various proteins lead to persistent aggregates that recruit proteins with ID domains; these inclusions would be

marked (*i.e.* immuno-positive) by specific aggregation-prone proteins. Alternatively, the diminution of protein-quality control with aging may create a cellular environment where proteins with long ID domains are increasingly susceptible to aggregation, thus polyQ models and their induced stress may be good tools for identifying proteins that are most vulnerable to aggregation under conditions of compromised protein-quality control.

Mechanisms of Toxicity

Both the yeast and mammalian model systems show a correlation between polyQ-expanded huntingtin aggregation and cellular toxicity (134; 239), however much debate still persists as to the mechanism by which aggregation leads to cell death (74). One possible mechanism is that sequestration of proteins to an aggregate may impair cellular function (174), which is arguably an indirect loss of function. Such sequestration by aggregates of polyQ-expanded Ataxin3 (spinocerebellar ataxia-causing protein) was proposed to cause a loss-of-function toxicity (243). Our observations suggest this indirect loss-of-function toxicity due to sequestration of essential proteins could occur with huntingtin aggregation as well. We observe altered cellular localization for a subset of proteins when Htt-polyQ aggregates are present. It is possible that these proteins may maintain some function while associated with the Htt-polyQ aggregate. If the ID domain of a protein becomes embedded in the polyQ aggregate, while globular or functional domains remain peripheral, function in the wrong place at the wrong time could be a gain-of-function toxicity associated with aggregates. Stoichiometrically, this may be more plausible than loss-of-function because we observe only a fraction of any given co-aggregating species mis-localized to the polyQ aggregate, which itself is quite abundant

due to over-expression of Htt (Figure 9). Of course, a combination of gain-of-function and loss-of-function mechanisms could contribute to the overall cellular toxicity.

Although many techniques have been employed to identify huntingtin-interacting proteins, few examine specifically the amyloid form. Our results show that a select group of proteins are trapped by polyQ amyloid-like aggregates. Proteins with long ID domains are disproportionately prone to inclusion, as are many proteins that are associated with other neurodegenerative diseases. The enrichment in ID domain-containing proteins in polyQ aggregates, and the elimination of this enrichment when the ID domains are deleted, reveals the significant role of protein structure in determining if a protein gets secondarily recruited into certain types of aggregates. Thus, while some proteins might be predicted to be recruited into aggregates because of their function (*i.e.* quality-control proteins or proteins that interact with the soluble form of an aggregating species), many proteins may be recruited simply as a consequence of their secondary and tertiary structural elements. The metastable structure and accessibility of long ID domains may render proteins particularly susceptible to aberrant inclusion in amyloid-like aggregates. Recently, Habch and colleagues put forth the idea that ID proteins represent a class of pharmacological targets (72). As our results suggest, if the recruitment of specific ID domain-containing proteins into pathological aggregates is critical to cellular degeneration, then targeting ID domains to reduce their sequestration may have therapeutic potential in a variety of neurodegenerative diseases.

Table 2: Molecular function of proteins associated with Htt-polyQ aggregates as identified by TAPI (N=52) in *Saccharomyces cerevisiae*

Category	% of total	Protein Name							
Protein Quality Control/ Chaperone	12%	Apj1	Bmh1	Def1	Mca1	Sgt2	Ydj1		
RNA/DNA Binding	44%	Ccr4	Cyc8	Dhh1	Eap1	Hrp1	Ixr1	Mbf1	Mcm1
		Mot3	Nab3	Nam8	New1	Nrd1	Pbp1	Pin4	Pop2
		Puf3	Snf5	Srp54	Taf5	Tup1	Whi3	Ygr250c	
Mitochondrial	8%	Apj1	Nam8	Puf3	Ynl208W				
Endocytosis, vessicle & cytoskeletal transport	21%	Akl1	Ent1	Ent2	Gts1	Pan1	Scd5	Sla1	
		Pin3	Sec24	Yap1801	Yap1802				
Other	21%	Cbk1	Epo1	Gal2	Mum2	Nup57	Nup100	Nup116	Pgd1
		Sml1	Slm1	Ylr177w					

Molecular function determined by gene ontology and *Saccharomyces* Genome Database. Proteins in bold represent those with overlapping cellular functions; totals exceed 100% because of multiple categorizations. Yeast ribosomal proteins Rpl30 and Rps6b were excluded.

Table 3: Molecular function of proteins associated with Htt-polyQ aggregates as identified by TAPI (N=91) in *Rattus norvegicus*

Category	% of total	Protein Name							
Protein Quality Control/ Chaperone	26%	Adrm1	Cltc	Ddi2	Dnaja2	Dnaja4	Dnajb1	Dnaje7	Hspa8
		Lap3	Ppia	Psmb2	Psmc1	Psmc2	Psmc3	Rad23b	Sgta
		Sumo2	Ubqln2	Ubqln4	Ubxn711	Usp7	Vps35		Sqstm1
RNA/DNA binding*	37%	Aars2	Akap81	Arl6ip4	Atp5a1	Atrx	Ddx5	Dync1h1	Eif4g1
		Fasn	Fus	Gigyf2	Hnrnpa3	Hnrnpf	Hnrnp2	Hnrnpm	Hnrnpu
		Matr3	Nono	Nr3c1	Pcbp1	Ppia	Prpf40a	Prrc2b	Rbms1
		Tardbp	Tcerg1	Tcf20	Tnrc6b	Tufm	Xrn2	Ythdf1	
Mitochondrial	18%	Aars2	Acad9	Acsf2	Atp5a1	Cltc	Etfb	Gls	Hadha
		Idh3B	Ndufs7	Ogdh	Pck2	Pdhb	Suclg2	Tufm	Idh2
Endocytosis, vessicle & cytoskeletal transport	14%	Aak1	Arl6ip4	Asap1	Clint1	Cltc	Cnn2	Dync1h1	Myo1d
		Rab10	Scyl2	Tfg	Vps35				Nsfl1c
Other	15%	Ep300	Gnao1	Kprp	Ldha	Maged1	Mlf2	Phgdh	Plekhh2
		Sik3	Sfn	Tgm3	Thy1	Ywhab			Ppp2r1a

Molecular function assignments were determined by gene ontology, RGD, and NCBI. *RNA-binding was assigned in some cases on empirical data (20), thus classification does not necessarily imply primary function. Proteins in **bold** represent those that were placed in multiple categories, thus totals can exceed 100%. Rat ribosomal proteins Rpl6 and Rpl13a were excluded.

Table 4: Htt-polyQ aggregate-associated proteins found in inclusions and aggregates in various neurodegenerative diseases.

Name	Disease or Disease Model	ID Domain	RBP	PQC	Reference
Aak1	ALS-SOD1	209, 81, 50			(185)
Fus	ALS	284, 71, 87	yes		(48; 116; 217)
Hnrnpa3	ALS	39	yes		(141)
Hspa8	HD & SCA1	46, 36		yes	(36; 87)
Matr3	ALS	31, 57, 42, 119, 72	yes		(92)
Mlf2	HD	136			(104)
Nono	ALS-FUS	66	yes		(183)
Ppia	ALS-SOD1	None	yes	yes	(7)
Rad23b	HD	120, 67, 41		yes	(10; 224)
Sqstm1	PD & ALS	153		yes	(115; 139; 248; 250)
Suclg2	AD	None			(159)
Sumo2	HD & ALS	None		yes	(146; 149)
Tardbp (TDP-43)	ALS	45, 55	yes		(96; 192)
Tcerg1	HD	138, 177, 70, 36	yes		(80)
Tcf20	HD	87, 739, 50, 32, 709, 34, 121	yes		(242)
Tfg	CMTD	35, 54, 105			(84)
Tgm3	HD	49			(246)
Ubqln2	ALS	62, 58, 42, 81, 31, 84		yes	(43)
Ywhab	ALS	None			(100)

The numbers for ID Domains indicate the length of distinct unstructured regions ≥ 30 , for the identified rat proteins from PC-12 cells, as determined by IUPred-L. RBP = RNA-binding protein as described in the legend of Table 3; PQC = protein quality control.

Table 5: Analogous proteins in yeast and rat associate with Htt-PolyQ aggregates.

Rat TAPI Protein	Yeast TAPI protein	Function
Aak1	Ak11	Ser/Thr kinases involved in endocytosis and actin cytoskeleton organization; recruit endocytic accessory factors
Sfn/Ywhab	Bmh1	14-3-3 proteins; associated with diverse protein binding and signaling activities
Dnaja2/Dnaja4/ Dnaja7/Dnaja1	Ydj1/Api1	HSP40 co-chaperones; function with HSP70s; generally involved in protein folding and quality control
Hnmpa3	Hrp1	Heterogeneous nuclear ribonucleoproteins; bind RNA; involved in RNA processing
Sgt1	Sgt2	Glutamine-rich cytoplasmic co-chaperone; functions in post-translational membrane insertion of proteins in yeast
Clint1	Ent1/Ent2	Epsin-like proteins involved in endocytosis; clathrin interactors
Ddx5	Dhh1	Cytoplasmic DExD/H-box RNA helicases; multiple RNA-related functions

Protein similarity determined using Ensembl comparative genomics, RGD, SGD, and % identity.

CHAPTER 3: Discussion and overall conclusions

MECHANISM OF HTT-POLYQ TOXICITY

Disease research has revealed that amyloid-like protein aggregates are associated with the pathology of many neurodegenerative diseases (33; 107; 165; 212). These aggregates were shown to contain not only the disease-associated protein, but also many other cellular proteins. In this work we have utilized TAPI to identify aggregate-associated cellular proteins. Using deletional analysis, we have also determined that the intrinsically disordered (ID) domains regions within two of these proteins – Sgt2 and FUS –are required for their association with Htt-polyQ aggregates.

ID DOMAINS IN HTT-POLYQ AGGREGATES

ID domains are regions of 30 amino acids or greater that contain a higher proportion of low hydrophobicity and high net charge residues, resulting in a less compact structure and defined by the lack of globular structure (168). Even though these regions do not form a folded structure, or share sequence homology, proteins containing ID domains perform many functions. These include post-translational modification to proteins, binding RNA or DNA, or acting as chaperones in protein folding (Tables 2 & 3). Intriguingly, biophysical analysis of the Htt-polyQ interacting proteins that we have identified reveals that the functions of proteins containing regions of intrinsic disorder are largely the same as those identified for ID domain proteins in general. Previous work by Ratovitski and colleagues as well as Raychaudhuri and colleagues suggests that intrinsically disordered (ID) domains may be shared among cellular proteins recruited to amyloid-like aggregates (160; 161).

Interest in intrinsic disorder has increased substantially in the last fifteen years as the identification of proteins containing ID domains has concomitantly grown (162). Currently it is estimated that between 15 and 45% of eukaryotic proteins contain ID domains 30 amino acids or greater in length (81; 140). ID domains are conformationally indistinct, defined by their lack of ordered structure these regions exist as highly dynamic ensembles of conformations, enabling them to interact with many binding partners (2; 140; 162; 168). The plasticity of these regions lends specific advantages to the proteins containing them.

First, the constant change in conformation and disorder-to-order transition upon target binding of ID domains results in a decoupling of specificity and affinity, resulting in high specificity/low affinity interactions. The disorder-to-order transition upon target binding of ID domains also increases the speed of interaction (168). The interaction surfaces of ID domains are large and contain multiple molecular recognition features (MoRFs) allowing for multiple distinct interactions to occur within a common binding surface (81; 129; 162; 168; 219). In addition, ID domains are promiscuous interactors, resulting in the classification of ID domain-containing proteins as ‘hub’ proteins in cellular networks (81; 129; 162; 168).

The constant fluctuation in ID domain conformation means a single structure for these regions does not exist. However, recently many methods have been utilized in attempts to characterize the structure ensembles of ID domains including: nuclear magnetic resonance (NMR) spectroscopy, small angle X-ray scattering, single molecule fluorescence, molecular dynamics (MD) simulation, paramagnetic relaxation enhancement (PRE), and circular dichroism (2; 90; 102; 106; 140). The most successful

of these methods are NMR spectroscopy and MD simulations resulting in a database of ID domain-containing protein conformational ensembles database (218).

Although ID domains are necessary and incredibly useful in cellular pathways, failure to regulate ID domain behavior is linked to variety of diseases (2; 94; 107; 189; 215). Many of the disease-linked proteins associated with neurodegenerative diseases contain ID domains (215). In fact, the mechanism of polyQ toxicity arose from the observation that proteins with longer polyQ repeats, irrespective of protein content, form insoluble aggregates (211). Biophysical characterization of polyQ peptides of varying length reveals they are intrinsically disordered (24). Here we have shown that the mechanism of toxicity in HD seems to be intrinsically linked to ID domains; both in the disease-linked huntingtin protein, and those proteins recruited to Htt-polyQ aggregates in disease models (232). Considering that there are at least eight other neurodegenerative diseases associated with aggregating proteins containing expanded polyQ regions, it is of great interest to investigate if this theory extends beyond HD to include all of these polyQ-associated neurodegenerative diseases.

CORRELATION BETWEEN TAPI FINDINGS WITH OTHER STUDIES OF Htt-POLYQ-ASSOCIATED PROTEINS

Our analysis also revealed that a significant fraction of the proteins identified in Htt-polyQ aggregates were indicated in other neurodegenerative diseases (19/91) (232) (Table 4). The finding that aggregate-associated proteins of one neurodegenerative disease are implicated in another suggests a unifying feature may exist linking all of these diseases. A significant portion of the aggregate-associated proteins also contain ID domains. In the future, it would be of interest to determine if ID domains represent a

shared biophysical attribute amongst proteins recruited to neurodegenerative disease-related aggregates.

Our TAPI analysis also revealed novel proteins associated with amyloid-like Htt-polyQ along with cellular proteins previously shown to be associated with HD (Tables 2 & 3). Of particular interest are two subsets; proteins conserved in both yeast and mammalian model systems, and cellular proteins identified in protein aggregates of other neurodegenerative diseases (Table 4). Our TAPI analysis revealed a cohort of proteins conserved between mammalian and yeast systems - YWHAB/Bmh1/2p, DDX5/Dhh1p, SGTA/Sgt2p, CLINT1/Ent1/2p, hnRNPA3/Hrp1 and HSPA8/Ssa1p. The association of two pairs of these SGTA/Sgt2p and CLINT1/Ent2p were further confirmed by immunoblotting for interaction with Htt-polyQ aggregates. These results suggest that comparison of our TAPI results with those of other studies may reveal a smaller subset to target for further investigation.

Several studies of protein aggregates associated with Htt-polyQ relied on *a priori* knowledge of target proteins or hypotheses of associated proteins (66; 101; 132; 155; 228; 237). These studies sought to confirm the aggregation of previously identified proteins with Htt-polyQ. However, two studies in the yeast system examine Htt-polyQ aggregate-associated proteins without bias (155; 229). Together with our own TAPI analysis, these three studies make up our meta-dataset. Each of these three studies utilized different techniques for the isolation and purification for aggregates. TAPI analysis identified 96 proteins specifically associated with Htt-polyQ (Appendix 4). Park and colleagues performed Stable Isotope Labeling with Amino Acids in Culture (SILAC): growth of Htt-polyQ expressing yeast strains in culture media containing

heavy ($^{13}\text{C}/^{15}\text{H}$) amino acids, coupled with LC-MS/MS analysis (155). This differential growth technique allows for inclusion of both the control (light amino acids) and Htt-polyQ expressing strains (heavy $^{13}\text{C}/^{15}\text{H}$ amino acids) in a single mass spectrometry run. This analysis identified 106 proteins specifically associated with Htt-polyQ (found in just the heavy $^{13}\text{C}/^{15}\text{H}$ mass spec. sample) in two of three biological replicates (Appendix 4) (155). Wang and colleagues expressed the Htt-Q103 plasmid construct followed by 3D antibody mesh immunoprecipitation (IP) coupled with 2D gel analysis and mass spec protein identification (228). In this method sequential antibody incubation with the sample forms an antibody mesh around the targeted aggregate, which is then precipitated and purified by centrifugation through a sucrose pad. The sample is then run on a 2D gel, bands of interest are removed and analyzed by mass spectrometry (Appendix 4) (228).

Comparison of these datasets – TAPI, 2D gel, and SILAC - revealed just five proteins shared identified in all three analyses – Bmh1, Ssa1, Sgt2, Pin3, and YNL208W (Figure 11). Bmh1 and Ssa1 have been investigated previously with regards to their contribution to Htt-polyQ aggregation (see below). However, the remaining three proteins have not previously been studied. Here, we will discuss the current state of our understanding of the association of these proteins with Htt-polyQ.

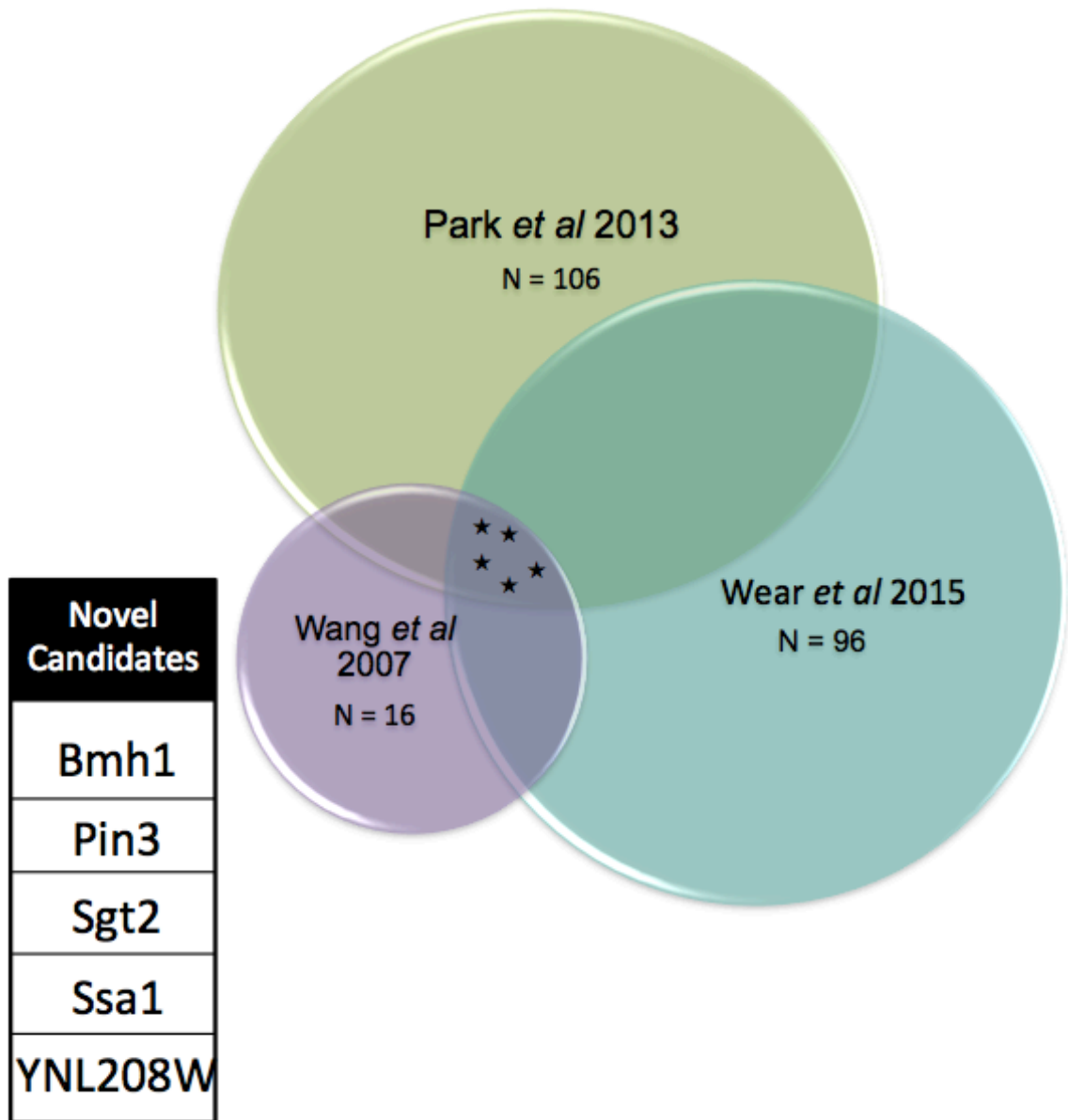


Figure 11: Meta-analysis identifies Htt-polyQ aggregate associated gene candidates. Three data sets were examined for shared Htt-polyQ aggregate-associated proteins – SILAC (Park 2013), 2D gel (Wang 2007), and TAPI (Wear 2015). Five candidate proteins found by all three association studies (see inset) identified for follow-up study.

Bmh1 is a 14-3-3 regulatory protein involved in regulation of exocytosis, vesicle transport, aggresome formation, apoptosis, and post-transcriptional protein levels (12; 15; 31; 63; 216; 241). Other 14-3-3 proteins have previously been reported to associate with

Htt-polyQ aggregates (223), and the levels of 14-3-3 protein in the cerebrospinal fluid (CSF) is currently used as a biomarker and diagnostic factor for the human prion disease Crutzfeld-Jacob Disease (CJD) (68; 179). Wang and colleagues showed that deletion of Bmh1, but not its homolog Bmh2, eliminated Htt-Q103 aggregation in the yeast system, suggesting that this protein may play an active role in aggregate formation or stabilization (229). However, deletion of Bmh1 also resulted in increased Htt-Q103 toxicity; this suggests that aggregates may have lower toxicity than free Htt-polyQ (229).

Ssa1 is a member of the HSP70 chaperone family and has three nearly identical additional family members, Ssa2, Ssa3, and Ssa4. Hsp70 proteins have been shown to work with Hsp100 disaggregase Hsp104 in fractionating amyloid aggregates in yeast. For a particular aggregate, the prion [URE3], overexpression of Ssa1, but not Ssa2, was able to completely eliminate the aggregate from yeast cells (181). Meriin and colleagues examined the effects of Ssa family proteins on Htt-Q103 toxicity in yeast and found that deletion of Ssa1/Ssa2 suppressed Htt-polyQ cell toxicity in a previous study (135). However, deletion of Ssa1/Ssa3 or Ssa1/Ssa4 did not affect Htt-polyQ toxicity (135), indicating each of these Ssa proteins may have unique functions that are masked in double deletions.

Sgt2 is implicated as a scaffold protein in the Guided Entry of Tail-anchored (TA) proteins (GET) trafficking pathway where it binds the hydrophobic tails of TA proteins to inhibit their aggregation prior to membrane insertion (103). Chartron and colleagues examined the interactions of Sgt2 with Get proteins, HSP chaperones, and TA proteins (22). They found that the C-terminus of Sgt2 mediates the interaction between HSP chaperones and TA proteins while the N-terminus of Sgt2 binds Get4/Get5 (22). We

found that deletion of the C-terminal ID domain of Sgt2, Sgt2 $\Delta^{300-346}$, interrupted interaction between Sgt2 and Htt-polyQ aggregates (232). The C-terminal region of Sgt2 previously shown to interact with HSP chaperones and TA proteins overlaps with the C-terminal ID domain, encompassing amino acids 220-346 (22; 103). It is possible that the C-terminal ID domain of Sgt2 is responsible for interaction with proteins with a higher propensity to aggregate, like TA proteins and Htt-polyQ. Of particular interest for future research is to determine if Sgt2/SGTA interacts with all HSP70 chaperone proteins, or only specific partners.

Pin3 acts in concert with paralog Lbs1 to negatively regulate nucleation of actin filaments (27; 190). Pin3 was also identified in a screen for inducers of the yeast prion [PIN+] suggesting a role in prion formation (44). Htt-polyQ forms prion-like aggregates, suggesting that Pin3 may influence Htt-polyQ aggregation similar to yeast prions. Further investigation will be necessary to determine if Pin3 associates with Htt-polyQ in a functional manner.

YNL208W is an uncharacterized protein with no known function. It is proposed to interact with ribosomes and mitochondria (56; 164). YNL208W is nearly entirely disordered. In fact, biophysical analysis of these overlapping proteins shows that the only characteristic shared among all five is that they all contain at least one ID domain (Figure 12). This further suggests that ID domains play a role in recruitment of proteins to Htt-polyQ.

In the future these proteins, particularly Sgt2, Pin3, and YNL208W are of interest to characterize their impact on Htt-polyQ associated toxicity.

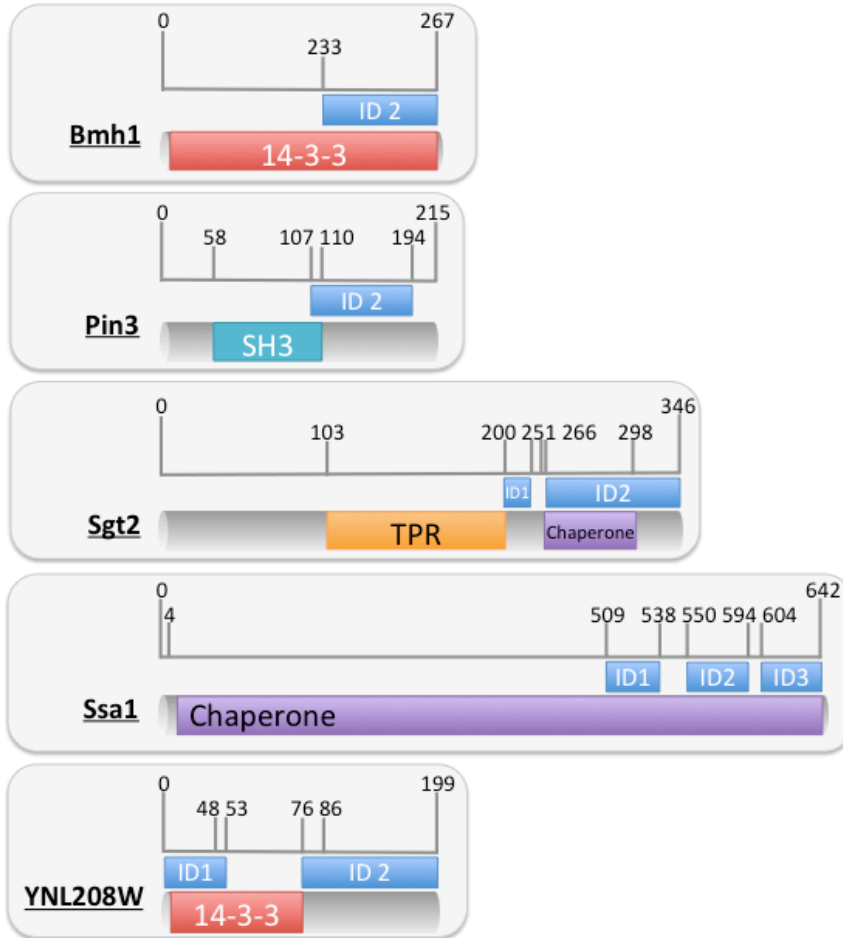


Figure 12: Proteins that alter Htt-polyQ toxicity contain intrinsically disordered (ID) domains.

TAPI-identified Htt-polyQ interacting cellular proteins confirmed by follow-up analysis were examined to identify conserved domains using NCBI conserved domains database (CDD) as well as ID domains using IUPred-L. Proteins containing 14-3-3 domains are regulatory molecules involved in signaling, SRC Homology 3 (SH3) domains are found in proteins of signaling pathways, tetratricopeptide repeat (TPR) domains are involved in protein-protein interactions form scaffolds, and chaperone domains are involved in protein folding and unfolding as well as the assembly and disassembly of protein complexes.

ID DOMAINS AS HUNTINGTON DISEASE THERAPEUTIC TARGETS

When considering the impact of this research on the area of neurodegenerative disease research and specifically HD, it is important to examine the possibility for therapeutic development. The survival of patients diagnosed with HD is only one to two decades beyond diagnosis. Unlike many other diseases, no therapeutics for HD are

currently able to prolong patient survival. The development of a therapeutic to extend HD patient survival would have significant impact on both patients and their families.

Research identifying huntingtin-interacting proteins has led to the identification of molecular pathways involved in HD-associated neuron death. However, no single molecular pathway identified has been able to account for all neuron death or reverse progression of HD pathology. The identification of a common link between all huntingtin-interacting proteins – the ID domain – represents a potential drug target for HD. The identification of ID domains as a common link is furthered by our observation that deletion of ID domains in select proteins – Sgt2, and FUS – results in attenuation of target protein interaction with Htt-polyQ. When considering this observation, it seems likely that ID domains may play a role in recruitment of proteins to Htt-polyQ and may impact disease-associated neuron death.

Recent studies indicate that ID domains represent drug targets for therapeutic development (72; 81; 94; 129; 140; 207; 211). However, the usual methods for protein-targeted drug development cannot be applied to ID domains due to the lack of stable three-dimensional structure. In fact, a search of the current “druggable proteome” by Hu and colleagues shows that current known drug targets contain two-fold less disorder than the human proteome (81). To circumnavigate these challenges a number of theoretical methods have been proposed for the intelligent design of drugs targeting ID domains. These include (i) inhibition or (ii) modulation of ID domain interactions with binding partners (iii) stabilization of the ID domain in natively disordered state and (iv) induction of allosteric inhibition (2; 94; 129).

Currently two protein-protein interactions involving ID domain-containing proteins have been successfully targeted. Peptides and small molecules have been identified that inhibit the interaction of p53 and Mdm2 (58; 225) and small molecules that bind to monomeric and disordered c-Myc and inhibit c-Myc-Max interaction (73; 77; 136) have been developed for cancer treatment. Currently, no small molecules targeting protein-protein interactions in neurodegenerative diseases have been identified.

Although no small molecules targeting protein-protein interactions have been identified, small molecules targeting the ID domain structure have been more successful with regards to neurodegenerative disease-associated proteins. Toth and colleagues utilized an *in silico* structure-based fragment mapping and docking screen to identify compounds that would bind to one or more of eight theoretical binding pockets in the various conformational states of the α -Synuclein protein (207). The compound, ELN484228, was effective at reducing neuronal toxicity in some, but not all, of the Parkinson's disease models tested (207). Although this is only preliminary, this study, along with a theoretical analysis of the Alzheimer's A β peptide (251) indicate that analysis of ID domain conformational ensembles for binding pockets has the potential to reveal viable small molecules (251). Sormanni and colleagues utilized a computational method to identify complementarity determining regions (CDRs) to target specific epitopes within ID domains (189). These CDRs were then grafted onto an antibody scaffold. These CDR antibodies were developed for three proteins - α -Synuclein, A β , and Islet Amyloid Polypeptide (IAPP) - and confirmed to target these proteins by enzyme-linked immunosorbent assay (ELISA) (189).

CONCLUSION

It was shown that Htt-polyQ aggregates and associated cellular proteins can be isolated and characterized without *a priori* knowledge of protein identities using the TAPI method. The proteins associated with Htt-polyQ in mammalian and yeast models show overlapping molecular functions and conserved ID domains. Deletion of ID domains in two identified proteins (FUS and Sgt2) resulted in loss of association between that protein and Htt-polyQ suggesting that ID domains are important for cellular protein association with Htt-polyQ.

Many of the mammalian proteins identified by TAPI are indicated in other neurodegenerative diseases and also contain ID domains. Recent publications indicate that ID domain containing proteins and protein-protein interactions involving these proteins can be targeted with small molecules and rationally designed antibodies. However, thus far, no ID domain-targeting drugs have been identified for the mutant huntingtin protein. Our studies indicate that huntingtin ID domain-targeting drugs represent the next step in HD research. If successful, these therapeutics would represent the first HD treatment to target symptoms of the disease.

CHAPTER 4: Materials and methods

Yeast Strains and Plasmids

Yeast strains BY4741 (MATa *his3 leu2 met15 ura3* [PIN+] [psi-]) or W303 (MATa *leu2 ade2-1 ura3 can1 trp1 his3 gal+*) were transformed with Gal-Inducible Huntingtin Exon 1 polyQ expansion plasmids (Htt-Q25-GFP or Htt-Q103-GFP) (109; 134). Genes (or truncated variants) were cloned into pFPS261 and pFPS262, which encode single HA tags in-frame with the multiple cloning site, thus adding c-terminal epitope tags to genes cloned into the XhoI site. Plasmids pFPS261 (*CEN LEU2 P_{GALI}*) and pFPS262 (*2 μ LEU2 P_{GALI}*) are respectively derivatives of previously-described pH316 (55) and pH317 (54). Truncated *SGT2- Δ ID* is *SGT2 Δ 300-346*, and truncated *FUS- Δ ID* is *FUS Δ 1-135*. Human α -synuclein-GFP was expressed from the previously-described plasmid DK258 (*2 μ LEU2 P_{GALI}*) (113). All strains were cultured in synthetic defined media with appropriate auxotrophic selection for plasmid maintenance. Protein expression was induced overnight with growth on selective galactose-containing medium.

Cell Lines and Maintenance

The PC-12 cell lines were previously described by Wyttenbach et al. (239). Briefly, the PC12 cells were stably transformed with Doxycycline-inducible GFP-tagged normal Htt-Q23 Exon 1 or expanded Htt-Q74. Cells were cultured on collagen IV (BD Biosciences) coated T-75 flasks and maintained in DMEM with 75 μ g/mL hygromycin, 100 U/mL penicillin/streptomycin, 2 mM L-glutamine, 10% heat-inactivated horse serum, 5% Tet-negative fetal bovine serum and 100 μ g/mL G418 at 37°C, 10% CO₂. Culture reagents were obtained from Corning.

Technique for Amyloid Purification and Identification (TAPI)

The yeast TAPI protocol was performed as previously described (109) with the following alterations: Buffer A - 30mM Tris-HCl pH=7.5, 5 mM DTT, 40 mM NaCl, 3 mM MgCl₂, 5% glycerol, 1x Complete protease inhibitors cocktail (Roche), 20 mM NEM (Sigma), 0.5 µl Benzonase nuclease (250 U/ul; Sigma); RNase A (200 µg/ml; Sigma) treatment for 15' at 4°C prior to ultracentrifugation at 300,000 g.

Mammalian TAPI samples were prepared as follows: a cell pellet of at least 1×10^9 cells (~100µl in volume) were lysed in cold modified 300 µl RIPA buffer (1% Triton X-100, 0.1% SDS, 1% Sodium deoxycholate, 150 mM NaCl, 10 mM Na₃PO₄, 50 mM NaF, 5 mM MgCl₂, 5mM DTT, 5mM Na₃VO₄, with 1x protease inhibitors (Roche) and 33 U DNase 1 (Sigma), 3 mg RNase A (Sigma) and 750 U Benzonase (Sigma)), followed with a 10-minute incubation at room temperature and then 20 minutes of mild rotation at 4°C. Lysates were then spun at low speed (5 minutes centrifugation at 100 g) and the pellets were subjected to a second lysing in modified RIPA buffer with rotation for 20 minutes at 4°C. Combined supernatants were run through a 30% sucrose gradient by ultracentrifugation (2 hours at 45000 rpm at 4°C with a Beckman SW-50A rotor). Some samples were analyzed at this point to determine the presence of specific proteins in the pellet fraction by western blotting. After ultracentrifugation, the pellet was re-suspended in high SDS buffer (1x TBS, 5 mM DTT, 5 mM EDTA, 2-4% SDS, with 1x protease inhibitors (Roche)) and incubated with gentle mixing for at least 20 minutes at 37°C. Samples were run at 200 volts on an acrylamide gel (Any kD™, Bio-Rad) in 10% glycerol with 0.1% bromophenol blue to monitor sample migration. The top 3 millimeters of the wells were excised and frozen. Frozen gel fragments were thawed and resuspended in elution buffer (10mM Tris pH 8.0, 0.4% SDS, 5mM DTT), then mixed and incubated

at 99°C for several minutes. Sample volume was reduced by half in a speed-vac and applied to a desalting column (Zeba spin, Thermo Scientific), pre-equilibrated with 25mM triethylammonium bicarbonate (TeABC). The flow-through was analyzed for elution efficiency by western blot prior to digestion.

Samples were digested for mass spectrometry (MS) analysis using a previously described method (238) with minor modifications. Briefly, DTT was added to the sample to obtain a 15 mM solution prior to addition of iodoacetamide to a concentration of 50 mM followed by a 20-minute incubation at 30°C in the dark. Next, 9M urea was added to the sample, which was then filtered (30 kDa amcon 30K spin filter – centrifuged at 16000 x g 5-10min), and washed with 25mM TeABC twice, and finally trypsin digested (5-10µg trypsin) overnight at room temperature. The sample was retrieved from the column by centrifugation (16000 x g for 4 min), and washed with 25mM TeABC prior to lyophilization. Lyophilized samples were analyzed by tandem MS/MS by Johns Hopkins Mass Spectrometry and Proteomics Facility.

Mass Spectrometry Analysis

Samples were run on Q-Exactive (Thermo Scientific) or Orbitrap Velos (Thermo Scientific) at 70,000 resolution for MS and 17,500 for MS2, or 30,000 resolution for MS and 15,000 for MS2 respectively. The data were collected in data dependent mode with the top 15 precursors chosen for MS/MS. The peptides were eluted with a 90 minute gradient at 300 nanoliters per minute after trapping and desalting for 5 minutes at 5 microliters per minute. Peptides were fragmented with a normalized collision energy of 27 for Q-Exactive, and 35 for Orbitrap Velos. Target values were 3E6 ions with 60 millisecond maximum injection time for MS and 5E4 with 250 milliseconds for MS2 for

Q-Exactive and 1E6 ions for MS with a 100 millisecond maximum injection time for MS and 5e4 with 300 milliseconds for MS2 for LTQ Orbitrap Velos.

All data were searched using Mascot (v2.6 Matrix Science) through Proteome Discoverer (v1.4 Thermo Scientific). The database for Yeast included Htt-Q25-GFP in addition to the RefSeq 2014 *Saccharomyces cerevisiae* and the database for Rat included GFP-Htt-Q23 in addition to the RefSeq 2014 *Rattus norvegicus*. Variable modifications included oxidation on Met, deamidation on N and Q, and carbamidomethylation on C. Data were searched with a 30 part per million (ppm) tolerance for precursor mass and 0.03 daltons for fragment masses. Data were searched with and without the MS2 processor node which deisotopes the MS2 spectra to the +1 charge state prior to searching. Data were filtered through the Target Decoy PSM Validator.

The resulting data were filtered through Scaffold software for Total Spectra Count at 5% FDR. Criteria for proteins to be defined as associated with Htt-Q74/103 (Htt-PolyQ aggregate) were as follows: 2 or more total spectra, and present in the expanded Htt-polyQ aggregate while absent in the short Htt-polyQ control sample (henceforth termed binary) in at least 2 of 4 samples examined. Thus, the requirement for a protein to be considered positive is that in at least 2 of the matched samples (containing a pair of Q-short and a Q-long samples) performed at the same time on the same instrument, the protein shows ≥ 2 spectra in the Q-long and not the Q-short. This resulted in 52 Htt-polyQ aggregate associated proteins identified in *Saccharomyces cerevisiae* (Table 1; Appendix 1) and 91 Htt-polyQ aggregate associated proteins identified in *Rattus norvegicus* (Table 3; Appendix 2). In the accompanying supplementary files, we also include the proteins that meet a less stringent threshold: present in the Htt-polyQ aggregate sample while

absent in the control for at least one sample set (binary) and 2-fold greater spectra number in Htt-Q74/103 (Htt-PolyQ aggregate) than the short Htt-polyQ control in at least one additional sample pair. As a control for the above method of processing, we also used normalized spectral counting (NCS), a label-free quantification method that compares the number of MS/MS spectra assigned to each protein normalized for the total spectral counting among samples (150; 253). The NCS processing yielded very similar data sets (data not shown).

Bioinformatic Analysis

The proteins identified to be associated with polyQ aggregates using TAPI and MS were further characterized by molecular function, Q/N content and intrinsic disorder. Gene Ontology, Saccharomyces Genome Database (SGD) and the Rat Genome Database (RGD) were used to determine the molecular function of each protein (5; 28; 117). Q/N-rich regions were defined as 30 or more Q/N in an 80-amino acid stretch (137). For the cases in which data were not available, we developed a PERL-based algorithm to examine protein sequences for Q/N-rich regions (Appendix 3).

Long intrinsically-disordered (ID) domains were determined using the IUPred-L structural prediction algorithm; ID domains were defined as 30 or more amino acids with a disorder score of 0.5 or greater (50). To approximate the percentage of proteins in the yeast and rat genomes with long ID regions, 100 and 200 proteins respectively (~ twice the size of each sample data set; S1 and Appendix 2s), were randomly selected using a random number generator in alignment with the full proteomes of yeast (*S. cerevisiae*) and Rat (*R. norvegicus*) downloaded from uniprot (www.uniprot.org). Domains with intrinsic disorder were evaluated for all proteins by IUPred-L. Chi-Square 2x2 Fisher's

Exact test (Graphpad software) was used to determine if proteins with long ID domains (≥ 100 amino acids) were significantly enriched in polyQ aggregates. Analysis was also performed to ensure that the TAPI methodology was not biased towards identifying proteins that are abnormally large or abundant. The size (kDa) and cellular abundance (molecules/cell) was determined for each protein in the yeast sample set and compared with the whole proteome (values accumulated from (62; 64; 122; 144); Figure 4.).

Western Blotting

Western blotting of cell lysates (input) and TAPI-purified samples were used to verify high molecular weight protein aggregation (observed as large species that cannot migrate beyond the top of an acrylamide stacking gel) and confirm that specific proteins are trapped in polyQ aggregates. Standard Western blotting techniques were employed using nitrocellulose or PVDF membranes, which were probed with primary antibodies against the following targets at dilutions of $\sim 1:5000$: α GFP (Roche), α HA (Santa Cruz and Sigma), α Erk (Santa Cruz sc-93), α FUS (Bethyl), α hnRNPA1 (Cell Signaling), α RAD23B (Protein Tech), α TDP43 (Protein Tech), and α UBQLN2 (Novus Bio - 5f5). Appropriate HRP-conjugated secondary antibodies were used at 1:1000 dilutions, followed by HRP chemiluminescent substrate (Pierce ECL) for visualization.

Lysate Partitioning

Analysis of proteins (Bmh1p, Def1p, FUS, Ent2p and Sgt2p) entrapped within aggregates was performed by observing the fraction of protein in the total lysate, supernatant or pellet fraction that partitioned to the stacking well of an acrylamide gel under standard SDS electrophoresis conditions. Briefly, yeast lysates were prepared by mechanical breakage using glass beads in TAPI Buffer A (with RNase A). Pellet and

supernatant fractions were prepared as described in the TAPI methodology described above. Cellular fractions were subjected to SDS-PAGE using Any kD™ gels (Bio-Rad) followed by Western Blotting. The effectiveness of the TAPI buffer to eliminate RNA from the aggregates was examined by RNase-Free Agarose gel electrophoresis with and without nuclease treatment (Figure 3A). The FUS protein is prone to degradation following cell lysis, thus denaturing buffer (10 mM Tris, pH 7.5, 8 M urea) was used to visualize protein levels under conditions in which degradation is greatly inhibited. This enables confirmation that the lysates contain equivalent initial amounts of FUS.

Confocal Microscopy

Confocal microscopy slides were prepared on Poly-lysine coated slides with 1×10^6 cells/spot, fixed with 4% paraformaldehyde in 1x PBS and permeablized with 0.1% Triton X-100 in PBS. Primary antibodies for RAD23B (Bethyl) or FUS (Bethyl) were added at 1:200 dilution, followed by type-specific Alexa Fluor conjugated secondary antibodies (α -rabbit 647 and α -mouse 568, Southern Biotech) at 1:1000 dilution in 1% fetal calf serum in 1x PBS. Sample slides were mounted with DAPI-containing fluoramount (EMS) and viewed on a Zeiss 710 confocal laser scanning microscope and analyzed using Zen software (2009).

Thioflavin-T Analysis

Thioflavin-T (Th-T; Sigma) fluorescence was used to determine if HttQ103-GFP aggregates adopted an amyloid-like conformation. Purified Sup35-NM fibres (amyloid positive control), HttQ25-GFP, and HttQ103-GFP samples were treated with 1 μ g Th-T in 50mM Tris pH 8.0, 50mM NaCl buffer in a black 96-well plate. Samples were analyzed on a BioTek Synergy H1 plate reader using an excitation of 440 nm and emission at 490

nm. To determine if Thioflavin-T absorbance was significantly different between HttQ25-GFP and HttQ103-GFP, a two-tailed T-test analysis was used.

Determination of overlapping Htt-polyQ associated proteins from multiple studies

Htt-polyQ-associated proteins identified by TAPI (231), SILAC, (155) and antibody isolation with 2D gel analysis (228) were collated and overlapping proteins were identified. These overlapping hits were then compared to the four overexpression and deletion studies (66; 101; 132; 237). This cohort was further analyzed for molecular function, Q/N content, and ID domain. Gene Ontology and Saccharomyces Genome Database (SGD) were used to determine the molecular function of each protein (5; 28; 117). Q/N-rich regions were defined as 30 or more Q/N in an 80-amino acid stretch (137). For the cases in which data were not available, we developed a PEARL-based algorithm to examine protein sequences for Q/N-rich regions (Appendix 3).

Long intrinsically disordered (ID) domains were determined using the IUPred-L structural prediction algorithm; ID domains were defined as 30 or more amino acids with a disorder score of 0.5 or greater (50). To approximate the percentage of proteins in the yeast genome with long intrinsically-disordered (ID) domains, 100 proteins (~ twice the size of the sample data set; Appendix 1), were randomly selected using a random number generator in alignment with the full proteome of yeast (*S. cerevisiae*) downloaded from uniprot (www.uniprot.org). Domains with intrinsic disorder were evaluated for all proteins by IUPred-L. Chi-Square 2x2 Fisher's Exact test (Graphpad software) was used to determine if proteins with long ID domains (≥ 100 amino acids) were significantly enriched in polyQ aggregates.

APPENDICES

Appendix 1: Supplemental File S1 - Proteins identified by mass spectrometry following TAPI purification of polyglutamine aggregates from yeast cells.

Supplemental File S1.1: Biophysical Characterization of Htt-polyQ aggregate-associated proteins identified by Mass Spec analysis of TAPI samples in *S. cerevisiae*.

					Total Spectra		Total Spec Total Spectra			Total Spec Total Spectra			Total Spec
Identified Proteins (Binary Zof4)	Protein Function	SGD Name	Accession #	Mol Weight	Q103 (8/13)	Q25	Q103	Ratio (10/13)	Q25	Q103	Ratio (12/13)	Q25	
Akl1p	Ser-Thr protein kina: YBR059C	gi 6319533	124 kDa	6	2	2	4	2	0	8	#DIV/0!	0	
Api1p	Chaperone with a ro YNL077W	gi 6324252	59 kDa	10	0	0	2	#DIV/0!	0	18	#DIV/0!	0	
Bmh1p	14-3-3 protein, major YER177W	gi 398365115	30 kDa	2	0	0	5	#DIV/0!	0	4	#DIV/0!	0	
Cbk1p	Serine/threonine pro YNL161W	gi 398364551	87 kDa	2	0	0	3	#DIV/0!	0	2	#DIV/0!	0	
Ccr4p	CCR4-NOT is involv YAL021C	gi 6319298	95 kDa	2	0	0	28	#DIV/0!	0	12	#DIV/0!	0	
Cyc8p	General transcriptio YBR112C	gi 398364855	107 kDa	2	2	2	8	4	0	2	#DIV/0!	0	
Def1p	RNAPII degradation YKL054C	gi 6322796	84 kDa	2	0	0	10	#DIV/0!	0	7	#DIV/0!	2	
Dhh1p	Cytoplasmic DExD/f YDL160C	gi 6320041	58 kDa	2	0	0	4	#DIV/0!	0	4	#DIV/0!	0	
Eap1p	accelerates mRNA c YKL204W	gi 6322645	70 kDa	2	0	0	16	#DIV/0!	0	6	#DIV/0!	0	
Ent1p	Epsin-like protein in YDL161W	gi 6320039	52 kDa	2	0	0	10	#DIV/0!	0	7	#DIV/0!	2	
Ent2p	Epsin-like protein in YLR206W	gi 6323235	72 kDa	2	0	0	4	#DIV/0!	0	2	#DIV/0!	0	
Gal2p	Galactose permease YLR081W	gi 6323110	64 kDa	2	0	0	2	#DIV/0!	0	2	#DIV/0!	0	
Gts1p	Protein involved in A YGL181W	gi 6321257	44 kDa	2	0	0	4	#DIV/0!	0	2	#DIV/0!	0	
Hrp1p	RRM-containing het YOL123W	gi 6324449	60 kDa	2	0	0	3	#DIV/0!	0	2	#DIV/0!	0	
Ixr1p	Transcriptional repre YKL032C	gi 398364635	68 kDa	2	0	0	2	#DIV/0!	0	2	#DIV/0!	0	
Mbf1p	Transcriptional coac YOR298C-A	gi 398366153	16 kDa	2	0	0	3	#DIV/0!	0	2	#DIV/0!	0	
Mca1p	Ca2+-dependent cy: YOR197W	gi 398365705	48 kDa	2	0	0	2	#DIV/0!	0	2	#DIV/0!	0	
Mcm1p	Transcription factor; YMR043W	gi 6323686	33 kDa	2	0	0	2	#DIV/0!	0	2	#DIV/0!	0	
Mot3p	Transcriptional repre YMR070W	gi 6323715	54 kDa	2	0	0	2	#DIV/0!	0	2	#DIV/0!	0	
Mum2p	cytoplasmic protein; YBR057C	gi 6319531	41 kDa	2	0	0	6	#DIV/0!	0	2	#DIV/0!	0	
Nab3p	RNA-binding protein YPL190C	gi 6325066	90 kDa	2	0	0	44	#DIV/0!	0	2	#DIV/0!	0	
Nam8p	RNA binding protein YHR086W	gi 6321878	57 kDa	2	2	2	16	8	0	2	#DIV/0!	0	
New1p	ATP binding cassette YPL226W	gi 6325030	134 kDa	2	0	0	44	#DIV/0!	0	10	#DIV/0!	0	
Nrd1p	RNA-binding subunit YNL251C	gi 6324078	64 kDa	2	0	0	44	#DIV/0!	0	4	#DIV/0!	0	
Nup100p	FG-nucleoporin com YKL068W	gi 6322782	100 kDa	2	2	2	8	4	0	10	#DIV/0!	0	
Nup116p	FG-nucleoporin com YMR047C	gi 6323691	116 kDa	2	0	0	2	#DIV/0!	0	2	#DIV/0!	0	
Nup57p	FG-nucleoporin com YGR119C	gi 6321557	57 kDa	2	2	2	8	4	0	2	#DIV/0!	0	
Pan1p	Part of actin cytoske YIR006C	gi 398364573	160 kDa	6	2	0	8	4	0	26	#DIV/0!	0	
Pbp1p	Component of gluco YGR178C	gi 398366039	79 kDa	2	0	0	7	#DIV/0!	0	6	#DIV/0!	2	
Pgd1p	Subunit of the RNA YGL025C	gi 37362652	43 kDa	2	0	0	4	#DIV/0!	0	2	#DIV/0!	0	
Pin3p	Negative regulator o YPR154W	gi 6325412	24 kDa	2	0	0	2	#DIV/0!	0	2	#DIV/0!	0	
Pin4p	Contains an RNA rei YBL051C	gi 6319420	74 kDa	2	0	0	2	#DIV/0!	0	2	#DIV/0!	0	
Pop2p	RNase of the DEDD YNR052C	gi 398365813	50 kDa	4	0	0	2	#DIV/0!	0	9	#DIV/0!	0	
Puf3p	mRNA-binding prote YLL013C	gi 6323016	98 kDa	4	0	0	2	#DIV/0!	0	2	#DIV/0!	0	
Scd5p	Protein required for YOR329C	gi 398366267	97 kDa	4	0	0	2	#DIV/0!	0	6	#DIV/0!	0	
Sec24p	Component of the S YIL109C	gi 398364269	104 kDa	2	0	0	32	#DIV/0!	0	24	#DIV/0!	4	
Sgt2p	Glutamine-rich cyto YOR007C	gi 6324580	37 kDa	2	0	0	6	#DIV/0!	0	13	#DIV/0!	0	
Sla1p	Cytoskeletal protein YBL007C	gi 6319464	136 kDa	2	0	0	2	#DIV/0!	0	2	#DIV/0!	0	
Slim1p	Phosphoinositide PI- YIL105C	gi 6322086	78 kDa	2	0	0	8	#DIV/0!	0	2	#DIV/0!	0	
Sml1p	Ribonucleotide redu YML058W	gi 6323582	12 kDa	2	0	0	4	#DIV/0!	0	2	#DIV/0!	0	
Snf5p	Subunit of the SWI/ YBR289W	gi 398365913	103 kDa	2	0	0	4	#DIV/0!	0	8	#DIV/0!	0	
Srp54p	Interacts with the SF YPR088C	gi 6325345	60 kDa	6	0	0	2	#DIV/0!	0	3	#DIV/0!	0	
Taf5p	Subunit (90 kDa) of YBR198C	gi 6319675	89 kDa	2	0	0	3	#DIV/0!	0	3	#DIV/0!	0	
Tup1p	General repressor o YCR084C	gi 6319926	78 kDa	2	0	0	19	#DIV/0!	0	2	#DIV/0!	0	
Whi3p	RNA binding protein YNL197C	gi 6324132	71 kDa	2	0	0	3	#DIV/0!	0	2	#DIV/0!	0	
Yap1801p	Protein of the AP18 YHR161C	gi 6321955	72 kDa	2	0	0	8	#DIV/0!	0	2	#DIV/0!	0	
Yap1802p	Protein of the AP18 YGR241C	gi 398366269	64 kDa	2	0	0	2	#DIV/0!	0	2	#DIV/0!	0	
Ydj1p	Type I HSP40 co-ch YNL064C	gi 6324265	45 kDa	2	0	0	29	#DIV/0!	0	12	#DIV/0!	0	
YGR250C	Putative RNA bindin YGR250C	gi 398366299	90 kDa	2	0	0	25	#DIV/0!	0	2	#DIV/0!	0	
YLR177W	Putative protein of u YLR177W	gi 6323206	71 kDa	6	0	0	7	#DIV/0!	0	4	#DIV/0!	0	
YMR124W	Protein involved in s YMR124W	gi 6323772	106 kDa	6	0	0	7	#DIV/0!	0	4	#DIV/0!	0	
YNL208W	Protein of unknown YNL208W	gi 42742305	20 kDa	9	0	0	4	#DIV/0!	0	4	#DIV/0!	0	

otal Spectra

Q103	Ratio (4/14)	%Q	Q/N-Rich	ID AA	%ID	AA	PI	ID domain IUPred-L	ID ≥50	ID ≥100	ID ≥150	ID ≥200	RBP	RBP W/ID	PQC	
18	#DIV/0!	8.75	655-792	709	64%	1108	7.73	79, 620	79, 620	620	620	620				
4	#DIV/0!	4.92	None	48	9%	528	6.87	42							Yes	
51	#DIV/0!	6.74	None	55	21%	267	4.65	35							Yes	
20	#DIV/0!	12.57	143-302	86	11%	756	8.01	263			263	263				
23	#DIV/0!	7.29	9-144	268	32%	837	6.84	115, 59, 42	115, 59	115			Yes	Yes		
4	#DIV/0!	17.18	441-682	599	62%	966	4.76	40, 172, 371	172, 371	172, 371	172, 371	371	Yes	Yes		
34	#DIV/0!	23.17	335-719	610	83%	738	4.76	387, 41, 32, 59	387, 59	387	387	387			Yes	
7	#DIV/0!	8.1	None	117	23%	506	7.97	70	70				Yes	Yes		
15	#DIV/0!	7.44	None	553	88%	632	10.3	41, 85, 95, 295	85, 95, 295	295	295	295	Yes	Yes		
29	#DIV/0!	15.2	202-360	303	67%	454	5.66	99, 42, 71, 84	99, 71, 84							
22	11	24.31	214-599	450	73%	613	5.21	81, 299	81, 299	299	299	299				
12		3.31	None	41	7%	574	7.21	None								
16		12.12	278-359	116	29%	396	7.46	43, 39								
6		8.99	None	288	54%	534	5.36	156, 54, 148	156, 54, 148	156, 148	156		Yes	Yes		
12		20.77	1-383	502	84%	597	8.45	158, 199, 106	158, 199, 106	158, 199, 106	158, 199		Yes	Yes		
2		4.64	None	171	113%	151	11.35	106	106	106			Yes			
6		11.34	15-108	65	15%	432	5.09	128	128	128					Yes	
2		26.92	107-239	243	85%	286	4.86	207	207	207	207	207	Yes	Yes		
6		9.59	1-90, 70-204, 133-227	143	29%	490	9.2	169, 133, 50	169, 133	169, 133	169		Yes	Yes		
12		4.92	None	200	55%	366	5.84	97, 40	97							
27		10.35	646-779	609	76%	802	4.22	217, 83, 256	217, 83, 256	217, 256	217, 256	217, 256	Yes	Yes		
10	#DIV/0!	7.07	None	83	16%	523	8.45	None					Yes			
48	#DIV/0!	4.26	15-129	199	17%	1196	5.67	32, 34					Yes	Yes		
8	#DIV/0!	7.48	None	169	29%	575	9.96	103, 108	103, 108	103, 108			Yes	Yes		
2	#DIV/0!	6.36	258-393	741	77%	959	10.15	52, 120, 377, 46, 34, 50	52, 120, 377	120, 377	377	377				
18	#DIV/0!	9.88	264-345, 385-467, 394	389	35%	1113	10.13	96, 197, 392, 86	96, 197, 392, 86	197, 392	197, 392	392				
16	#DIV/0!	11.09	207-291	323	60%	541	10.39	260	260	260	260	260				
31	#DIV/0!	10.2	1-120	1097	74%	1480	5.02	231, 75, 51, 37, 36, 509	231, 75, 51, 509	231, 509	231, 509	231, 509				
16	8	3.19	None	473	66%	722	6.91	43, 45, 90, 88, 30, 42	90, 88				Yes	Yes		
2		10.33	272-359	274	69%	397	9.13	256	256	256	256	256				
6		13.49	None	126	59%	215	7.55	88	88							
8		10.03	415-508	465	70%	668	6.21	78, 93, 32, 161, 34	78, 93, 161	161	161		Yes	Yes		
7		14.55	11-162	123	28%	433	6.11	82, 38	82				Yes	Yes		
24		7.74	379-480	383	44%	879	7.21	146, 198	146, 198	146, 198	198		Yes	Yes		
6		11.93	598-695, 734-839	307	35%	872	7.94	53, 45, 36, 476	53, 476	476	476	476				
34		#DIV/0!	6.26	None	40	4%	926	6.16	134	134	134					Yes
65		4.62	None	164	47%	346	4.51	48, 81	81							Yes
15		#DIV/0!	6.67	None	812	65%	1244	5.69	124, 71, 40, 40, 143, 34, 84, 77	124, 71, 143, 84, 77	124, 143					Yes
8		#DIV/0!	11.08	1-129	271	40%	686	8.16	117, 81	117, 81	117					
6	#DIV/0!	8.65	None	76	73%	104	4.55	None								
8	#DIV/0!	16.69	1-95, 113-330	460	51%	905	8.49	131, 31, 54, 46, 110	131, 54, 110	131, 110		Yes	Yes			
42	#DIV/0!	7.39	None	147	27%	541	9.04	71	71				Yes	Yes		
11	#DIV/0!	4.39	None	256	32%	798	7.39	134, 56	134, 56	134			Yes	Yes	Yes	
4	#DIV/0!	8.56	None	326	46%	713	5.47	192, 36	192	192	192		Yes	Yes	Yes	
14	#DIV/0!	7.41	None	347	52%	661	8.53	42, 125, 146, 32, 35, 41	125, 146	125, 146			Yes	Yes		
2	#DIV/0!	11.46	484-595	307	48%	637	7.04	73, 200	73, 200	200	200	200				
8	#DIV/0!	12.15	None	293	52%	568	6.77	50, 91, 86	91, 86							
22	#DIV/0!	3.42	None	162	40%	409	6.23	None							Yes	
16	#DIV/0!	5.38	None	233	30%	781	5.1	41, 32					Yes	Yes		
4	#DIV/0!	7.17	125-224	191	30%	628	9.38	59, 63	59, 63							
16	#DIV/0!	8.27	17-139, 301-383	515	55%	943	6.88	143, 37, 345, 70, 45, 33, 32	143, 345, 70	143, 345	345	345				
36	#DIV/0!	11.06	None	168	84%	199	7.9	53, 124	53, 124	124						
		8.62	54%		48%	686	7.28	48/52 = 92%	43/52 = 83%	33/52 = 63%	22/52 = 42%	16/52 = 31%	46%	96%	19%	

Appendix 2: Supplemental File S2 - Proteins identified by mass spectrometry following TAPI purification of polyglutamine aggregates from rat cells.

Supplemental File S2.1: Biophysical Characterization of Htt-polyQ aggregate-associated proteins identified by Mass Spec analysis of TAPI samples in *R. norvegicus*.

Note: Rows highlighted in blue are disease-linked proteins.

Notes: Rows highlighted in blue are disease-linked proteins.		Total Sp		Total Sp Total Spectra			Total Sp Total Spectra			Total Sp Total Spectra			Q/N-rich			ID Domain >30	IDD >50	IDD >100	IDD >150	IDD >200	RBP	RBP W/ID	PQC	
Identified PnProtein Function	NCBI Ref Seq	Accession #	Mol Weight	Q74 (Ap)	Q23	Q74	Ratio (L)	Q23Q	74Q	Ratio (A)	Q23Q	74Q	Ratio (L)	Q23Q	74Q									Ratio (A)
Aars2	alanine--RNA ligase, mitochondrial (NP_061433.2)	gi 157818067 (+2)	108 kDa		0	2	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	0	14	#DIV/0!	None	209, 81, 50	209, 81, 50	209	209			
Acad9	acyl-CoA dehydrogenase family n (NP_001030123.1)	gi 197313734 (+1)	69 kDa		0	6	#DIV/0!	0	22	#DIV/0!	0	22	#DIV/0!	0	22	#DIV/0!	None	None	None	None	None	yes		
Acst2	acyl-CoA synthetase family mem (NP_113896.1)	gi 77993368	68 kDa		0	4	#DIV/0!	0	6	#DIV/0!	0	6	#DIV/0!	0	4	#DIV/0!	None	None	None	None	None			
Adrm1	proteasomal ubiquitin receptor ACXP_006241116.1	gi 13928990 (+1)	42 kDa		0	9	#DIV/0!	2	6	3	0	6	3	0	6	#DIV/0!	None	None	None	None	None			yes
Akap8l	PREDICTED: A kinase (PRKA) a (XP_006241116.1)	gi 564358445 (+2)	68 kDa		0	2	#DIV/0!	0	2	#DIV/0!	0	2	#DIV/0!	0	2	#DIV/0!	None	119, 100	119, 100	119, 100	119, 100	yes	yes	
Arbp4	ADP-ribosylation factor-like protein (NP_001020801.1)	gi 71043808	26 kDa		0	4	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	None	199	199	199	199	yes	yes	
Asap1	PREDICTED: arf-GAP with SH3 (NP_075581.1)	gi 564307456 (+2)	109 kDa	5	0	2	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	None	364	364	364	364			
Atp5a1	ATP synthase subunit alpha, mito (XP_008756735.1)	gi 40533742	60 kDa		0	2	#DIV/0!	0	2	#DIV/0!	4	20	5.00	0	20	5.00	None	None	None	None	None	yes		
Atrx	PREDICTED: transcriptional regul (XP_008756735.1)	gi 564322975 (+10)	279 kDa		0	13	#DIV/0!	0	5	#DIV/0!	0	5	#DIV/0!	0	5	#DIV/0!	None	126	126	126	126	yes		
Clint1	PREDICTED: clathrin interactor 1 (NP_001002022.1)	gi 564371473	68 kDa		0	60	#DIV/0!	0	18	#DIV/0!	0	91	#DIV/0!	0	91	#DIV/0!	None	124, 35, 46	124	124	124			
Ctic	PREDICTED: clathrin heavy chain (XP_006247169.1)	gi 564373736 (+3)	192 kDa		0	5	#DIV/0!	0	11	#DIV/0!	0	11	#DIV/0!	0	11	#DIV/0!	None	None	None	None	None			yes
Cnn2	PREDICTED: calponin-2-like [Rar] (XP_006241100.1)	gi 564358403	29 kDa		0	2	#DIV/0!	0	3	#DIV/0!	0	3	#DIV/0!	0	5	#DIV/0!	None	None	None	None	None			yes
Daf2	PREDICTED: protein DDH1 homo (XP_006239367.1)	gi 564353874 (+3)	106 kDa		0	4	#DIV/0!	0	3	#DIV/0!	0	5	#DIV/0!	0	4	#DIV/0!	None	None	None	None	None			yes
Ddx5	probable ATP-dependent RNA he (NP_001076141.1)	gi 56090441	69 kDa	2	0	3	#DIV/0!	0	5	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	None	38, 36				yes	yes	
Dnaj2	dnaJ homolog subfamily A memb (NP_114468.2)	gi 56799412	46 kDa		0	24	#DIV/0!	0	10	#DIV/0!	0	18	#DIV/0!	0	18	#DIV/0!	None	62	62	209	209			yes
Dnaj4	dnaJ homolog subfamily A memb (NP_001020582.1)	gi 70794764	62 kDa		0	4	#DIV/0!	0	2	#DIV/0!	0	2	#DIV/0!	0	2	#DIV/0!	None	209	209	209	209			yes
Dnajb1	dnaJ homolog subfamily B memb (NP_001019111.1)	gi 157823165 (+1)	42 kDa		0	3	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	None	30						yes
Dnajc7	dnaJ homolog subfamily C memb (NP_998790.1)	gi 47155561 (+3)	57 kDa		0	2	#DIV/0!	0	2	#DIV/0!	0	5	#DIV/0!	0	5	#DIV/0!	None	None	None	None	None			yes
Dync1h1	cytoplasmic dynein 1 heavy chain (NP_062099.3)	gi 148491097	532 kDa		0	8	#DIV/0!	0	8	#DIV/0!	0	84	#DIV/0!	0	84	#DIV/0!	None	None	None	None	None	yes		
Eif4g1	PREDICTED: eukaryotic translati (NP_006249843.1)	gi 654314679 (+11)	176 kDa		0	11	#DIV/0!	0	3	#DIV/0!	0	3	#DIV/0!	0	4	#DIV/0!	None	91, 356, 75, 83, 30, 247	91, 356, 75, 83, 247	356, 247	356, 247	356, 247	yes	yes
Eif4g2	eukaryotic translation initiation fac (NP_001017374.2)	gi 153267461	102 kDa		0	11	#DIV/0!	0	3	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	None	35, 54, 105	54, 105	105	105	yes	yes	
Ep300	PREDICTED: histone acetyltrans (XP_001076610.1)	gi 109482524 (+1)	264 kDa		0	20	#DIV/0!	0	2	#DIV/0!	0	7	#DIV/0!	0	7	#DIV/0!	None	230, 58, 38, 48, 50, 413, 65, 257	230, 58, 50, 413, 65, 257	230, 413, 257	230, 413, 257	230, 413, 257	yes	yes
Etib	electron transfer flavoprotein sub (NP_001004220.1)	gi 51948412	28 kDa		0	2	#DIV/0!	0	2	#DIV/0!	9	40	4.44	0	9	40	None	None	None	None	None			
Fasn	fatty acid synthase [Rattus norveg (NP_059028.1)	gi 8394158	273 kDa		7	0	17	#DIV/0!	0	3	#DIV/0!	0	12	#DIV/0!	0	12	#DIV/0!	None	None	None	None	None	yes	
Fus	PREDICTED: RNA-binding protein (NP_006230352.1)	gi 564331043 (+1)	63 kDa		1	28	28	0	7	#DIV/0!	0	23	#DIV/0!	0	23	#DIV/0!	None	284, 71, 87	284	284	284	yes	yes	
Garg2	PREDICTED: PERQ amino acid (XP_003754561.1)	gi 302342451 (+3)	151 kDa		0	5	#DIV/0!	0	5	#DIV/0!	0	5	#DIV/0!	0	5	#DIV/0!	None	50, 33, 259, 52, 195, 38, 39	50, 259, 52, 195	259, 195	259, 195	yes	yes	
Gls	glutaminase kidney isoform, mito (NP_036701.2)	gi 158303294 (+1)	74 kDa		0	8	#DIV/0!	0	8	#DIV/0!	0	8	#DIV/0!	0	8	#DIV/0!	None	30, 74	74					
Gnao1	guanine nucleotide-binding protein (NP_059023.1)	gi 8394152	40 kDa		0	2	#DIV/0!	0	12	#DIV/0!	0	12	#DIV/0!	0	12	#DIV/0!	None	None	None	None	None			
Hadha	trifunctional enzyme subunit alpha (NP_570839.2)	gi 148747393	83 kDa		0	4	#DIV/0!	0	4	#DIV/0!	12	36	3.00	0	12	36	None	None	None	None	None			
Hnrnp3a	heterogeneous nuclear ribonucleo (NP_00104764.1)	gi 162329577 (+3)	40 kDa		0	13	#DIV/0!	0	4	#DIV/0!	0	10	#DIV/0!	0	10	#DIV/0!	None	39				yes	yes	
Hnrnpf	heterogeneous nuclear ribonucleo (NP_071782.1)	gi 257425737 (+1)	46 kDa	2	0	7	#DIV/0!	0	2	#DIV/0!	0	13	#DIV/0!	0	13	#DIV/0!	None	None	None	None	None	yes		
Hnrnp2	PREDICTED: heterogeneous nuc (XP_006257281.1)	gi 564399943 (+3)	49 kDa		0	3	#DIV/0!	0	4	#DIV/0!	0	27	#DIV/0!	0	27	#DIV/0!	None	None	None	None	None			
Hnrnmp	heterogeneous nuclear ribonucleo (NP_446328.2)	gi 158186696 (+1)	74 kDa		0	3	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	None	58, 30, 31	58			yes	yes	
Hnrnpu	heterogeneous nuclear ribonucleo (NP_476840.2)	gi 148747541 (+2)	88 kDa		0	8	#DIV/0!	0	6	#DIV/0!	0	6	#DIV/0!	0	6	#DIV/0!	None	None	None	None	None	yes		
Hsp90aa1	heat shock protein HSP 90-alpha (NP_786937.1)	gi 28467005	85 kDa		0	10	#DIV/0!	0	37	#DIV/0!	0	37	#DIV/0!	0	37	#DIV/0!	None	55, 30	55			yes	yes	yes
Hsp8a	heat shock cognate 71 kDa prote (NP_077327.1)	gi 13242237	71 kDa	6	0	31	#DIV/0!	4	11	2.75	0	20	#DIV/0!	0	20	#DIV/0!	None	46, 36				yes		yes
Irb3	isolate dehydrogenase [NAD(P)N (NP_001014163.1)	gi 60709055	51 kDa		0	2	#DIV/0!	0	4	#DIV/0!	14	32	2.29	0	14	32	None	None	None	None	None			
Irb3b	isolate dehydrogenase [NAD] (NP_446033.1)	gi 56268203 (+2)	42 kDa		0	2	#DIV/0!	0	2	#DIV/0!	0	14	#DIV/0!	0	14	#DIV/0!	None	None	None	None	None			
Kprp	keratinocyte proline-rich protein [(NP_001002290.1)	gi 50511334 (+1)	76 kDa	6	0	12	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	None	None	None	None	None			
Lap3	cytosol aminopeptidase [Rattus n (NP_001011910.1)	gi 58865398	56 kDa		0	2	#DIV/0!	0	2	#DIV/0!	0	8	#DIV/0!	0	8	#DIV/0!	None	None	None	None	None			
Ldha	PREDICTED: L-lactate dehydrogen (NP_058721.1)	gi 392341842 (+3)	37 kDa		0	2	#DIV/0!	0	2	#DIV/0!	0	2	#DIV/0!	0	2	#DIV/0!	None	None	None	None	None			
Maged1	melanoma-associated antigen D1NP_445861.1	gi 16758144	86 kDa		0	12	#DIV/0!	0	2	#DIV/0!	0	2	#DIV/0!	0	2	#DIV/0!	None	35, 40, 30						yes
Matf3	PREDICTED: matrix-3 isoform X (XP_006254966.1)	gi 564392976 (+2)	39 kDa		0	6	#DIV/0!	0	6	#DIV/0!	0	6	#DIV/0!	0	6	#DIV/0!	None	31, 57, 42, 119, 72	354	354	354			
Mif2	myeloid leukemia factor 2 [Rattus (NP_001013359.1)	gi 157823873 (+2)	28 kDa	2	2	6	3	0	4	#DIV/0!	0	4	#DIV/0!	0	4	#DIV/0!	None	57, 119, 72	119	119	119	yes	yes	
Myo1d	unconventional myosin-Ib [Rattus (NP_037115.2)	gi 56799396	116 kDa		0	2	#DIV/0!	0	2	#DIV/0!	0	6	#DIV/0!	0	6	#DIV/0!	None	438	136	136	136			
Ndufs7	PREDICTED: NADH dehydrogen (NP_001008525.1)	gi 564358240 (+2)	25 kDa		0	2	#DIV/0!	0	2	#DIV/0!	0	5	#DIV/0!	0	5	#DIV/0!	None	None	None	None	None			
Nono	PREDICTED: non-POU domain-c (NP_001012356.1)	gi 564399550 (+1)	55 kDa		0	8	#DIV/0!	0	6	#DIV/0!	0	6	#DIV/0!	0	6	#DIV/0!	None	66	66			yes	yes	
Nr3c1	glucocorticoid receptor [Rattus n (NP_036708.2)	gi 158303300	87 kDa		0	4	#DIV/0!	0	2	#DIV/0!	0	2	#DIV/0!	0	2	#DIV/0!	None	146	146	146	146			
Nsf1c	NSF1 cofactor p4 [Rattus norveg (NP_114187.1)	gi 140100371 (+1)	41 kDa		0	20	#DIV/0!	0	2	#DIV/0!	0	2	#DIV/0!	0	2	#DIV/0!	None	76, 45	76					
Ogdh	PREDICTED: 2-oxoglutarate dehy (NP_006251508.1)	gi 564384801 (+3)	116 kDa		0	9	#DIV/0!	0	4	#DIV/0!	15	71	4.73	0	15	71	None	None	None	None	None			
Pcbp1	PREDICTED: poly(rC)-binding pr (XP_008774006.1)	gi 564304213 (+1)	41 kDa	2	0	14	#DIV/0!	0	8	#DIV/0!</														

Appendix 3: Supplemental File S3 - PEARL-based algorithm for examining protein sequences for Q/N-rich regions.

QSEARCH(1)

User Contributed Perl Documentation

QSEARCH(1)

NAME

Qsearch – Amino acid sequence linear-search program

SYNOPSIS

qsearch [**options**] *filename*

DESCRIPTION

A program to analyze Q occurrence in protein sequences contained in the FASTA file named *filename*. Valid amino acid sequences are processed one at a time starting at the top of the FASTA file.

Each sequence is broken up into a moving sub-domain containing the number of amino acids specified by the **--length** option. Domains are created sequentially starting with the first amino acid in the sequence and moving linearly in a manner analogous to a moving average. Domains with a smaller than requested length at the end of the sequence are not analyzed.

OPTIONS

Command line options are processed from left to right and may be specified more than once. If conflicting options are specified, later options override earlier ones.

SHORT FORM

-a *letters*

-aletters

Equivalent to **--aminoacids** *letters*

-f *num*

-fnum

Equivalent to **--frequency** *num*

-h

Equivalent to **--help** and can be bundled.

-v [*num*]

-vnum

Equivalent to **--verbose** [*num*]

LONG FORM

--aminoacids *letters*

Set the list of amino acid one-letter abbreviations. This list is used in the regular expression to parse each sequence in the FASTA file. If this list does not include all possible amino acid one-letter codes contained in the file, the parsing will be inaccurate.

The default value of *letters* is ABCDEFGHIJKLMNOPQRSTUVWXYZ.

--debug [*num*]

--verbose [*num*]

Set the debug print level to *num*. The default value of *num* is one.

--frequency *num*

Set the frequency threshold for the occurrence of desired amino acids. The default value of *num* is 30.

--help

Print full qsearch documentation via perldoc then exit.

--length *num*

Set the amino acid domain length. The default value of *num* is 80.

--version

Report version number then exit.

NOTES

The FASTA file must be free of blank lines and have proper newline characters for the operating system this program is running upon. These conversions can be achieved with programs such as dos2unix, unix2dos or PERL.

This program relies on a regular expression to extract a sequence from each FASTA entry. The regular expression was built to work with the FASTA files associated with the corresponding research, rather than

an extensive test set of FASTA files. Consequently said regular expression might require modification to read your FASTA file.

Regular expression manipulation is nontrivial. Seek the assistance of an expert before modifying the qsearch.pl file.

OUTPUT

For each valid sequence in the FASTA file, the following items are returned as output.

- *FASTA Entry*
The unwrapped, valid FASTA entry currently being processed
- Protein: *sequence*
The extracted sequence from the FASTA entry
- Number of Amino Acids: *integer*
Total number of amino acids in the extracted sequence
- Q total: *integer*
Total number of glutamines in the sequence
- Percent Q: *decimal %*
Percentage of glutamines in the sequence as computed by (Q total)/(Amino acid total)
- Z total: *integer*
Total number of glutamine or glutamic acids in the sequence
- Percent QZ: *decimal %*
Percentage of glutamines and glutamic acids in the sequence as computed by [(Q total) + (Z total)]/(Amino acid total)

For each domain in the sequence containing a number of Q, Z, N, and B amino acids greater than or equal to the value specified with the **--frequency** option, the following items are also returned as output.

- *length* Amino Acid Regions
Header noting the start of domain analysis. *length* is the value set with the **--length** option.
- *first – last*
Integer index values indicating the beginning and ending amino acids from the sequence defining the current domain
- *domain*
Amino acid sequence for the current domain
- Q total: *integer*
Total number of glutamines in the domain
- Z total: *integer*
Total number of glutamine or glutamic acids in the domain
- N total: *integer*
Total number of asparagines in the domain
- B total: *integer*
Total number of asparagine or aspartic acids in the domain
- QZNB total: *integer*
Total number of Q, Z, N or B amino acids in the domain

EXAMPLES

These examples assume that qsearch has been aliased to qsearch.pl. Since qsearch reports output to the screen, it may be helpful to pipe output into a new file.

```
qsearch rat > rat.out
```

Search the sequences contained in the file rat returning the usual information and piping it to a new file called rat.out.

```
qsearch -f 10 file.fasta
```

Search the sequences contained in file.fasta reporting the usual output information for all domains containing at least 10 occurrences of Q, Z, N or B.

```
qsearch -l 50 fasta.txt
```

Search the sequences contained in fasta.txt reporting the usual output information for domains containing 50 amino acids.

```
qsearch -h
```

Print the full qsearch documentation and then exit.

```
qsearch --version
```

Print the software version number for qsearch and then exit.

INSTALLATION

To install qsearch, just place qsearch.pl in a folder and make it executable. Then execute the program passing the necessary options and file on the command line.

On UNIX systems such as Mac OS X, use the following command in Terminal to make a file executable.

```
chmod a+x qsearch.pl
```

Then use the following command from within the folder containing qsearch.pl to execute the program and return the version in Terminal.

```
./qsearch.pl --version
```

Other operating systems follow a similar procedure to impart executability. See Google or your neighborhood programmer for more syntax details. :-)

This program was written and tested in PERL version 5.18.2 on a Mac but should be backward compatible with previous PERL versions on other operating systems.

VERSION

2.0

PROGRAMMER

Dr. Jason L. Sonnenberg, <sonnenberg.11@osu.edu>

COPYRIGHT

This utility is free software; you can redistribute it and/or modify it under the same terms as PERL itself. I would like to hear of any suggestions for improvement.

Appendix 4: Supplemental File S4 - Proteins identified by meta-analysis of Htt-polyQ associated protein studies in yeast.

Supplemental File 4: Overlapping Htt-polyQ aggregate-associated proteins between TAPI, SILAC (Park 2013) and 2D gel analysis (Wang 2008) in yeast. Bio overlapping protein hits including ID domains and molecular function.

SGD Name	TAPI	Park 2013	Wang 2007	Mason 2013	Willingham 2003	Kayatekin 2014	Giorgini 2005	Protein Function	SGD Name	%Q
1 YER177W	Bmh1p	BMH1	BMH1					14-3-3 protein, major isoform	YER177W	6.74
2 YDR171W		HSP42	HSP42						YDR171W	
3 YCR012W	Pgk1		PGK1						YCR012W	
4 YPR154W	Pin3p	PIN3	PIN3					Negative regulator of actin nucleation	YPR154W	13.49
5 YOR007C	Sgt2p	SGT2	SGT2					Glutamine-rich cytoplasmic cochaperone	YOR007C	4.62
6 YNL007C	Sis1p	SIS1	SIS1			Sis1			YNL007C	
7 YAL005C	SSA1	SSA1	SSA1						YAL005C	
8 YLL024C		SSA2	SSA2						YLL024C	
9 YNL208W	YNL208W	YNL208W	YNL208W					Protein of unknown function; may interact with	YNL208W	11.06
10 YFL039C	Act1	ACT1							YFL039C	
11 YDR099W	Bmh2p	BMH2							YDR099W	
12 YAL021C	Ccr4p	CCR4						CCR4-NOT is involved in regulation of ζ	YAL021C	7.29
13 YBR112C	Cyc8p	CYC8						General transcriptional co-repressor; co	YBR112C	17.18
14 YKL054C	Def1p	DEF1					def1	RNAPII degradation factor	YKL054C	23.17
15 YDL161W	Ent1p	ENT1						Epsin-like protein involved in endocytosis	YDL161W	15.2
16 YDL226C	Gcs1p	GCS1		GCS1				ADP-ribosylation factor GTPase activator	YDL226C	6.25
17 YLL026W	HSP104						hsp104		YLL026W	
18 YKL032C	Ixr1p	IXR1						Transcriptional repressor that regulates	YKL032C	20.77
19 YOR197W	Mca1p	MCA1						Ca ²⁺ -dependent cysteine protease; req	YOR197W	11.34
20 YPL190C	Nab3p	NAB3						RNA-binding protein, subunit of Nrd1	cc YPL190C	7.07
21 YKL068W	Nup100p	NUP100						FG-nucleoporin component of central α	YKL068W	6.36
22 YMR047C	Nup116p	NUP116						FG-nucleoporin component of central α	YMR047C	9.88
23 YIR006C	Pan1p	PAN1						Part of actin cytoskeleton-regulatory complex	YIR006C	10.2
24 YGR178C	Pbp1p	PBP1						Component of glucose deprivation induced	YGR178C	3.19
25 YDL053C	Pbp4p	PBP4		PBP4				Pbp1p binding protein; interacts strongly	YDL053C	7.57
26 YNL016W	Pub1p	PUB1						Poly (A) ⁺ RNA-binding protein; abundant	YNL016W	12.8
27 YLL013C	Puf3p	PUF3						mRNA-binding protein: Protein of the mi	YLL013C	7.74
28 YGL014W	Puf4p	PUF4						preferentially binds mRNAs encoding nu	YGL014W	4.84
29 YCL028W	Rnq1p	RNQ1					rnq1		YCL028W	
30 YIL109C	Sec24p	SEC24						Component of the Sec23p-Sec24p hetero	YIL109C	6.26
31 YBL007C	Slp1p	SLA1						Cytoskeletal protein binding protein; req	YBL007C	6.67
32 YHR030C	Slf2p	SLT2						Serine/threonine MAP kinase; involved i	YHR030C	9.5
33 YPR088C	Srp54p	SRP54						Signal recognition particle (SRP) subun	YPR088C	7.39
34 YDR172W	Sup35p	SUP35				Sup35			YDR172W	
35 YBR198C	Taf5p	TAF5						Subunit (90 kDa) of TFIID and SAGA co	YBR198C	4.39
36 YNL064C	Ydj1p	YDJ1						Type I HSP40 co-chaperone; involved in	YNL064C	3.42
37 YGR250C	YGR250C	YGR250C						Putative RNA binding protein; localizes	YGR250C	5.38
38 YMR124W	YMR124W	YMR124W						Protein involved in septin-ER tethering;	YMR124W	8.27
										9.22

physical characterization of

AA	ID domain IUPred-L	ID >50	ID >100	ID >150	ID >200	RBP	RBP_W/ID	>100	>150	>200	PQC/Chaperone
267	35										Y
	52, 166	52, 166	166	166							Y
	None					Y	None				
215	88	88									
346	48, 81	81									Y
	39, 34										Y
	30, 45, 42										Y
	30, 47, 42										Y
199	53, 124	53, 124	124								
	None					y	None				
	41										Y
837	115, 59, 42	115, 59	115			Y	115, 59, 42	115			
966	40, 172, 371	172, 371	172, 371	172, 371	371	Y	40, 172, 371	172, 371	172, 371	371	
738	387, 41, 32, 59	387, 59	387	387	387						Y
454	99, 42, 71, 84	99, 71, 84									
352	105, 33,40	105	105								
	30										Y
597	158, 199, 106	158, 199, 106	158, 199, 106	158, 199		Y	158, 199, 106	158, 199, 106	158, 199		
432	128	128	128								Y
802	217, 83, 256	217, 83, 256	217, 256	217, 256	217, 256	Y	217, 83, 256	217, 256	217, 256	217, 256	
959	52, 120, 377, 46, 34, 50	52, 120, 377	120, 377	377	377						
1113	96, 197, 392, 86	96, 197, 392, 86	197, 392	197, 392	392						
1480	231, 75, 51, 37, 36, 509	231, 75, 51, 509	231, 509	231, 509	231, 509						
722	43, 45, 90, 88, 30, 42	90, 88				Y	45, 90, 88, 30, 42				Y
185	122	122	122								
453	67, 119, 46	67, 119	119			Y	67, 119, 46	119			Y
879	146, 198	146, 198	146, 198	198		Y	146, 198	146, 198	198		
888	70, 35, 33, 46, 36, 201	70, 201	201	201	201	Y	35, 33, 46, 36,	201	201	201	
	41, 227	227	227	227	227						
926	134	134	134								
1244	124, 71, 40, 40, 143, 34, 84, 77 4, 71, 143, 84,		124, 143								Y
484	106	106	106			Y	106	106			
541	71	71				Y	71				
	37, 199	199	199	199		Y	37, 199	199	199		
798	134, 56	134, 56	134			Y	134, 56	134			Y
409	None										Y
781	41, 32					Y	41, 32				
943	143, 37, 345, 70, 45, 33, 32	143, 345, 70	143, 345	345	345						
	92%	74%	61%	34%	24%		87%	67%	40%	20%	
	31%	16%	9%	6%	0%	Proteome RE	53%	17%	10%	3%	
P=	0.0001	0.0001	0.0001		0.0001		0.0146	0.0001	0.0038	0.0181	

REFERENCES

1. Aguzzi A, Calella AM. 2009. Prions: protein aggregation and infectious diseases. *Physiological reviews* 89:1105-52
2. Ambadipudi S, Zweckstetter M. 2015. Targeting intrinsically disordered proteins in rational drug discovery. *Expert opinion on drug discovery*:1-13
3. Antzutkin ON, Balbach JJ, Leapman RD, Rizzo NW, Reed J, Tycko R. 2000. Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of beta-sheets in Alzheimer's beta-amyloid fibrils. *Proceedings of the National Academy of Sciences of the United States of America* 97:13045-50
4. Arrasate M, Mitra S, Schweitzer ES, Segal MR, Finkbeiner S. 2004. Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* 431:805-10
5. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics* 25:25-9
6. Baltz AG, Munschauer M, Schwanhauser B, Vasile A, Murakawa Y, et al. 2012. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Molecular cell* 46:674-90
7. Basso M, Massignan T, Samengo G, Cheroni C, De Biasi S, et al. 2006. Insoluble mutant SOD1 is partly oligoubiquitinated in amyotrophic lateral sclerosis mice. *The Journal of biological chemistry* 281:33325-35
8. Baxa U. 2008. Structural basis of infectious and non-infectious amyloids. *Curr Alzheimer Res* 5:308-18
9. Bence NF, Sampat RM, Kopito RR. 2001. Impairment of the ubiquitin-proteasome system by protein aggregation. *Science* 292:1552-5
10. Bergink S, Severijnen LA, Wijgers N, Sugasawa K, Yousaf H, et al. 2006. The DNA repair-ubiquitin-associated HR23 proteins are constituents of neuronal inclusions in specific neurodegenerative disorders without hampering DNA repair. *Neurobiology of disease* 23:708-16
11. Bernstein SL, Dupuis NF, Lazo ND, Wytenbach T, Condrón MM, et al. 2009. Amyloid-beta protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. *Nature chemistry* 1:326-31
12. Bertram PG, Zeng C, Thorson J, Shaw AS, Zheng XF. 1998. The 14-3-3 proteins positively regulate rapamycin-sensitive signaling. *Current biology : CB* 8:1259-67
13. Bodner RA, Outeiro TF, Altmann S, Maxwell MM, Cho SH, et al. 2006. Pharmacological promotion of inclusion formation: a therapeutic approach for Huntington's and Parkinson's diseases. *Proceedings of the National Academy of Sciences of the United States of America* 103:4246-51
14. Bonini NM. 2002. Chaperoning brain degeneration. *Proceedings of the National Academy of Sciences of the United States of America* 99 Suppl 4:16407-11
15. Bruckmann A, Hensbergen PJ, Balog CI, Deelder AM, de Steensma HY, van Heusden GP. 2007. Post-transcriptional control of the *Saccharomyces cerevisiae* proteome by 14-3-3 proteins. *Journal of proteome research* 6:1689-99

-
16. Bucciantini M, Giannoni E, Chiti F, Baroni F, Formigli L, et al. 2002. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* 416:507-11
 17. Butland SL, Devon RS, Huang Y, Mead CL, Meynert AM, et al. 2007. CAG-encoded polyglutamine length polymorphism in the human genome. *BMC genomics* 8:126
 18. Buxbaum JN, Linke RP. 2012. A molecular history of the amyloidoses. *Journal of molecular biology* 421:142-59
 19. Calvo SE, Mootha VK. 2010. The mitochondrial proteome and human disease. *Annual review of genomics and human genetics* 11:25-44
 20. Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, et al. 2012. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149:1393-406
 21. Cha JH. 2007. Transcriptional signatures in Huntington's disease. *Progress in neurobiology* 83:228-48
 22. Chartron JW, Gonzalez GM, Clemons WM, Jr. 2011. A structural model of the Sgt2 protein and its interactions with chaperones and the Get4/Get5 complex. *The Journal of biological chemistry* 286:34325-34
 23. Chen S, Berthelie V, Hamilton JB, O'Nuallain B, Wetzel R. 2002. Amyloid-like features of polyglutamine aggregates and their assembly kinetics. *Biochemistry* 41:7391-9
 24. Chen S, Berthelie V, Yang W, Wetzel R. 2001. Polyglutamine aggregation behavior in vitro supports a recruitment mechanism of cytotoxicity. *Journal of molecular biology* 311:173-82
 25. Chen S, Ferrone FA, Wetzel R. 2002. Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. *Proceedings of the National Academy of Sciences of the United States of America* 99:11884-9
 26. Chen-Plotkin AS, Lee VM, Trojanowski JQ. 2010. TAR DNA-binding protein 43 in neurodegenerative disease. *Nat Rev Neurol* 6:211-20
 27. Chernova TA, Romanyuk AV, Karpova TS, Shanks JR, Ali M, et al. 2011. Prion induction by the short-lived, stress-induced protein Lsb2 is regulated by ubiquitination and association with the actin cytoskeleton. *Molecular cell* 43:242-52
 28. Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, et al. 2012. Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic acids research* 40:D700-5
 29. Chiti F, Dobson CM. 2006. Protein misfolding, functional amyloid, and human disease. *Annual review of biochemistry* 75:333-66
 30. Chow MK, Lomas DA, Bottomley SP. 2004. Promiscuous beta-strand interactions and the conformational diseases. *Current medicinal chemistry* 11:491-9
 31. Clapp C, Portt L, Khoury C, Sheibani S, Norman G, et al. 2012. 14-3-3 protects against stress-induced apoptosis. *Cell death & disease* 3:e348
 32. Cooper JK, Schilling G, Peters MF, Herring WJ, Sharp AH, et al. 1998. Truncated N-terminal fragments of huntingtin with expanded glutamine repeats form nuclear and cytoplasmic aggregates in cell culture. *Human molecular genetics* 7:783-90

-
33. Corsaro A, Thellung S, Villa V, Nizzari M, Florio T. 2012. Role of prion protein aggregation in neurotoxicity. *International journal of molecular sciences* 13:8648-69
 34. Couthouis J, Hart MP, Shorter J, DeJesus-Hernandez M, Erion R, et al. 2011. A yeast functional screen predicts new candidate ALS disease genes. *Proceedings of the National Academy of Sciences of the United States of America* 108:20881-90
 35. Craufurd D, Thompson JC, Snowden JS. 2001. Behavioral changes in Huntington Disease. *Neuropsychiatry, neuropsychology, and behavioral neurology* 14:219-26
 36. Cummings CJ, Mancini MA, Antalffy B, DeFranco DB, Orr HT, Zoghbi HY. 1998. Chaperone suppression of aggregation and altered subcellular proteasome localization imply protein misfolding in SCA1. *Nature genetics* 19:148-54
 37. Da Cruz S, Cleveland DW. 2011. Understanding the role of TDP-43 and FUS/TLS in ALS and beyond. *Curr Opin Neurobiol* 21:904-19
 38. Daigle JG, Lanson NA, Jr., Smith RB, Casci I, Maltare A, et al. 2013. RNA-binding ability of FUS regulates neurodegeneration, cytoplasmic mislocalization and incorporation into stress granules associated with FUS carrying ALS-linked mutations. *Human molecular genetics* 22:1193-205
 39. Davies SW, Turmaine M, Cozens BA, DiFiglia M, Sharp AH, et al. 1997. Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell* 90:537-48
 40. Davranche A, Aviolat H, Zeder-Lutz G, Busso D, Altschuh D, et al. 2011. Huntingtin affinity for partners is not changed by polyglutamine length: aggregation itself triggers aberrant interactions. *Human molecular genetics* 20:2795-806
 41. Dawson S, Kristjanson LJ, Toye CM, Flett P. 2004. Living with Huntington's disease: need for supportive care. *Nursing & health sciences* 6:123-30
 42. de Pril R, Fischer DF, Roos RA, van Leeuwen FW. 2007. Ubiquitin-conjugating enzyme E2-25K increases aggregate formation and cell death in polyglutamine diseases. *Molecular and cellular neurosciences* 34:10-9
 43. Deng HX, Chen W, Hong ST, Boycott KM, Gorrie GH, et al. 2011. Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* 477:211-5
 44. Derkatch IL, Bradley ME, Hong JY, Liebman SW. 2001. Prions affect the appearance of other prions: the story of [PIN(+)]. *Cell* 106:171-82
 45. DiFiglia M, Sapp E, Chase KO, Davies SW, Bates GP, et al. 1997. Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* 277:1990-3
 46. Divino V, Dekoven M, Warner JH, Giuliano J, Anderson KE, et al. 2013. The direct medical costs of Huntington's disease by stage. A retrospective commercial and Medicaid claims data analysis. *Journal of medical economics* 16:1043-50
 47. Doi H, Mitsui K, Kurosawa M, Machida Y, Kuroiwa Y, Nukina N. 2004. Identification of ubiquitin-interacting proteins in purified polyglutamine aggregates. *FEBS letters* 571:171-6
 48. Doi H, Okamura K, Bauer PO, Furukawa Y, Shimizu H, et al. 2008. RNA-binding protein TLS is a major nuclear aggregate-interacting protein in huntingtin

-
- exon 1 with expanded polyglutamine-expressing cells. *The Journal of biological chemistry* 283:6489-500
49. Dong J, Castro CE, Boyce MC, Lang MJ, Lindquist S. 2010. Optical trapping with high forces reveals unexpected behaviors of prion fibrils. *Nature structural & molecular biology* 17:1422-30
50. Dosztanyi Z, Csizmok V, Tompa P, Simon I. 2005. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21:3433-4
51. Dosztanyi Z, Csizmok V, Tompa P, Simon I. 2005. The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *Journal of molecular biology* 347:827-39
52. Duennwald ML, Jagadish S, Giorgini F, Muchowski PJ, Lindquist S. 2006. A network of protein interactions determines polyglutamine toxicity. *Proceedings of the National Academy of Sciences of the United States of America* 103:11051-6
53. Eanes ED, Glenner GG. 1968. X-ray diffraction studies on amyloid filaments. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* 16:673-7
54. Edskes HK, Wickner RB. 2000. A protein required for prion generation: [URE3] induction requires the Ras-regulated Mks1 protein. *Proceedings of the National Academy of Sciences of the United States of America* 97:6625-9
55. Edskes HK, Wickner RB. 2002. Conservation of a portion of the *S. cerevisiae* Ure2p prion domain that interacts with the full-length protein. *Proceedings of the National Academy of Sciences of the United States of America* 99 Suppl 4:16384-91
56. Fleischer TC, Weaver CM, McAfee KJ, Jennings JL, Link AJ. 2006. Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes. *Genes & development* 20:1294-307
57. Frieden C. 2007. Protein aggregation processes: In search of the mechanism. *Protein science : a publication of the Protein Society* 16:2334-44
58. Fry D, Huang KS, Di Lello P, Mohr P, Muller K, et al. 2013. Design of libraries targeting protein-protein interfaces. *ChemMedChem* 8:726-32
59. Fuentealba RA, Udan M, Bell S, Wegorzewska I, Shao J, et al. 2010. Interaction with polyglutamine aggregates reveals a Q/N-rich domain in TDP-43. *The Journal of biological chemistry* 285:26304-14
60. Furukawa Y, Kaneko K, Matsumoto G, Kurosawa M, Nukina N. 2009. Cross-seeding fibrillation of Q/N-rich proteins offers new pathomechanism of polyglutamine diseases. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29:5153-62
61. Fushimi K, Long C, Jayaram N, Chen X, Li L, Wu JY. 2011. Expression of human FUS/TLS in yeast leads to protein aggregation and cytotoxicity, recapitulating key features of FUS proteinopathy. *Protein Cell* 2:141-9
62. Fitcher B, Latter GI, Monardo P, McLaughlin CS, Garrels JI. 1999. A sampling of the yeast proteome. *Molecular and cellular biology* 19:7357-68
63. Gelperin D, Weigle J, Nelson K, Roseboom P, Irie K, et al. 1995. 14-3-3 proteins: potential roles in vesicular transport and Ras signaling in *Saccharomyces*

-
- cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America* 92:11539-43
64. Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, et al. 2003. Global analysis of protein expression in yeast. *Nature* 425:737-41
 65. Gidalevitz T, Ben-Zvi A, Ho KH, Brignull HR, Morimoto RI. 2006. Progressive disruption of cellular protein folding in models of polyglutamine diseases. *Science* 311:1471-4
 66. Giorgini F, Guidetti P, Nguyen Q, Bennett SC, Muchowski PJ. 2005. A genomic screen in yeast implicates kynurenine 3-monooxygenase as a therapeutic target for Huntington disease. *Nature genetics* 37:526-31
 67. Goehler H, Lalowski M, Stelzl U, Waelter S, Stroedicke M, et al. 2004. A protein interaction network links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease. *Molecular cell* 15:853-65
 68. Golanska E, Hulas-Bigoszewska K, Sikorska B, Liberski PP. 2005. [Analyses of 14-3-3 protein in the cerebrospinal fluid in Creutzfeldt-Jakob disease. Preliminary report]. *Neurologia i neurochirurgia polska* 39:358-65
 69. Gourfinkel-An I, Cancel G, Duyckaerts C, Faucheux B, Hauw JJ, et al. 1998. Neuronal distribution of intranuclear inclusions in Huntington's disease with adult onset. *Neuroreport* 9:1823-6
 70. Gunawardena S, Her LS, Brusch RG, Laymon RA, Niesman IR, et al. 2003. Disruption of axonal transport by loss of huntingtin or expression of pathogenic polyQ proteins in Drosophila. *Neuron* 40:25-40
 71. Gutekunst CA, Li SH, Yi H, Mulroy JS, Kuemmerle S, et al. 1999. Nuclear and neuropil aggregates in Huntington's disease: relationship to neuropathology. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 19:2522-34
 72. Habchi J, Tompa P, Longhi S, Uversky VN. 2014. Introducing protein intrinsic disorder. *Chemical reviews* 114:6561-88
 73. Hammoudeh DI, Follis AV, Prochownik EV, Metallo SJ. 2009. Multiple independent binding sites for small-molecule inhibitors on the oncoprotein c-Myc. *Journal of the American Chemical Society* 131:7390-401
 74. Hands SL, Wytenbach A. 2010. Neurotoxic protein oligomerisation associated with polyglutamine diseases. *Acta Neuropathol* 120:419-37
 75. Harjes P, Wanker EE. 2003. The hunt for huntingtin function: interaction partners tell many different stories. *Trends in biochemical sciences* 28:425-33
 76. Harper JD, Lansbury PT, Jr. 1997. Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annual review of biochemistry* 66:385-407
 77. Harvey SR, Porrini M, Stachl C, MacMillan D, Zinzalla G, Barran PE. 2012. Small-molecule inhibition of c-MYC:MAX leucine zipper formation is revealed by ion mobility mass spectrometry. *Journal of the American Chemical Society* 134:19384-92
 78. Henshall TL, Tucker B, Lumsden AL, Nornes S, Lardelli MT, Richards RI. 2009. Selective neuronal requirement for huntingtin in the developing zebrafish. *Human molecular genetics* 18:4830-42

-
79. Hoffner G, Djian P. 2014. Monomeric, oligomeric and polymeric proteins in huntington disease and other diseases of polyglutamine expansion. *Brain Sci* 4:91-122
 80. Holbert S, Denghien I, Kiechle T, Rosenblatt A, Wellington C, et al. 2001. The Gln-Ala repeat transcriptional activator CA150 interacts with huntingtin: neuropathologic and genetic evidence for a role in Huntington's disease pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 98:1811-6
 81. Hu G, Wu Z, Wang K, Uversky VN, Kurgan L. 2015. Untapped potential of disordered proteins in current druggable human proteome. *Current drug targets*
 82. Hu NW, Smith IM, Walsh DM, Rowan MJ. 2008. Soluble amyloid-beta peptides potently disrupt hippocampal synaptic plasticity in the absence of cerebrovascular dysfunction in vivo. *Brain : a journal of neurology* 131:2414-24
 83. Invernizzi G, Papaleo E, Sabate R, Ventura S. 2012. Protein aggregation: mechanisms and functional consequences. *The international journal of biochemistry & cell biology* 44:1541-54
 84. Ishiura H, Sako W, Yoshida M, Kawarai T, Tanabe O, et al. 2012. The TRK-fused gene is mutated in hereditary motor and sensory neuropathy with proximal dominant involvement. *American journal of human genetics* 91:320-9
 85. Ivanyi-Nagy R, Davidovic L, Khandjian EW, Darlix JL. 2005. Disordered RNA chaperone proteins: from functions to disease. *Cellular and molecular life sciences : CMLS* 62:1409-17
 86. Jacobsen JC, Gregory GC, Woda JM, Thompson MN, Coser KR, et al. 2011. HD CAG-correlated gene expression changes support a simple dominant gain of function. *Human molecular genetics* 20:2846-60
 87. Jana NR, Tanaka M, Wang G, Nukina N. 2000. Polyglutamine length-dependent interaction of Hsp40 and Hsp70 family chaperones with truncated N-terminal huntingtin: their role in suppression of aggregation and cellular toxicity. *Human molecular genetics* 9:2009-18
 88. Jarrett JT, Lansbury PT, Jr. 1993. Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* 73:1055-8
 89. Jimenez JL, Guijarro JJ, Orlova E, Zurdo J, Dobson CM, et al. 1999. Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing. *The EMBO journal* 18:815-21
 90. Jin F, Yu C, Lai L, Liu Z. 2013. Ligand clouds around protein clouds: a scenario of ligand binding with intrinsically disordered proteins. *PLoS computational biology* 9:e1003249
 91. Jin M, Shepardson N, Yang T, Chen G, Walsh D, Selkoe DJ. 2011. Soluble amyloid beta-protein dimers isolated from Alzheimer cortex directly induce Tau hyperphosphorylation and neuritic degeneration. *Proceedings of the National Academy of Sciences of the United States of America* 108:5819-24
 92. Johnson JO, Pioro EP, Boehringer A, Chia R, Feit H, et al. 2014. Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nature neuroscience* 17:664-6

-
93. Johnston JA, Ward CL, Kopito RR. 1998. Aggresomes: a cellular response to misfolded proteins. *The Journal of cell biology* 143:1883-98
 94. Joshi P, Vendruscolo M. 2015. Druggability of Intrinsically Disordered Proteins. *Advances in experimental medicine and biology* 870:383-400
 95. Ju S, Tardiff DF, Han H, Divya K, Zhong Q, et al. 2011. A yeast model of FUS/TLS-dependent cytotoxicity. *PLoS biology* 9:e1001052
 96. Kabashi E, Valdmanis PN, Dion P, Spiegelman D, McConkey BJ, et al. 2008. TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nature genetics* 40:572-4
 97. Kaltenbach LS, Romero E, Becklin RR, Chettier R, Bell R, et al. 2007. Huntingtin interacting proteins are genetic modifiers of neurodegeneration. *PLoS genetics* 3:e82
 98. Kar K, Hoop CL, Drombosky KW, Baker MA, Kodali R, et al. 2013. beta-hairpin-mediated nucleation of polyglutamine amyloid formation. *Journal of molecular biology* 425:1183-97
 99. Karlin S, Brocchieri L, Bergman A, Mrazek J, Gentles AJ. 2002. Amino acid runs in eukaryotic proteomes and disease associations. *Proceedings of the National Academy of Sciences of the United States of America* 99:333-8
 100. Kawamoto Y, Akiguchi I, Fujimura H, Shirakashi Y, Honjo Y, Sakoda S. 2005. 14-3-3 proteins in Lewy body-like hyaline inclusions in a patient with familial amyotrophic lateral sclerosis with a two-base pair deletion in the Cu/Zn superoxide dismutase (SOD1) gene. *Acta Neuropathol* 110:203-4
 101. Kayatekin C, Matlack KE, Hesse WR, Guan Y, Chakrabortee S, et al. 2014. Prion-like proteins sequester and suppress the toxicity of huntingtin exon 1. *Proceedings of the National Academy of Sciences of the United States of America* 111:12085-90
 102. Kikhney AG, Svergun DI. 2015. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS letters* 589:2570-7
 103. Kiktev DA, Patterson JC, Muller S, Bariar B, Pan T, Chernoff YO. 2012. Regulation of chaperone effects on a yeast prion by cochaperone Sgt2. *Molecular and cellular biology* 32:4960-70
 104. Kim WY, Fayazi Z, Bao X, Higgins D, Kazemi-Esfarjani P. 2005. Evidence for sequestration of polyglutamine inclusions by Drosophila myeloid leukemia factor. *Molecular and cellular neurosciences* 29:536-44
 105. King OD, Gitler AD, Shorter J. 2012. The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain research* 1462:61-80
 106. Knott M, Best RB. 2012. A preformed binding interface in the unbound ensemble of an intrinsically disordered protein: evidence from molecular simulations. *PLoS computational biology* 8:e1002605
 107. Knowles TP, Vendruscolo M, Dobson CM. 2014. The amyloid state and its association with protein misfolding diseases. *Nature reviews. Molecular cell biology* 15:384-96
 108. Koo EH, Lansbury PT, Jr., Kelly JW. 1999. Amyloid diseases: abnormal protein aggregation in neurodegeneration. *Proceedings of the National Academy of Sciences of the United States of America* 96:9989-90

-
109. Kryndushkin D, Pripuzova N, Burnett B, Shewmaker F. 2013. Non-targeted identification of prions and amyloid-forming proteins from yeast and mammalian cells. *The Journal of biological chemistry* 288:27100-11
 110. Kryndushkin D, Wear MP, Shewmaker F. 2013. Amyloid cannot resist identification. *Prion* 7
 111. Kryndushkin D, Wickner RB, Shewmaker F. 2011. FUS/TLS forms cytoplasmic aggregates, inhibits cell growth and interacts with TDP-43 in a yeast model of amyotrophic lateral sclerosis. *Protein Cell*
 112. Kryndushkin DS, Alexandrov IM, Ter-Avanesyan MD, Kushnirov VV. 2003. Yeast [PSI⁺] prion aggregates are formed by small Sup35 polymers fragmented by Hsp104. *The Journal of biological chemistry* 278:49636-43
 113. Kryndushkin DS, Engel A, Edskes H, Wickner RB. 2011. Molecular chaperone Hsp104 can promote yeast prion generation. *Genetics* 188:339-48
 114. Kuemmerle S, Gutekunst CA, Klein AM, Li XJ, Li SH, et al. 1999. Huntington aggregates may not predict neuronal death in Huntington's disease. *Annals of neurology* 46:842-9
 115. Kuusisto E, Salminen A, Alafuzoff I. 2001. Ubiquitin-binding protein p62 is present in neuronal and glial inclusions in human tauopathies and synucleinopathies. *Neuroreport* 12:2085-90
 116. Kwiatkowski TJ, Jr., Bosco DA, Leclerc AL, Tamrazian E, Vanderburg CR, et al. 2009. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* 323:1205-8
 117. Laulederkind SJ, Hayman GT, Wang SJ, Smith JR, Lowry TF, et al. 2013. The Rat Genome Database 2013--data, tools and users. *Brief Bioinform* 14:520-6
 118. Li SH, Li XJ. 2004. Huntingtin-protein interactions and the pathogenesis of Huntington's disease. *Trends Genet* 20:146-54
 119. Li YR, King OD, Shorter J, Gitler AD. 2013. Stress granules as crucibles of ALS pathogenesis. *The Journal of cell biology* 201:361-72
 120. Liao L, Cheng D, Wang J, Duong DM, Losik TG, et al. 2004. Proteomic characterization of postmortem amyloid plaques isolated by laser capture microdissection. *The Journal of biological chemistry* 279:37061-8
 121. Lin MT, Beal MF. 2006. Alzheimer's APP mangles mitochondria. *Nature medicine* 12:1241-3
 122. Lu P, Vogel C, Wang R, Yao X, Marcotte EM. 2007. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nature biotechnology* 25:117-24
 123. Ma J, Lindquist S. 2002. Conversion of PrP to a self-perpetuating PrP^{Sc}-like conformation in the cytosol. *Science* 298:1785-8
 124. Ma J, Wollmann R, Lindquist S. 2002. Neurotoxicity and neurodegeneration when PrP accumulates in the cytosol. *Science* 298:1781-5
 125. Mackenzie IR, Rademakers R, Neumann M. 2010. TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. *The Lancet. Neurology* 9:995-1007

-
126. MacPhee CE, Dobson CM. 2000. Chemical dissection and reassembly of amyloid fibrils formed by a peptide fragment of transthyretin. *Journal of molecular biology* 297:1203-15
 127. Makin OS, Atkins E, Sikorski P, Johansson J, Serpell LC. 2005. Molecular basis for amyloid fibril formation and stability. *Proceedings of the National Academy of Sciences of the United States of America* 102:315-20
 128. Malchiodi-Albedi F, Paradisi S, Matteucci A, Frank C, Diociaiuti M. 2011. Amyloid oligomer neurotoxicity, calcium dysregulation, and lipid rafts. *International journal of Alzheimer's disease* 2011:906964
 129. Marasco D, Scognamiglio PL. 2015. Identification of inhibitors of biological interactions involving intrinsically disordered proteins. *International journal of molecular sciences* 16:7394-412
 130. Margittai M, Langen R. 2008. Fibrils with parallel in-register structure constitute a major class of amyloid fibrils: molecular insights from electron paramagnetic resonance spectroscopy. *Quarterly reviews of biophysics* 41:265-97
 131. Martindale D, Hackam A, Wieczorek A, Ellerby L, Wellington C, et al. 1998. Length of huntingtin and its polyglutamine tract influences localization and frequency of intracellular aggregates. *Nature genetics* 18:150-4
 132. Mason RP, Casu M, Butler N, Breda C, Campesan S, et al. 2013. Glutathione peroxidase activity is neuroprotective in models of Huntington's disease. *Nature genetics* 45:1249-54
 133. Mayer AM, Hall ML, Lynch SM, Gunasekera SP, Sennett SH, Pomponi SA. 2005. Differential modulation of microglia superoxide anion and thromboxane B2 generation by the marine manzamines. *BMC pharmacology* 5:6
 134. Meriin AB, Zhang X, He X, Newnam GP, Chernoff YO, Sherman MY. 2002. Huntington toxicity in yeast model depends on polyglutamine aggregation mediated by a prion-like protein Rnq1. *The Journal of cell biology* 157:997-1004
 135. Meriin AB, Zhang X, Miliaras NB, Kazantsev A, Chernoff YO, et al. 2003. Aggregation of expanded polyglutamine domain in yeast leads to defects in endocytosis. *Molecular and cellular biology* 23:7554-65
 136. Michel J, Cuchillo R. 2012. The impact of small molecule binding on the energy landscape of the intrinsically disordered protein C-myc. *PloS one* 7:e41070
 137. Michelitsch MD, Weissman JS. 2000. A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. *Proceedings of the National Academy of Sciences of the United States of America* 97:11910-5
 138. Mitsui K, Doi H, Nukina N. 2006. Proteomics of polyglutamine aggregates. *Methods Enzymol* 412:63-76
 139. Mizuno Y, Amari M, Takatama M, Aizawa H, Mihara B, Okamoto K. 2006. Immunoreactivities of p62, an ubiquitin-binding protein, in the spinal anterior horn cells of patients with amyotrophic lateral sclerosis. *Journal of the neurological sciences* 249:13-8
 140. Moree B, Yin G, Lazaro DF, Munari F, Strohaker T, et al. 2015. Small Molecules Detected by Second-Harmonic Generation Modulate the Conformation of Monomeric alpha-Synuclein and Reduce Its Aggregation in Cells. *The Journal of biological chemistry* 290:27582-93

-
141. Mori K, Lammich S, Mackenzie IR, Forne I, Zilow S, et al. 2013. hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations. *Acta Neuropathol* 125:413-23
 142. Nagai Y, Inui T, Popiel HA, Fujikake N, Hasegawa K, et al. 2007. A toxic monomeric conformer of the polyglutamine protein. *Nature structural & molecular biology* 14:332-40
 143. Nelson DL, Orr HT, Warren ST. 2013. The unstable repeats--three evolving faces of neurological disease. *Neuron* 77:825-43
 144. Newman JR, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, et al. 2006. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* 441:840-6
 145. Nielsen L, Frokjaer S, Carpenter JF, Brange J. 2001. Studies of the structure of insulin fibrils by Fourier transform infrared (FTIR) spectroscopy and electron microscopy. *Journal of pharmaceutical sciences* 90:29-37
 146. Niikura T, Kita Y, Abe Y. 2014. SUMO3 modification accelerates the aggregation of ALS-linked SOD1 mutants. *PloS one* 9:e101080
 147. Novoselova TV, Margulis BA, Novoselov SS, Sapozhnikov AM, van der Spuy J, et al. 2005. Treatment with extracellular HSP70/HSC70 protein can reduce polyglutamine toxicity and aggregation. *Journal of neurochemistry* 94:597-606
 148. Nucifora FC, Jr., Ellerby LM, Wellington CL, Wood JD, Herring WJ, et al. 2003. Nuclear localization of a non-caspase truncation product of atrophin-1, with an expanded polyglutamine repeat, increases cellular toxicity. *The Journal of biological chemistry* 278:13047-55
 149. O'Rourke JG, Gareau JR, Ochaba J, Song W, Rasko T, et al. 2013. SUMO-2 and PIAS1 modulate insoluble mutant huntingtin protein accumulation. *Cell reports* 4:362-75
 150. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, et al. 2005. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Molecular & cellular proteomics : MCP* 4:1487-502
 151. Olshina MA, Angley LM, Ramdzan YM, Tang J, Bailey MF, et al. 2010. Tracking mutant huntingtin aggregation kinetics in cells reveals three major populations that include an invariant oligomer pool. *The Journal of biological chemistry* 285:21807-16
 152. Orr HT, Zoghbi HY. 2007. Trinucleotide repeat disorders. *Annu Rev Neurosci* 30:575-621
 153. Outeiro TF, Lindquist S. 2003. Yeast cells provide insight into alpha-synuclein biology and pathobiology. *Science* 302:1772-5
 154. Panov AV, Gutekunst CA, Leavitt BR, Hayden MR, Burke JR, et al. 2002. Early mitochondrial calcium defects in Huntington's disease are a direct effect of polyglutamines. *Nature neuroscience* 5:731-6
 155. Park SH, Kukushkin Y, Gupta R, Chen T, Konagai A, et al. 2013. PolyQ proteins interfere with nuclear degradation of cytosolic proteins by sequestering the Sis1p chaperone. *Cell* 154:134-45

-
156. Paulson HL, Perez MK, Trotter Y, Trojanowski JQ, Subramony SH, et al. 1997. Intranuclear inclusions of expanded polyglutamine protein in spinocerebellar ataxia type 3. *Neuron* 19:333-44
 157. Poirier MA, Li H, Macosko J, Cai S, Amzel M, Ross CA. 2002. Huntingtin spheroids and protofibrils as precursors in polyglutamine fibrilization. *The Journal of biological chemistry* 277:41032-7
 158. Ramaswami M, Taylor JP, Parker R. 2013. Altered ribostasis: RNA-protein granules in degenerative disorders. *Cell* 154:727-36
 159. Ramirez A, van der Flier WM, Herold C, Ramonet D, Heilmann S, et al. 2014. SUCLG2 identified as both a determinant of CSF Aβ₁₋₄₂ levels and an attenuator of cognitive decline in Alzheimer's disease. *Human molecular genetics* 23:6644-58
 160. Ratovitski T, Chighladze E, Arbez N, Boronina T, Herbrich S, et al. 2012. Huntingtin protein interactions altered by polyglutamine expansion as determined by quantitative proteomic analysis. *Cell cycle* 11:2006-21
 161. Raychaudhuri S, Dey S, Bhattacharyya NP, Mukhopadhyay D. 2009. The role of intrinsically unstructured proteins in neurodegenerative diseases. *PloS one* 4:e5566
 162. Reed BJ, Locke MN, Gardner RG. 2015. A Conserved Deubiquitinating Enzyme Uses Intrinsically Disordered Regions to Scaffold Multiple Protein Interaction Sites. *The Journal of biological chemistry* 290:20601-12
 163. Reilly MM. 1998. Genetically determined neuropathies. *Journal of neurology* 245:6-13
 164. Reinders J, Zahedi RP, Pfanner N, Meisinger C, Sickmann A. 2006. Toward the complete yeast mitochondrial proteome: multidimensional separation techniques for mitochondrial proteomics. *Journal of proteome research* 5:1543-54
 165. Renner M, Melki R. 2014. Protein aggregation and prionopathies. *Pathologie-biologie* 62:162-8
 166. Roos RA. 2010. Huntington's disease: a clinical review. *Orphanet journal of rare diseases* 5:40
 167. Ross CA, Poirier MA. 2004. Protein aggregation and neurodegenerative disease. *Nature medicine* 10 Suppl:S10-7
 168. Rossetti G, Dibenedetto D, Calandrini V, Giorgetti A, Carloni P. 2015. Structural predictions of neurobiologically relevant G-protein coupled receptors and intrinsically disordered proteins. *Archives of biochemistry and biophysics* 582:91-100
 169. Rub U, de Vos RA, Brunt ER, Sebesteny T, Schols L, et al. 2006. Spinocerebellar ataxia type 3 (SCA3): thalamic neurodegeneration occurs independently from thalamic ataxin-3 immunopositive neuronal intranuclear inclusions. *Brain pathology* 16:218-27
 170. Rubinsztein DC, Carmichael J. 2003. Huntington's disease: molecular basis of neurodegeneration. *Expert Rev Mol Med* 5:1-21
 171. Rutherford NJ, Lewis J, Clippinger AK, Thomas MA, Adamson J, et al. 2013. Unbiased screen reveals ubiquitin-1 and -2 highly associated with huntingtin inclusions. *Brain research* 1524:62-73

-
172. Sabate R, Estelrich J. 2005. Evidence of the existence of micelles in the fibrillogenesis of beta-amyloid peptide. *The journal of physical chemistry. B* 109:11027-32
 173. Sabate R, Gallardo M, Estelrich J. 2003. An autocatalytic reaction as a model for the kinetics of the aggregation of beta-amyloid. *Biopolymers* 71:190-5
 174. Sakahira H, Breuer P, Hayer-Hartl MK, Hartl FU. 2002. Molecular chaperones as modulators of polyglutamine protein aggregation and toxicity. *Proceedings of the National Academy of Sciences of the United States of America* 99 Suppl 4:16412-8
 175. Saudou F, Finkbeiner S, Devys D, Greenberg ME. 1998. Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions. *Cell* 95:55-66
 176. Scherzinger E, Lurz R, Turmaine M, Mangiarini L, Hollenbach B, et al. 1997. Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. *Cell* 90:549-58
 177. Scherzinger E, Sittler A, Schweiger K, Heiser V, Lurz R, et al. 1999. Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. *Proceedings of the National Academy of Sciences of the United States of America* 96:4604-9
 178. Schiffer NW, Broadley SA, Hirschberger T, Tavan P, Kretzschmar HA, et al. 2007. Identification of anti-prion compounds as efficient inhibitors of polyglutamine protein aggregation in a zebrafish model. *The Journal of biological chemistry* 282:9195-203
 179. Schmitz M, Ebert E, Stoeck K, Karch A, Collins S, et al. 2015. Validation of 14-3-3 Protein as a Marker in Sporadic Creutzfeldt-Jakob Disease Diagnostic. *Molecular neurobiology*
 180. Schwartz R, Ting CS, King J. 2001. Whole proteome pI values correlate with subcellular localizations of proteins for organisms within the three domains of life. *Genome Res* 11:703-9
 181. Schwimmer C, Masison DC. 2002. Antagonistic interactions between yeast [PSI(+)] and [URE3] prions and curing of [URE3] by Hsp70 protein chaperone Ssa1p but not by Ssa2p. *Molecular and cellular biology* 22:3590-8
 182. Serio TR, Cashikar AG, Kowal AS, Sawicki GJ, Moslehi JJ, et al. 2000. Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* 289:1317-21
 183. Shelkownikova TA, Robinson HK, Troakes C, Ninkina N, Buchman VL. 2014. Compromised paraspeckle formation as a pathogenic factor in FUSopathies. *Human molecular genetics* 23:2298-312
 184. Shewmaker F, McGlinchey RP, Wickner RB. 2011. Structural insights into functional and pathological amyloid. *The Journal of biological chemistry* 286:16533-40
 185. Shi B, Conner SD, Liu J. 2014. Dysfunction of endocytic kinase AAK1 in ALS. *International journal of molecular sciences* 15:22918-32
 186. Sieradzan KA, Mechan AO, Jones L, Wanker EE, Ninkina N, Mann DM. 1999. Huntington's disease intranuclear inclusions contain truncated, ubiquitinated huntingtin protein. *Experimental neurology* 156:92-9

-
187. Sipe JD, Cohen AS. 2000. Review: history of the amyloid fibril. *Journal of structural biology* 130:88-98
 188. Slow EJ, Graham RK, Osmand AP, Devon RS, Lu G, et al. 2005. Absence of behavioral abnormalities and neurodegeneration in vivo despite widespread neuronal huntingtin inclusions. *Proceedings of the National Academy of Sciences of the United States of America* 102:11402-7
 189. Sormanni P, Aprile FA, Vendruscolo M. 2015. Rational design of antibodies targeting specific epitopes within intrinsically disordered proteins. *Proceedings of the National Academy of Sciences of the United States of America* 112:9902-7
 190. Spiess M, de Craene JO, Michelot A, Rinaldi B, Huber A, et al. 2013. Lsb1 is a negative regulator of las17 dependent actin polymerization involved in endocytosis. *PloS one* 8:e61147
 191. Squitieri F, Andrew SE, Goldberg YP, Kremer B, Spence N, et al. 1994. DNA haplotype analysis of Huntington disease reveals clues to the origins and mechanisms of CAG expansion and reasons for geographic variations of prevalence. *Human molecular genetics* 3:2103-14
 192. Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, et al. 2008. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science* 319:1668-72
 193. Stefani M. 2004. Protein misfolding and aggregation: new examples in medicine and biology of the dark side of the protein world. *Biochimica et biophysica acta* 1739:5-25
 194. Stefani M, Dobson CM. 2003. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J Mol Med* 81:678-99
 195. Sugaya K. 2012. Modeling the polyglutamine aggregation pathway in Huntington's disease: from basic studies to clinical applications. *Sub-cellular biochemistry* 65:353-88
 196. Sun Z, Diaz Z, Fang X, Hart MP, Chesi A, et al. 2011. Molecular determinants and genetic modifiers of aggregation and toxicity for the ALS disease protein FUS/TLS. *PLoS biology* 9:e1000614
 197. Sunde M, Blake C. 1997. The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Advances in protein chemistry* 50:123-59
 198. Sunde M, Blake CC. 1998. From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. *Quarterly reviews of biophysics* 31:1-39
 199. Sunde M, Serpell LC, Bartlam M, Fraser PE, Pepys MB, Blake CC. 1997. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of molecular biology* 273:729-39
 200. Szebenyi G, Morfini GA, Babcock A, Gould M, Selkoe K, et al. 2003. Neuropathogenic forms of huntingtin and androgen receptor inhibit fast axonal transport. *Neuron* 40:41-52
 201. Takahashi T, Kikuchi S, Katada S, Nagai Y, Nishizawa M, Onodera O. 2008. Soluble polyglutamine oligomers formed prior to inclusion body formation are cytotoxic. *Human molecular genetics* 17:345-56

-
202. Tang TS, Tu H, Chan EY, Maximov A, Wang Z, et al. 2003. Huntingtin and huntingtin-associated protein 1 influence neuronal calcium signaling mediated by inositol-(1,4,5) triphosphate receptor type 1. *Neuron* 39:227-39
203. Taylor JP, Hardy J, Fischbeck KH. 2002. Toxic proteins in neurodegenerative disease. *Science* 296:1991-5
204. Tompa P, Csermely P. 2004. The role of structural disorder in the function of RNA and protein chaperones. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 18:1169-75
205. Tompa P, Fuxreiter M, Oldfield CJ, Simon I, Dunker AK, Uversky VN. 2009. Close encounters of the third kind: disordered domains and the interactions of proteins. *BioEssays : news and reviews in molecular, cellular and developmental biology* 31:328-35
206. Torok M, Milton S, Kaye R, Wu P, McIntire T, et al. 2002. Structural and dynamic features of Alzheimer's A β peptide in amyloid fibrils studied by site-directed spin labeling. *The Journal of biological chemistry* 277:40810-5
207. Toth G, Gardai SJ, Zago W, Bertoni CW, Cremades N, et al. 2014. Targeting the intrinsically disordered structural ensemble of alpha-synuclein by small molecules as a potential therapeutic strategy for Parkinson's disease. *PloS one* 9:e87133
208. Tourette C, Li B, Bell R, O'Hare S, Kaltenbach LS, et al. 2014. A large scale Huntingtin protein interaction network implicates Rho GTPase signaling pathways in Huntington disease. *The Journal of biological chemistry* 289:6709-26
209. Tycko R. 2011. Solid-state NMR studies of amyloid fibril structure. *Annu Rev Phys Chem* 62:279-99
210. Tydlacka S, Wang CE, Wang X, Li S, Li XJ. 2008. Differential activities of the ubiquitin-proteasome system in neurons versus glia may account for the preferential accumulation of misfolded proteins in neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 28:13285-95
211. Uversky VN. 2010. Targeting intrinsically disordered proteins in neurodegenerative and protein dysfunction diseases: another illustration of the D(2) concept. *Expert Rev Proteomics* 7:543-64
212. Uversky VN. 2014. The triple power of D(3): protein intrinsic disorder in degenerative diseases. *Frontiers in bioscience* 19:181-258
213. Uversky VN. 2015. Intrinsically disordered proteins and their (disordered) proteomes in neurodegenerative disorders. *Front Aging Neurosci* 7:18
214. Uversky VN, Gillespie JR, Fink AL. 2000. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* 41:415-27
215. Uversky VN, Oldfield CJ, Dunker AK. 2008. Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annual review of biophysics* 37:215-46
216. van Heusden GP, Griffiths DJ, Ford JC, Chin AWTF, Schrader PA, et al. 1995. The 14-3-3 proteins encoded by the BMH1 and BMH2 genes are essential in the yeast *Saccharomyces cerevisiae* and can be replaced by a plant homologue. *European journal of biochemistry / FEBS* 229:45-53
-

-
217. Vance C, Rogelj B, Hortobagyi T, De Vos KJ, Nishimura AL, et al. 2009. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 323:1208-11
 218. Varadi M, Kosol S, Lebrun P, Valentini E, Blackledge M, et al. 2014. pE-DB: a database of structural ensembles of intrinsically disordered and of unfolded proteins. *Nucleic acids research* 42:D326-35
 219. Varadi M, Zsolyomi F, Guharoy M, Tompa P. 2015. Functional Advantages of Conserved Intrinsic Disorder in RNA-Binding Proteins. *PloS one* 10:e0139731
 220. Vendruscolo M, Knowles TP, Dobson CM. 2014. Protein solubility and protein homeostasis: a generic view of protein misfolding disorders. *Cold Spring Harb Perspect Biol* 3
 221. Vucetic S, Xie H, Iakoucheva LM, Oldfield CJ, Dunker AK, et al. 2007. Functional anthology of intrinsic disorder. 2. Cellular components, domains, technical terms, developmental processes, and coding sequence diversities correlated with long disordered regions. *Journal of proteome research* 6:1899-916
 222. Wacker JL, Zareie MH, Fong H, Sarikaya M, Muchowski PJ. 2004. Hsp70 and Hsp40 attenuate formation of spherical and annular polyglutamine oligomers by partitioning monomer. *Nature structural & molecular biology* 11:1215-22
 223. Waelter S, Boeddrich A, Lurz R, Scherzinger E, Lueder G, et al. 2001. Accumulation of mutant huntingtin fragments in aggresome-like inclusion bodies as a result of insufficient protein degradation. *Molecular biology of the cell* 12:1393-407
 224. Wang G, Sawai N, Kotliarova S, Kanazawa I, Nukina N. 2000. Ataxin-3, the MJD1 gene product, interacts with the two human homologs of yeast DNA repair protein RAD23, HHR23A and HHR23B. *Human molecular genetics* 9:1795-803
 225. Wang J, Cao Z, Zhao L, Li S. 2011. Novel strategies for drug discovery based on Intrinsically Disordered Proteins (IDPs). *International journal of molecular sciences* 12:3205-19
 226. Wang Q, Woltjer RL, Cimino PJ, Pan C, Montine KS, et al. 2005. Proteomic analysis of neurofibrillary tangles in Alzheimer disease identifies GAPDH as a detergent-insoluble paired helical filament tau binding protein. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 19:869-71
 227. Wang X, Chapman MR. 2008. Curli provide the template for understanding controlled amyloid propagation. *Prion* 2:57-60
 228. Wang Y, Meriin AB, Costello CE, Sherman MY. 2007. Characterization of proteins associated with polyglutamine aggregates: a novel approach towards isolation of aggregates from protein conformation disorders. *Prion* 1:128-35
 229. Wang Y, Meriin AB, Zaarur N, Romanova NV, Chernoff YO, et al. 2009. Abnormal proteins can form aggresome in yeast: aggresome-targeting signals and components of the machinery. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 23:451-63
 230. Ward JJ, McGuffin LJ, Bryson K, Buxton BF, Jones DT. 2004. The DISOPRED server for the prediction of protein disorder. *Bioinformatics* 20:2138-9

-
231. Wear MKD, O'Meally R, Cole RN, Sonnenberg JL, Shewmaker F. 2015. Proteins with intrinsically disordered domains are preferentially recruited to polyglutamine aggregates. *PloS one*
232. Wear MP, Kryndushkin D, O'Meally R, Sonnenberg JL, Cole RN, Shewmaker FP. 2015. Proteins with Intrinsically Disordered Domains Are Preferentially Recruited to Polyglutamine Aggregates. *PloS one* 10:e0136362
233. Weiss A, Klein C, Woodman B, Sathasivam K, Bibel M, et al. 2008. Sensitive biochemical aggregate detection reveals aggregation onset before symptom development in cellular and murine models of Huntington's disease. *Journal of neurochemistry* 104:846-58
234. Westermarck P, Benson MD, Buxbaum JN, Cohen AS, Frangione B, et al. 2005. Amyloid: toward terminology clarification. Report from the Nomenclature Committee of the International Society of Amyloidosis. *Amyloid : the international journal of experimental and clinical investigation : the official journal of the International Society of Amyloidosis* 12:1-4
235. Wiederhold E, Veenhoff LM, Poolman B, Slotboom DJ. 2010. Proteomics of *Saccharomyces cerevisiae* Organelles. *Molecular & cellular proteomics : MCP* 9:431-45
236. Williams AJ, Paulson HL. 2008. Polyglutamine neurodegeneration: protein misfolding revisited. *Trends in neurosciences* 31:521-8
237. Willingham S, Outeiro TF, DeVit MJ, Lindquist SL, Muchowski PJ. 2003. Yeast genes that enhance the toxicity of a mutant huntingtin fragment or alpha-synuclein. *Science* 302:1769-72
238. Wisniewski JR, Zougman A, Nagaraj N, Mann M. 2009. Universal sample preparation method for proteome analysis. *Nat Methods* 6:359-62
239. Wytenbach A, Swartz J, Kita H, Thykjaer T, Carmichael J, et al. 2001. Polyglutamine expansions cause decreased CRE-mediated transcription and early gene expression changes prior to cell death in an inducible cell model of Huntington's disease. *Human molecular genetics* 10:1829-45
240. Xia Q, Liao L, Cheng D, Duong DM, Gearing M, et al. 2008. Proteomic identification of novel proteins associated with Lewy bodies. *Frontiers in bioscience : a journal and virtual library* 13:3850-6
241. Xu G, Stevens SM, Jr., Moore BD, McClung S, Borchelt DR. 2013. Cytosolic proteins lose solubility as amyloid deposits in a transgenic mouse model of Alzheimer-type amyloidosis. *Human molecular genetics*
242. Yamanaka T, Wong HK, Tosaki A, Bauer PO, Wada K, et al. 2014. Large-scale RNA interference screening in Mammalian cells identifies novel regulators of mutant huntingtin aggregation. *PloS one* 9:e93891
243. Yang H, Li JJ, Liu S, Zhao J, Jiang YJ, et al. 2014. Aggregation of polyglutamine-expanded ataxin-3 sequesters its specific interacting partners into inclusions: implication in a loss-of-function pathology. *Sci Rep* 4:6410
244. Yao J, Ong SE, Bajjalieh S. 2014. Huntingtin is associated with cytomatrix proteins at the presynaptic terminal. *Molecular and cellular neurosciences* 63:96-100
245. Yong W, Lomakin A, Kirkitadze MD, Teplow DB, Chen SH, Benedek GB. 2002. Structure determination of micelle-like intermediates in amyloid beta -protein

-
- fibril assembly by using small angle neutron scattering. *Proceedings of the National Academy of Sciences of the United States of America* 99:150-4
246. Zainelli GM, Dudek NL, Ross CA, Kim SY, Muma NA. 2005. Mutant huntingtin protein: a substrate for transglutaminase 1, 2, and 3. *J Neuropathol Exp Neurol* 64:58-65
247. Zala D, Benchoua A, Brouillet E, Perrin V, Gaillard MC, et al. 2005. Progressive and selective striatal degeneration in primary neuronal cultures using lentiviral vector coding for a mutant huntingtin fragment. *Neurobiology of disease* 20:785-98
248. Zatloukal K, Stumpfner C, Fuchsbichler A, Heid H, Schnoelzer M, et al. 2002. p62 Is a common component of cytoplasmic inclusions in protein aggregation diseases. *The American journal of pathology* 160:255-63
249. Zeron MM, Hansson O, Chen N, Wellington CL, Leavitt BR, et al. 2002. Increased sensitivity to N-methyl-D-aspartate receptor-mediated excitotoxicity in a mouse model of Huntington's disease. *Neuron* 33:849-60
250. Zhou L, Wang H, Ren H, Hu Q, Ying Z, Wang G. 2014. Bcl-2 Decreases the Affinity of SQSTM1/p62 to Poly-Ubiquitin Chains and Suppresses the Aggregation of Misfolded Protein in Neurodegenerative Disease. *Molecular neurobiology*
251. Zhu M, De Simone A, Schenk D, Toth G, Dobson CM, Vendruscolo M. 2013. Identification of small-molecule binding pockets in the soluble monomeric form of the Abeta42 peptide. *The Journal of chemical physics* 139:035101
252. Zuccato C, Valenza M, Cattaneo E. 2010. Molecular mechanisms and potential therapeutic targets in Huntington's disease. *Physiological reviews* 90:905-81
253. Zybailov B, Mosley AL, Sardi ME, Coleman MK, Florens L, Washburn MP. 2006. Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. *Journal of proteome research* 5:2339-47