STRUCTURAL DYNAMICS OF HUNTINGTIN EXON1 AGGREGATION STUDIED BY SANS

A Dissertation Presented
For a Doctor of Philosophy
Degree
University of Tennessee, Knoxville

Tatiana Perevozchikova

May 2012

Copyright © 2012 Tatiana Perevozchikova
All Right Reserved

DEDICATION

This dissertation is dedicated to:

My loving Parents:

Nadezhda Petrova

&

Vladimir Petrov

And also to my son:

Gleb Perevozchikov

ACKNOWLEDGEMENTS

Thank you, Dr. Valerie Berthelier for giving me an opportunity of being a part of your laboratory and for teaching me how to be an effective scientist and hopefully successful researcher. The knowledge and experience you shared with me provided the stable ground for my PhD work and will benefit all my future endeavors. I appreciate the mentorship, guidance and encouragement you provided me with during the last four years. With all your patience and persistence, I learned how to set and reach high expectations throughout all aspects of my life.

I would like to express special gratitude to Dr. Christopher Stanley, who introduced me to the theory and application of neutron scattering and guided me through many other experimental approaches. His vast knowledge in the field of biophysics made my learning experience very diverse and fruitful. I will greatly miss our scientific discussions and I feel very happy to have a chance to learn from you. I am also thankful to Penny McWilliams-Koeppen and Monique LeMieux for helping me in everyday work in the lab and keeping me updated on the recent happenings in the real world.

I am really thankful to my committee members, Drs. Elizabeth Howell and Jonathan Wall for their great advice and concerns. Without your unlimited expertise I would not be able to finish my research work in the set time frame.

I would like to thank the Graduate School of Genome Science and Technology for providing me with an opportunity to learn from the greatest scientist from many fields of biology, working at UT and ORNL. I would like to say a special thank you to Dr. Von Arnim for being very understanding and flexible throughout my study at GST.

I also want to express my sincere gratitude to Dr. Hirsh Nanda for his help in computational coarse-grained modeling and suggestions in the stylistics and content of this work. I really appreciate all the support and encouragement he provided me with during the period of preparation for a thesis defense.

And finally, my biggest gratitude is going to my loving family. My son, Gleb, you are the most resilient, understanding and forgiving child. You have been my most sincere fan and follower growing up with me in the lab and spending your afternoons at the seminar lectures. I really love you and hope that all this time I could not afford to spend with you will be paid off later. I owe so much gratitude to my parents, Nadezhda and Vladimir Petrovs, who have shown a great patience and a strong belief in my success. It is impossible to express how much I appreciate your support and dedication. I am very grateful to my sister, Olga, for her encouragement and advice and for being with me in the hardest moments of my life. Thank you all for teaching me how to work hard and do my best but remain a warm and caring person.

ABSTRACT

Huntington's disease (HD) is a genetic neurodegenerative disorder associated with the mutant huntington (Htt) protein containing an abnormally long stretch of glutamine (polyQ) residues. Upon proteolityc cleavage, this protein forms a htt-exon1 fragment that aggregates into highly stable and organized β-sheet structures. Currently, no clear evidence exists to determine if Htt protein aggregates are epiphenomena, if they are beneficial to or pathogenic for neurons. Although the correlation between the length of polyQ repeats, their propensity for aggregation, and disease is undeniable – the longer the polyQ region, the earlier the onset of HD and its symptoms are more severe – several cell and HD animal models studies demonstrated an absence of a link between aggregate presence and neuronal toxicity. It was proposed that neuronal toxicity is associated with the early stages of protein fibril formation and that mature aggregates actually represent an inert end stage, serving as a rescue mechanism. At present, a detailed understanding of the structures of different intermediate species, formed both on- and off-pathway to Htt fibril formation, is not established. Molecular insights in amyloid research and protein aggregation suffer from fundamental difficulties in controlling the formation of early intermediate assemblies and characterizing them. Within the scope of this research project we performed a comprehensive analysis on the aggregation of the htt-exon1 fragments containing normal (22Gln) and pathological (42Gln) length of polyglutamine repeats. To unravel their aggregation pathways, we performed time-resolved small-angle neutron scattering (TR-SANS) with *ab-initio* reconstruction approaches. We obtained

the 3-D structures of the earliest intermediates, followed their progression into protofilbrills and obtained the internal composition of the mature fibrils. Using SANS and other biophysical techniques we found that the length of polyGln repeat within the htt-exon1 fragment not only affects the kinetics of aggregate formation, but also drastically influence the structural dynamics and mechanism of aggregation. We also performed osmotic stress experiments coupled with contrast variation techniques to determine quantitatively the differences in the internal structures of the mature fibrils. The described results illustrate the utility of SANS for identification of various intermediates associated with amyloid and neurodegenerative diseases.

TABLE OF CONTENTS

Chapter 1: **Literature Review..... *1***

1.1 Amyloid fibrils..... 1

1.1.1 Overview.....5

1.1.2 Properties and structure of amyloid aggregates.....7

1.1.3 Mechanism of amyloid formation.....16

1.2 Neurodegenerative diseases..... 26

1.2.1 Overview 26

1.2.2 Mechanism of neuronal death..... 26

1.2.3 Role of inclusions..... 29

1.2.4 Cellular protective methods..... 32

1.3 Polyglutamine diseases 33

1.3.1 Overview.....33

1.4 Huntington's disease.....39

1.4.1 Overview.....39

1.4.2 Historical background.....40

1.4.3 Clinical findings.....41

1.4.4 Clinical genetics.....45

1.4.5 Neuropathology in HD.....50

1.4.6 Modeling HD.....54

1.5 Huntingtin protein.....63

1.5.1 Structure.....63

1.5.2 Huntington's cleavage……65

1.5.3 Post-translational modification……67

1.5.4 Huntingtin functions……69

1.5.4 Studies on polyQ aggregation in vitro……71

Chapter 2: **Major Methodology*****……76***

2.1 SANS as a powerful tool to study protein aggregation and fibrillization…76

2.1.1 Overview……76

2.1.1 Investigation of protein aggregation mechanisms by SANS……80

2.2 The Theoretical background on SAS……81

2.2.1 Underlying physics of SANS……81

2.2.2 Contrast variation……87

2.2.3 *Ab initio* reconstruction……89

Chapter 3:**Results and Discussions*****……92***

3.1 Structural Formation of Huntingtin Exon 1 Aggregates Probed by Small-Angle Neutron Scattering……92

3.1.1 Introduction……92

3.1.2 Materials and methods……94

Peptide solubilization and disaggregation……94

High Performance Liquid Chromatography……94

Thiaflavin T fluorescence……95

Congo Red birefringence……95

Transmission Electron Microscopy……96

Small-angle neutron scattering……96

3.1.3 Results and discussion……98

$NtQ_{42}P_{10}$ aggregates formed at high concentration are amyloid-like…..98

Quantifying the $NtQ_{42}P_{10}$ aggregates by SANS……100

Early-stage aggregate structural growth……104

Later-stage two-dimensional fibril growth……111

Fibril structure determined from SANS……116

3.1.4 Conclusions……120

3.2 Glutamine Expansion Length Dictates Structural Properties of Huntingtin Exon 1 Assembly……122

3.2.1 Introduction……122

3.1.2 Experimental Procedures……124

Preparation of peptide……124

RP-HPLC……125

ThT Fluorescence assays……125

TR-SANS and SANS measurements……126

TR-SANS analysis……127

TEM……128

3.2.3 Results……129

$NtQ_{22}P_{10}$ does not form dimer/trimer intermediates.……129

The GNOM-calculated Pair-distribution functions confirm the previous structural results for $NtQ_{22}P_{10}$ and $NtQ_{42}P_{10}$ early-oligomers ……134

Ab initio experiments allow shape reconstruction of $NtQ_{22}P_{10}$ and $NtQ_{22}P_{10}$ fibrils demonstrate distinctive structural features.....136

$NtQ_{22}P_{10}$ fibrils demonstrate distinctive structural features.....139

3.1.4 Conclusions.....141

Chapter 4: **Conclusions and suggestions for the future work*****.....144***

4.1 $NtQ_{22}P_{10}$ and $NtQ_{42}P_{10}$: similar peptides, but different mechanisms of aggregation 144

4.2 Selected methodology to study htt exon1 aggregation.....147

4.3 Future directions.....149

4.3.1 X-ray fiber diffraction studies 149

4.3.2 Selective deuteration/contrast matching studies.....150

4.3.3 Transition to a full-length htt exon-1 fragment.....151

List of References:**.....153**

Appendix**.....171**

Vita**.....184**

LIST OF TABLES

Table 1.1: Human amyloid-associated diseases……5

Table 1.2: The major characteristics of proteins associated with polyQ diseases and the observed clinical manifestations of these disorders…… 34

Table 1.3: Potential functions and native cellular location of polyglutamine proteins, causing cellular toxicity in the CNS..... 37

Table 1.4: Rodent genetic models designed to study Huntington's disease…… 55

Table 3.1: Aggregation kinetics and structural parameters of $NtQ_{22}P_{10}$ and $NtQ_{42}P_{10}$ peptides obtained from Guinier analysis of the SANS profiles……132

Table 3.2: Structural parameters of $NtQ_{22}P_{10}$ and $NtQ_{42}P_{10}$ peptides derived from P(r) function.....132

LIST OF FIGURES

Figure 1.1: Biological factors influencing protein misfolding and aggregation in vivo.....4

Figure 1.2: TEM images of Aβ1–40 aggregates.....10

Figure 1.3: The cross-sectional organization of amyloid aggregates..... 11

Figure 1.4: The overview of the three-dimensional structures of fibrillar aggregates formed by various amyloids. 13

Figure 1.5: Idealized schematic representation of a general protein aggregation pathway.....19

Figure 1.6: Nucleation barrier..... 21

Figure 1.7: Free-energy picture for downhill polymerization..... 25

Figure 1.8: Cytoplasmic, intranuclear and extracellular inclusions revealed by electron microscopy..... 31

Figure 1.9: EMG recordings of chorea and motor impersistence in HD patients.....44

Figure 1.10: Observed correlation between CAG repeat expansion and age of onset and severity of HD..... 47

Figure 1.11: Composite graphs illustrating the age of the onset against the number of CAG repeats in 8 polyQ disorders..... 49

Figure 1.12: Hypothesized time line of alteration in striatal D2 binding, pattern expression and localized metabolism observed in the presymptomatic phase of HD.....53

Figure 1.13: Schematic illustration of the huntingtin amino acid sequence 66

Figure 1.14: A model for simple polyQ aggregate nucleation and elongation.....72

Figure 2.1: The association pathway of insulin fibrils.....79

Figure 2.2: SANS experimental set-up and data analysis. 85

Figure 3.1: Aggregation kinetics of $NtQ_{42}P_{10}$ and the resulting fibrils.....101

Figure 3.2: Time-resolved SANS on $NtQ_{42}P_{10}$.....102

Figure 3.3: Time-resolved SANS data analysis on $NtQ_{42}P_{10}$ aggregation 105

Figure 3.4: Early aggregates of $NtQ_{42}P_{10}$ formed within the first 71 min..... 107

Figure 3.5: Structural insights on $NtQ_{42}P_{10}$ aggregation.....110

Figure 3.6: Aggregates of $NtQ_{42}P_{10}$ beyond 1 hr are predominately fibrillar and their two-dimensional cross-section is analyzed.....112

Figure 3.7: Prediction of the internal organization of $NtQ_{42}P_{10}$ fibril.....114

Figure 3.8: SANS data obtained over a larger *Q*-range allow the $NtQ_{42}P_{10}$ fibril structure to be more carefully probed..... 117

Figure 3.9: Time-resolved SANS on aggregation of $NtQ_{22}P_{10}$.....130

Figure 3.10: Guinier analysis on the early stages of $NtQ_{22}P_{10}$ aggregation observed by TR-SANS..... 133

Figure 3.11: Dynamics of monomer loss and amyloid aggregate formation for normal and pathological-like htt exon-1 peptides.....135

Figure 3.12: P(r) analysis of scattering data on the early $NtQ_{22}P_{10}$ and $Nt_{42}P_{10}$ intermediates 131

Figure 3.13: 3-D models of the early intermediates formed at the initial stages of $NtQ_{42}P_{10}$ and $NtQ_{22}P_{10}$ aggregation.....138

Figure 3.14: SANS data on $NtQ_{22}P_{10}$ fibril growth.....140

Figure A.1: The $NtQ_{42}P_{10}$ zero-angle neutron scattering intensity ratio, $I_{c,s}(0)/I_{c,0}(0)$; as a function of TEG volume fraction, fv.....138

Figure A.2: Measured change in $NtQ_{42}P_{10}$ R_c versus solute f_v.....140

CHAPTER 1.

Literature Review.

1.1 Amyloid Fibrils

1.1.1 Overview

Numerous proteins and peptides show a natural tendency to self-assemble into fibrillar, β-sheet-rich structures termed amyloid fibrils. Functional amyloids have been found in bacteria, fungi and mammals with roles as diverse as biofilm formation, development of aerial structures, scaffolding, regulation of melanin and activation of haemostatic factors, among others [1-3]. Nevertheless, not all amyloid fibrils are connected to functional events; they are often hallmarks of different diseases [4]. Fibrils of extracellular amyloid beta peptide (Aβ) and neurofibrillary tangles formed by cytosolic tau protein are the main constituents of the neuritic plaques formed in the brains affected by Alzheimer's disease. Parkinson's disease, Huntington's disease and spongiform encephalopathies are other neurodegenerative pathologies characterized by the presence of amyloid fibrils of α-synuclein, huntingtin and prion protein, respectively (Table 1.1). Apo serum amyloid A and β2-microglobulin are further examples of proteins whose accumulation as fibrillar aggregates in different tissues is related to amyloidoses.

One of the leading causes for the formation of amyloid fibrils is mutation, ultimately leading to aggregation (Figure 1.1) ([4-8].) Another factor contributing to the aggregation is an increased concentration of the protein: there is a direct correlation between cell degeneration and the level of the expression of amyloid proteins, such as Aβ and α-synuclein [9]. The occurrence of the unfolded state of natively globular proteins, as observed immediately after translation, also offers a significantly large pool of aggregation-prone species [10]. An additional biological factor, potentially contributing to formation of aggregates is imbalance between generation of oxidants and antioxidants in the system [11]. In vivo, oxidation of protein can cause cross-linking, significantly increasing the propensity of proteins to aggregate [12], however, the direct relationship between the formation of amyloid and oxidative stress has not been found. Since the majority of protein-deposition disorders demonstrate a late-onset nature, the obvious explanation for this phenomenon can be the key role aging plays in promoting protein aggregation in the organism (Figure 1.1). Both inside and outside the cell, proteins must undergo conformational changes, at least partially, to aggregate into amyloid fibrils. Entirely or somewhat unfolded intermediates on the pathway to their native states (or produced as a result of chemical variations, stress conditions or a genetic mutation) are especially susceptible to aggregation. Proteins and fragments that are intrinsically disordered or peptides created by the proteolytic cleavage and not capable of folding without the remainder of the sequence, can also form aggregates under particular circumstances, such as their elevated concentration. Sometimes early amorphous aggregates dissociate, but some other species can reorganize to conform oligomers, performing as a structural

base for the formation of subsequent spherical, annular or protofibrillar amyloid intermediates found in many aggregation systems. A crucial numbers of the molecules must be present to form create a long-range order in the mentioned above structures, so that the favorable enthalpic contribution from their organized stacking can actually offset the associated loss of free rotational entropy [4]. The complex presence of the before mentioned biological and thermodynamic factors increase the protein's propensity to aggregate, making it toxic to cells. As of today there is number of human diseases, which is believed to arise from the formation of fibrillar amyloid aggregates (Table 1.1). Within this work, I will focus on cerebral amyloids.

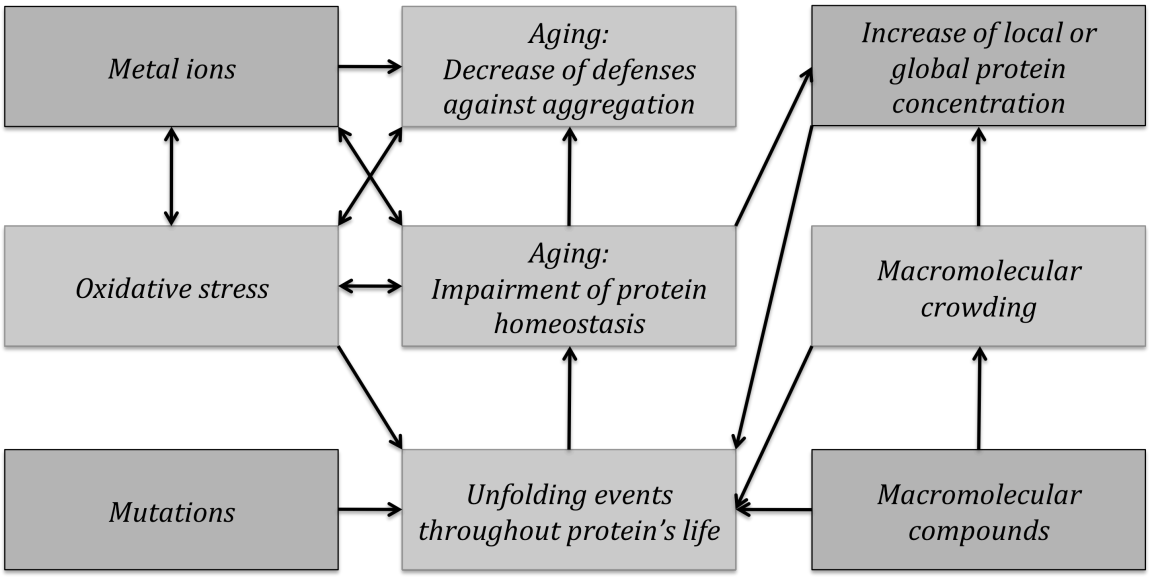


Figure 1.1. Biological factors influencing protein misfolding and aggregation in vivo. (Adapted from Campioni, Protein Misfolding Diseases, 2010.) Some aspects influence protein aggregation directly, and therefore their effects can be studied in vitro (dark gray rectangles.) Other factors act at the higher, organism's level, engaged in a significant level of complexity (light grey rectangles.) Arrows indicated associations between the factors.

Table 1.1: Human amyloid-associated diseases.

Amyloid condition	Aggregating peptide or protein	Native structure(s) of peptide or protein	Affected tissue/ syndrome
Cerebral (Localized)			
Expanded polyglutamine diseases	Various polyglutamine containing proteins (CAG)	Intrinsically unfolded; β- sheet; α-helices; combination α+β; unknown (See Table 1.2)	Cerebral cortex, Purkinje cells, basal ganglia, etc.
Alzheimer's disease	β-amyloid precursor protein (Aβ)	Intrinsically unfolded	Cerebral cortex, hippocampus, basal forebrain, etc.
Spongiform encephalopathies	Prion protein (APrP)	Intrinsically unfolded; α- helices	Brain
Parkinson's disease	α-synuclein	Intrinsically disordered	Substantia nigra
Amyotrophic lateral sclerosis	Superoxide dismutase 1	β-sheet	Motor neurons
Pituitary prolactinoma	Prolactin (APro)	α-helices, 4-helical cytokines	Pituitary
Icelandic cerebral amyloid angiopathy	Cystatin C (ACys)	α+β, cystatin	Brain
Peripheral: Systemic (S) or Localized (L)			
AL amyloidosis – Light chain (S,L)	Immunoglobulin light chains (AL)	β-sheet	Skin, nerves, heart, GI tract (includes tongue), kidney, liver, spleen, blood vessels
AH amyloidosis – Heavy chain (S,L)	Immunoglobulin heavy chains (AH)	β-sheet	Myeloma-associated primary tumors
AA amyloidosis – Inflammation-associated (S)	Serum amyloid A protein (AA)	α/β	Spleen, liver, kidneys, adrenals, lymph nodes. vascular Secondary to chronic infections, reactive tumors
Familial & senile systemic amyloidosis (S,L)	Transthyretin (ATTR)	β-sheet	Predominately liver, most organs & tissues

Continued from table 1.1

Amyloid condition	Aggregating peptide or protein	Native structure(s) of peptide or protein	Affected tissue/ syndrome
Hemodialysis amyloidosis (S,L)	β_2-microglobulin (Aβ_2M)	β-sheet	Musculoskeletal, GI tract
Lysozyme amyloidosis (S)	Lysozyme (ALys)	α+β, lysozyme fold	Visceral organs & tissues
Familial, sporadic & aortic (S,L)	Apolipoprotein (AApoAI, AApoAII, AApoAIV)	α-helices	Visceral organs & tissues
Familial amyloidosis (S)	Gelsolin (AGel)	α+β, gelsolin-like	Eyes
Hereditary fibrinogen α-chain amyloidosis (S)	Fibrinogen α-chain (AFib)	Coil?	Kidney
Familial British dementia (S)	ABriPP (ABri)	Majority coil, some α+β	Central nervous system
Familial Danish dementia (S)	ADanPP (ADan)	Predicted β-sheet	Central nervous system
Diabetes mellitus Type II (L)	Islet amyloid polypeptide (AIAPP) or amylin	Natively unfolded	Islets of Langerhans (pancreas)
Medullary thyroid carcinoma (L)	Pro(calcitonin) (ACal)	50% coil : 50% α-helix	C-cell thyroid tumors
Isolated atrial amyloidosis	Atrial natriuretic factor	50% coil : 50% α-helix	Heart
Iatrogenic insulin-dependent amyloidosis (L)	Insulin (AIns)	Largely α-helix	Liver, kidney, spleen, pancreas
Senile aortic, medin amyloidosis (L)	Lactadherin (AMed)	Coil	Heart
Familial corneal amyloidosis	Kerato-epithelin (AKer)	Largely α-helix	Eyes
Familial corneal amyloidosis	Lactoferrin (ALac)	α+β combination	Eyes

Adapted from Protein misfolding diseases: current and emerging principles and therapies, 2009.

The table is based on the Structural Classification of Proteins (SCOP), Protein Data Bank (PDB), or Advanced Protein Secondary Structure Prediction Server (APSSP).

1.1.2 Properties and structure of amyloid aggregates

The formed amyloid aggregates consistently demonstrate fibrillar characteristics when observed by transmission electron microscopy (TEM) and atomic force microscopy (AFM) [13](Figure1.2). These techniques also demonstrated the protofilament substructure of amyloid fibrils, which generally consist of 2 to 6 filaments, with a diameter of 2 to 5 nm (Figure1.3). The filaments often twist around each other to form a final 7-13 nm structure or associate laterally producing thick ribbons reaching 30 nm as a maximum width [14]. As of today, amyloid fibrils composed from different proteins and peptides have been studied using various techniques, however the majority of the detailed information available on the internal structure of the mature fibrils came from aggregates formed by Aβ fragment, associated with Alzheimer's disease, as its aggregation pathway was an initial system under intensive investigations [15].

A. Identification of amyloid fibrils using fluorescent techniques.

The formation of amyloid fibrils is commonly identified by several fluorescent techniques. Binding of Congo red dye, observed by the characteristic apple-green birefringence under crossed polarized light has been a most widely used diagnostic test for the presence of amyloid in tissues**.** In vitro, extrinsic fluorescence of the benzothiazole dye, thioflavin T (ThT), is extensively utilized for the detection and quantification of amyloid aggregates, being the primary technique for observing the

kinetics of fibril formation in real-time [16, 17]. Upon addition of ThT to samples containing β-sheet-rich deposits, such as β-sheet quaternary structure of amyloid aggregates, it demonstrates significant fluorescence with excitation maxima at 440 and emission at 490nm [16, 18]. The mechanism of interaction between ThT and amyloid is still unknown, the most currently supported theory suggests the intercalation of ThT molecules into the grooves flanked by solvent-exposed side chains of the fibril, running parallel to the main (fibril) axis [19-21]. Binding within the channels is proposed to supply rigidity and to avert the establishment of a less-fluorescent twisted internal charge-transfer state of the ThT molecule [22]. Since there is a detectable stoichiometric and saturable contact between ThT and amyloid fibrils, fluorescence arising from the amyloid–ThT complex provides robust quantification of amyloid growth as a function of fibril length and number [17].

B. High-resolution structure of mature fibrils

Information on the high-resolution structure of fibrils is important for unveiling the mechanism of amyloid formation and could show a direction for the development of agents potent in affecting this process. However, no detailed molecular structure of the amyloid fibrillar aggregates has been obtained using traditional X–ray crystallography and solution NMR, due to fibrils' intrinsic insolubility and a significant length distribution. Although electron microscopy can be potentially used to describe the fibrils' structure, the major pitfall lies in their relatively small radius.

The primary technique, originally used for studying the internal organization of mature fibrils is X-ray diffraction. The sample containing amyloid aggregates is

prepared by aligning the fibrils into a macroscopic fibre and the obtained diffraction provides information on the β-sheet organization within the formed amyloid. This diffraction pattern usually contains two main reflections located orthogonally to each other: at 4.7 Å (length of the β-sheet repeat, sharp peak) and 10 Å (intersheet spacing, broad peak) and signifies the perpendicular arrangement of β-strands relative to the fibril axis. This internal cross-β-sheet packing has been detected for synthetic and *ex-vivo* Aβ fibrils [15, 23, 24], Met 30 variant of transthyretin [25], the amyloidogenic SH3 domain from phosphatidylinositol-3'-kinase [26], β_2-microglobulin [27] and other amyloidogenic fragments.

The opportunity to drastically enhance the information provided from X-ray diffraction studies and significantly improve our understanding on the molecular organization of amyloid aggregates came from a major progress in the utilization of solid-state NMR (SSNMR) spectroscopy to study of amyloid fibrils [28-30] and from successes in growing nano- or microcrystals of small peptide fragments that have features of amyloid fibrils yet are susceptible to single crystal X-ray crystallography [31, 32].

Using SSNMR technique, the Tycko group [28, 33, 34] proposed a structure of the amyloid fibrils formed from the 40-residue form of the amyloid β peptide (Aβ1−40. According to their structure, each Aβ1−40 molecule contributes a pair of β-strands to the core part of the fibrils (Figure 1.4.A). These strands take part in the formation of two discrete β-sheets constituting same protofilament [35]. Several Aβ molecules are placed onto each other, in a parallel manner and in register [33, 34].

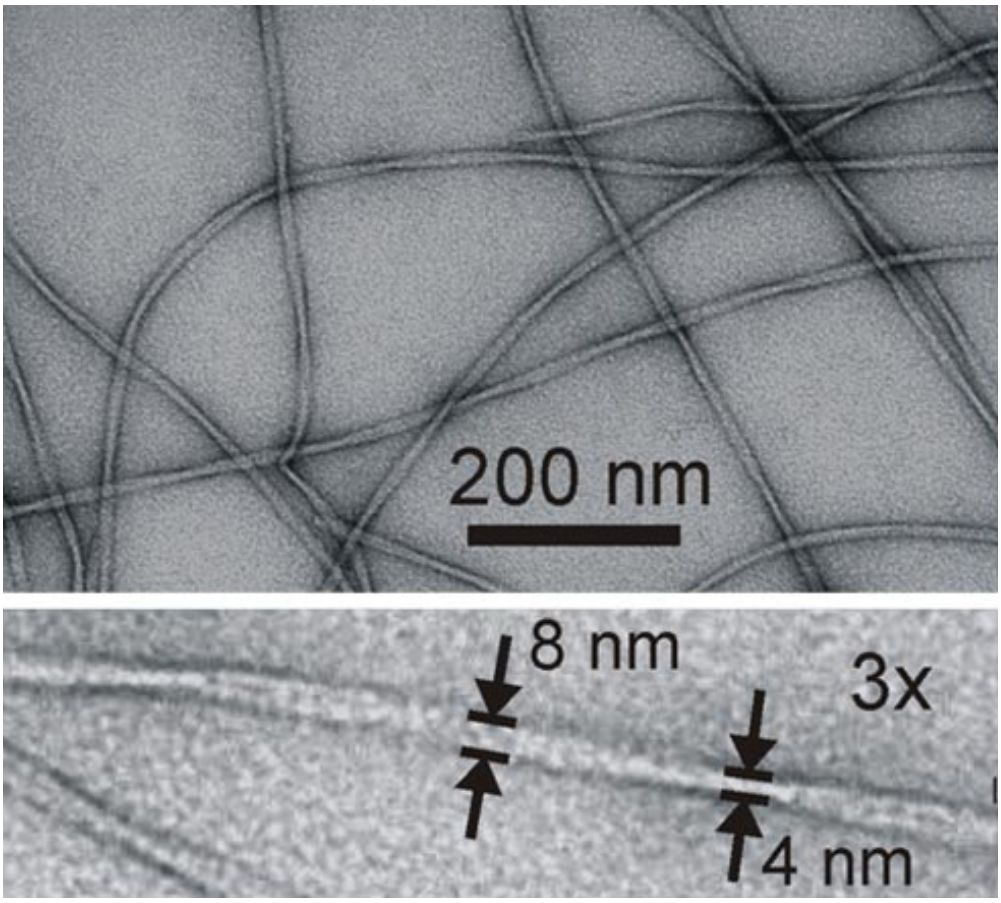


Figure 1.2. TEM images of Aβ1–40 aggregates. (Adapted form Paravastu, PNAS, 2008).

Negatively stained electron microscopy image of Aβ1–40 aggregates, demonstrating the fibrillar nature and twisted morphology with some noticeable width variation.

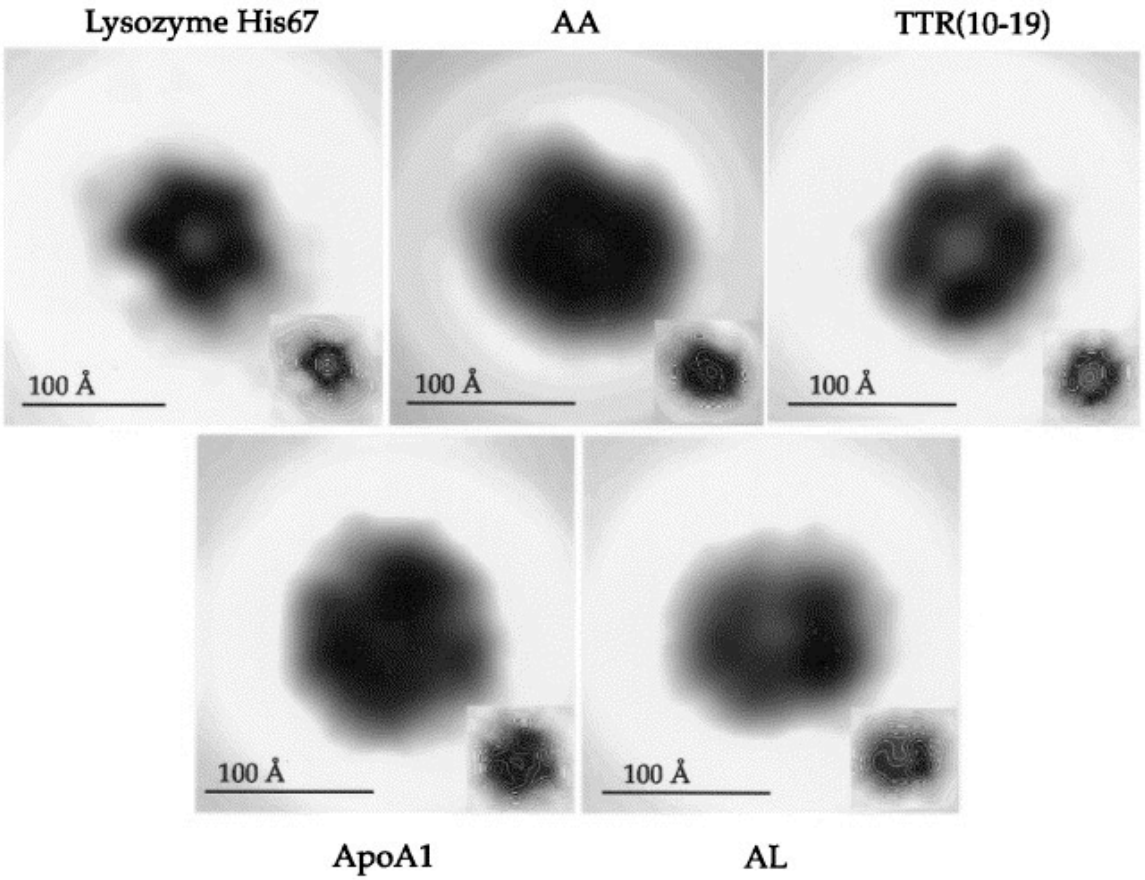


Figure 1.3 The cross-sectional organization of amyloid aggregates. Images taken from (Serpell, JMB, 2000). The reconstructed substructurte of ex vivo amyloid fibrils obtained by cross-correlation methods on the averaged electron microscopy images of: Asp67His variant lysozyme, amyloid A protein, a synthetic ten-residue peptide corresponding to the A-strand of transthyretin, Leu60Arg variant apolipoprotein A1 and monoclonal immunoglobulin λ light chain. The results are produced by centering and aligning particles for each amyloid. Final images demonstrate the center hollow core of each particle and the organization of protofilaments around it. The inset shows the contoured image to emphasize the regions of the increased density.

Inclusion of the additional experimental constraints, such as the diameter of the protofilaments and the mass per unit length obtained by TEM, and scanning transmission electron microscopy (STEM) respectively [28, 36], yielded a proposal that a single protofilament consists of four β-sheets (Figure 1.4.A).

The major support for the proposed model was provided by experiments of site-directed spin labeling coupled to electron paramagnetic resonance (SDSL-EPR)[37]. The observed EPR spectra indicated the values of the inverse central line width for a series of labeled residues, which suggested that the regions of the Aβ1−42 molecule spanning residues 13–21 and 30–39 represent structured, parallel and in register elements in the fibrils.

SSNMR, in combination with site-directed fluorescence labeling and hydrogen/deuterium exchange experiments was used to identify the portions of β2-microglobulin fibrils that contribute to repeating structure [38], and led to elucidation of the regions of the C-terminal fragment of fungal prion HET-s that are involved in the formation of the fibrils' core [30]. In the given structure, each molecule yields four β-strands, with the first and third strands belonging to the same parallel β-sheet and second and forth strands contributing to another parallel β-sheet, located at the distance of ∼10 A° away (Figure 1.4b).

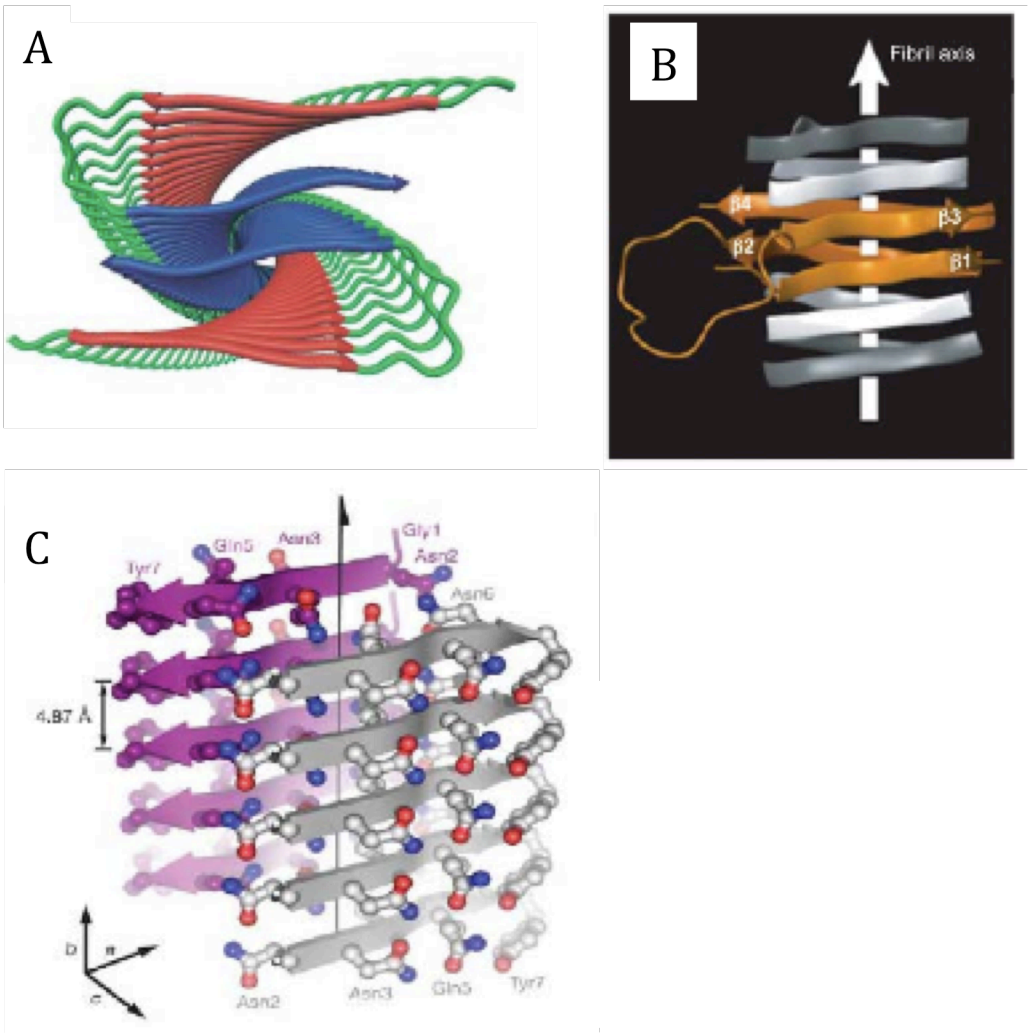


Figure 1.4. The overview of the three-dimensional structures of fibrillar aggregates formed by various amyloids. (A) The protofilament formed by Aβ and viewed down the long axis of the fibrils. (Adapted from Tycko, R. Biochemistry, 2006). The shown segments are 12-24 (red) and 30-40 (blue). (B) Fibrils formed by the C-terminal fragment 218-289 of the fungal prion HET. (Ritter, C. Nature, 2005). The ribbon diagram demonstrates the beta-strands spanning residues 226-234, 237-245, 262-270, and 273-282 and the loop connecting β2 and β3 within one molecule. The residues flanking the fibrils are shown in grey. (C) The atomic structure of the microcrystals formed by GNNQQNY peptide. (Adapted from Nelson, R. Nature, 2005.)

The developments in SSNMR techniques, permitting the measurements of particular internuclear distances and torsion angles allowed obtaining an atomic resolution structure of amyloid-like fibrils formed by 11-residue fragment of transthyretin [28, 39]. This experiments revealed that this construct forms an extended β-strand conformation within the fibrils. The most important finding, although, is that the molecules forming the fibrils demonstrate a high degree of uniformity, including that of the side-chain torsion angles, a characteristic previously attributed solely to crystalline materials. Because this uniformity is clearly expressed as the very narrow resonance lines in the SSNMR spectra, it is expected that the atomic-level conformations will soon be obtained for a series of other aggregating systems, enriching our insights into this component of amyloid formation.

The structure of the Sup35p (GNNQQNY) and KFFEAAAKKFFE peptides and the way these molecules are arranged together was determined with 1.8 Å resolution after significant achievements in growing three-dimensional crystals retaining major properties of amyloid fibrils have been made [31, 32]. The crystal of the Sup35p fragment contains pairs of parallel β-sheets where each individual peptide molecule produces a single β-strand (Figure 1.4.C). Both sheets consist of stacked parallel and in register β-strands. The outer faces of the sheets are hydrated and placed away from the next pair of β-sheets, possibly resulting from a crystal contact rather being a feature of the fibrillar structure.

The described structures, as observed by X-ray and SSNMR are noticeably similar to proposed model of the amyloid fibrils formed from an SH3 domain and from insulin by cryo-electron microscopy (EM) analysis, in which electron density maps

were interpreted as the result from pairs of sufficiently flat β- sheets[26, 40]. The apparent similarity between structural features of many amyloid fibrils could create a plausible suggestion of a common internal organization for all amyloid fibrils, probably driven by the intrinsic conformational requirements of polypeptide chains. The observed pattern however, can also be attributed to the dehydrated nature of the fibrils, as a consequence of the chosen experimental methods. The alternative cross-sectional organization of amyloid aggregates can be observed if the solution-based techniques are utilized. For example, different models have been proposed to characterize amyloid fibrils' structure, including the water filled nanotube [41], the solenoid structures [15-17] and the stacking of protein in native conformations to form fibrils [18]. Another unconventional β-helical cross-sectional organization was proposed for the mature fibrils formed by the simple polyGln peptides [41]. The polyglutamine fibrils represent a particular interest due to the possible additional hydrogen-bonding interactions between the side chains, which can yield a structure drastically different from those described above. Evidence for the unconventional cross-sectional organization of these fibrils came from the X-ray fiber diffraction experiments with the discovery of the absence of the 10 Å peak from reflection patterns. The conformation and arrangement of the polyGln peptides within the formed fibrils will be extensively discussed within the following chapters of this document.

1.1.3. Mechanisms of amyloid formation

The significant similarity in the final structure of the mature aggregates, both functional and pathogenic, might suggest the presence of the common aggregation mechanism for many amyloidogenic proteins. For the purposes of studying protein misfolding and association pathways, the aggregation can be categorized into three classes [42]: (i) productive or naturally occurring aggregation, (ii) unwanted biological aggregation and (iii) unwanted industrial aggregation. An example of the first class is n(G-actin)→(F-actin)n polymerization, taking place throughout many types of organisms controlling the shape and the mobility of the cells[43]. The enzyme glutamate hydrogenase (GH) represents another protein undergoing the naturally occurring aggregation [44-46]. Both actin and GH function with polymerization of the protein when in its native state. The second type of aggregation includes α-synuclein, amyloid β, prions and polyglutamine expansions' amyloid formation and potentially causes diseases [42]. This category of aggregation involves the aggregation of the protein in the non-native state (vida infra.) The type (iii) aggregation often results in production of amorphous aggregates and its comprehension as well as control is highly important for biotechnology industry for keeping protein of interest in "bottleable", non-aggregated, long-shelf-life form [47].

Generally aggregation starts with a protein in its monomeric form. However, in the case of undesired aggregation, it can be initiated with protein either in its native or denatured form. It is known that several misfolded proteins participate in the first stages of the unwanted aggregation pathway[48]. One of the best-studied examples is

α-synuclein, which was shown to misfold and aggregate more rapidly under oxidative conditions, mimicking the oxidative stress within a cell[49]. Fink and co-workers have also demonstrated that α-synuclein obtains a partially folded forms as its aggregates[50]. This observation gave a rise to a hypothesis that in particular cases a protein monomer undergoes a conformational alteration resulting in a higher content of β-sheet in its structure or more exposure of the hydrophobic regions[50], which make this protein more prone to aggregate, yielding it being more "active"[51].

As it was described above, the aggregation of amyloidogenic proteins results in the formation of stable, β-sheet rich fibrils. In several cases, some amorphous aggregates have also been observed during the process and were proposed to be off-pathway products. There is still no consensus reached which intermediates can be considered "on" and which "off" the pathway of aggregation; the increasing body of evidence suggests that the early intermediates formed at the initial stages of aggregation are the primary suspects for neural toxicity. Two types of intermediates are currently included into this class: earlier oligomers and later formed protofibrils. The description of these intermediates as formed by polyglutamine constructs is one of the subjects studied within the scope of current work and described in chapter 3.

A. Nucleated Polymerization Aggregation.

In the kinetic models, the protein aggregation is usually described in terms of polymer formation; and its mechanism is usually related to the rate of polymer formation, which, in some cases, is proposed to be governed by nucleation (Figure 1.5.). In the naturally occurring aggregation, a nucleus represents a stable entity that

serves as a site for the initiation of the polymer formation at the biologically relevant site. In the case of unwanted, pathological aggregation, the nucleus population might be an unstable species, formation of which represents a limiting step of the reaction. Figure 1.6 demonstrates the free energy change (ΔG) present during the process of forming aggregates via nucleation pathway. The free energy of monomer is zero by definition. Unfavorable species are described by the positive free energy and the presence of the unfavorable regime signifies the barrier that needs to be overcome in transition from monomers to long polymers. The aggregates demonstrating a positive free energy are suppressed comparing to monomers and are difficult to observe. The presence of the described barrier accounts for the effect of seeding, characteristic for nucleation polymerization: introducing short polymers to a concentrated monomer solution permits these polymers to elongate immediately, without the kinetic barrier that induce the slowdown of the initial stages of their production. Another important feature for identification of nucleation-driven aggregation is a concentration dependence of the rate constant, which yields very important information on critical nucleus size and nucleation equilibrium constant.

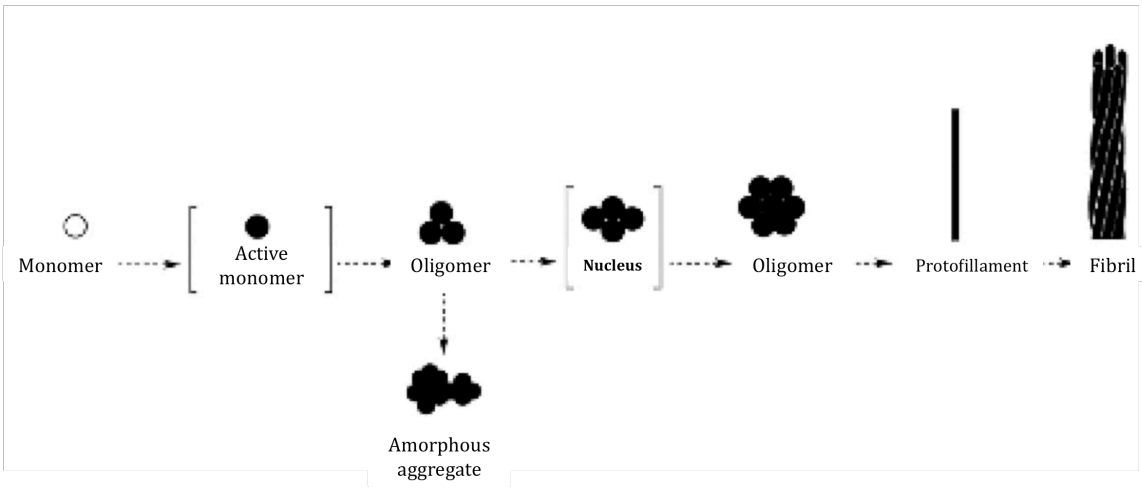


Figure 1.5. Idealized schematic representation of a general protein aggregation pathway. (Adapted from Morris et al., Biochimica et Biophysica Acta, 2009.)

The scheme includes all putative intermediates, shown in brackets, potentially present in several mechanisms: in several cases, monomeric protein has to undergo conformational changes to obtain an "active" form; nucleation formation step is a basic for nucleated polymerization mechanism. Protofilaments are represented in the generic way, since there is no agreement if they are formed by association of oligomers or the monomer addition.

One of the currently used models to describe nucleated polymerization as a mechanism of subsequent monomer addition, was proposed by Ferrone in 1999, where he characterizes the critical nucleus as an aggregate size, upon reaching which "the association rates exceed the dissociation rates for the first time.[52]" Scheme 1 represents the translation of Ferrone's equations into chemical reactions:

$$n^*M \underset{}{\overset{K_{n^*}}{\rightleftharpoons}} M_{n^*}$$

$$M + M_{n^*} \xrightarrow{k_+} M_{m>n^*}$$

$$M + M_{m>n^*} \xrightarrow{k_+} M_{m>n^*}$$

,

where M is the protein monomer, M_{n^*} represents the nucleus of size n*, $M_{m>n^*}$ is a polymer with a size m>n*, K_{n^*} is a the nucleation equilibrium constant, and k_+ is the growth rate constant. After excepting such assumptions as: (i) all polymers can be counted by their ends; (ii) the concentration of nuclei is very small, so that monomer depletion occurs mostly through polymer growth; and (iii) the concentration of the monomer incorporated into polymer should be small relative to the overall monomer concentration at the initial stages (10-15%) of aggregation, so that concentration of the free monomer can be considered to be constant and be equal to the original value, the proposed mechanism can be interpreted in terms of the equations:

(a) $[M_{n^*}] = K_{n^*}[M]\text{^}n^*$

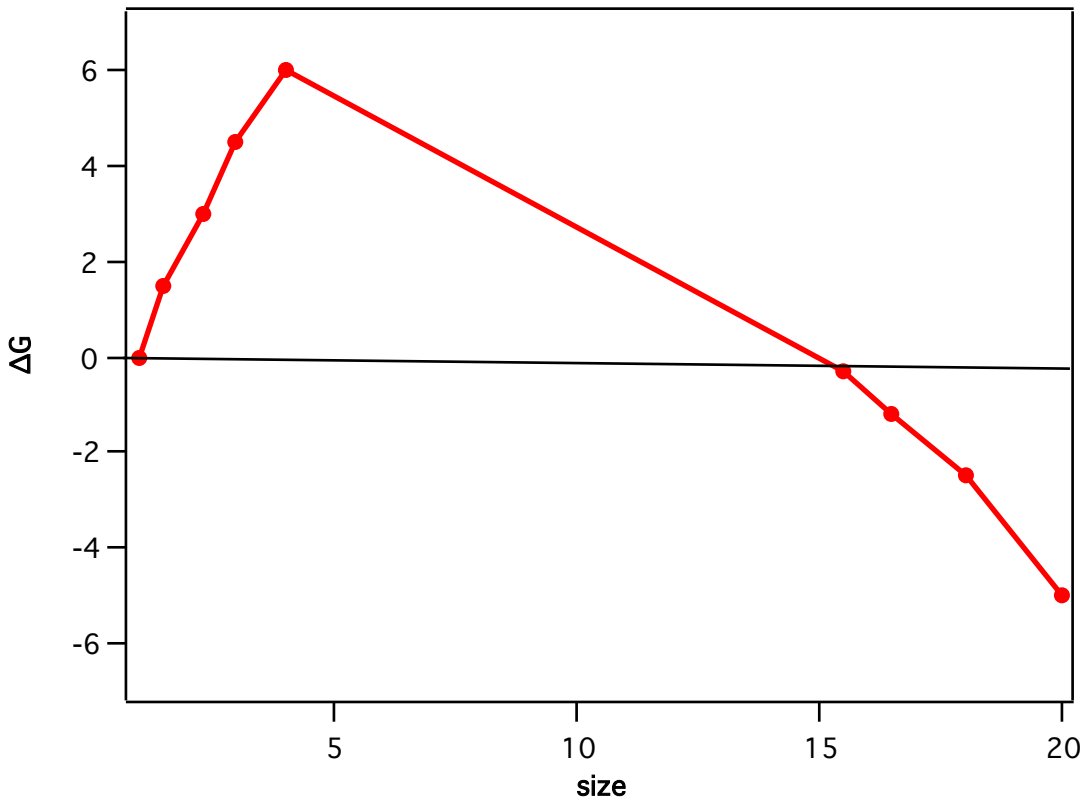


Figure 1.6. Nucleation barrier. (Adapted from Powers and Ferrone., Kinetic Models for Protein Misfolding and Association, 2010.)

The free energy change, ΔG, for forming a given aggregate relative to the monomer as a reference is demonstrated for aggregates of different size. The reason for the shown barrier arises from the changing balance between the entropy of the free motion of monomers versus the created bonds within the aggregate. As aggregates grow in size, more bonds are formed within the molecule, eventually offsetting the entropic cost that determines instability of smaller aggregates. Numbers on the Y-axis are given for visual guidance only.

(b) $$\frac{d[M_{m>n^*}]}{dt} = k_+[M][M_{n^*}] \qquad (1)$$

(c) $$-\frac{d[M]}{dt} = k_+[M][M_{mNn^*}]$$

The set of the equation (1) was applied toward polyglutamine aggregation, as described by the equation (2), where Δ is a concentration of the monomer incorporated in polymers, c0 is the initial concentration of the monomers, c represents the concentration of the free monomer, k+ is the second order elongation rate constant, Kn* is the nucleation equilibrium constant, and n* is the critical nucleus.

$$\Delta_{(t)} = c_0 - c_{(t)} = \frac{1}{2}k^2{}_+K_{n^*}c^{(n^*+2)}t^2 \qquad (2)$$

The equation (2) is applicable for the first stages of the kinetic curve, in which the concentration of the monomer is considered to be constant and equal to the original value. In the first 15% of the aggregation curve, a plot of $\{\Delta_{(t)}$ vs. $t^2\}$, provides with a slope of $[1/2k_+^2K_{n^*}c^{(n^*+2)}]$; and the slope of {ln(slope) vs. ln(c0)} yields [(n*+2)], while intercept gives [ln(1/2k+2Kn*)], providing information on the critical nucleus size and nucleation equilibrium constant respectively.

B. Downhill Polymerization Aggregation.

Downhill polymerization, also called association-limited aggregation, is another model for the aggregation-limited kinetics. According to this mechanism, the dimer formation is rate-limiting step in formation of the earliest irreversible species [53, 54], and also the rate coefficient of the subsequent growth is assumed to be the same as for

dimer formation (Figure 1.7.) The simplest downhill mechanism represents a system where prepared monomers are able to polymerize immediately. The rate of the polymer formation in this case is expressed as:

$$\frac{dp}{dt} = ka^2 \qquad (3),$$

where a is a concentration of monomers able to add to polymers and k is a dimer association rate constant. Using a pertubation approach, the equation (3) can be solved with a result:

$$\Delta(t) = \frac{1}{2}c_0(1 - \cos Bt) \qquad (4)$$

where $B^2=2kc_0J$, with c_0 being an initial concentration of monomers and J is a net elongation rate of polymers. A useful addition to downhill polymerization is a consideration of a slow conformational change, which forms species capable of associating directly and more rapidly than the conformational change[55]. In this case, the rate limiting step is conformational change, and the process represents a simple decay. Monomers are depleted at the rate $dc/dt = -k_{ca}C$ (c corresponds to the initial species) and their depletion correlates exactly with polymer formation:

$$\Delta = c_0\left[1 - \exp\left(-k_{ca}t\right)\right] \qquad (5)$$

This behaviour possesses a leading term proportional to t, with linear dependence on c_0. Such system, with a rate-limiting step of spontaneous conformational change of monomers, is not affected by the introduction of seeds.

The aggregation process of monomeric transthyretin variant (M-TTR) was studied at low pH and did not show any acceleration of reaction upon seeding,

suggesting the absence of nucleation stage in the growth[56]. Kinetic experiments supported this observation, when showed a linear (not t^2) early time course in TfT fluorescence. This led to a deduction that polymerization followed a downhill pathway.

Despite the importance of aggregation as biological and industrial problem and over 50 years of research on the determination of the mechanisms and rate constant for protein association, the general pathway of protein aggregation remains unclear. The mechanism was shown to vary from one amyloidogenic protein to another, sometimes even within the same protein with a slightest sequence variation[42].

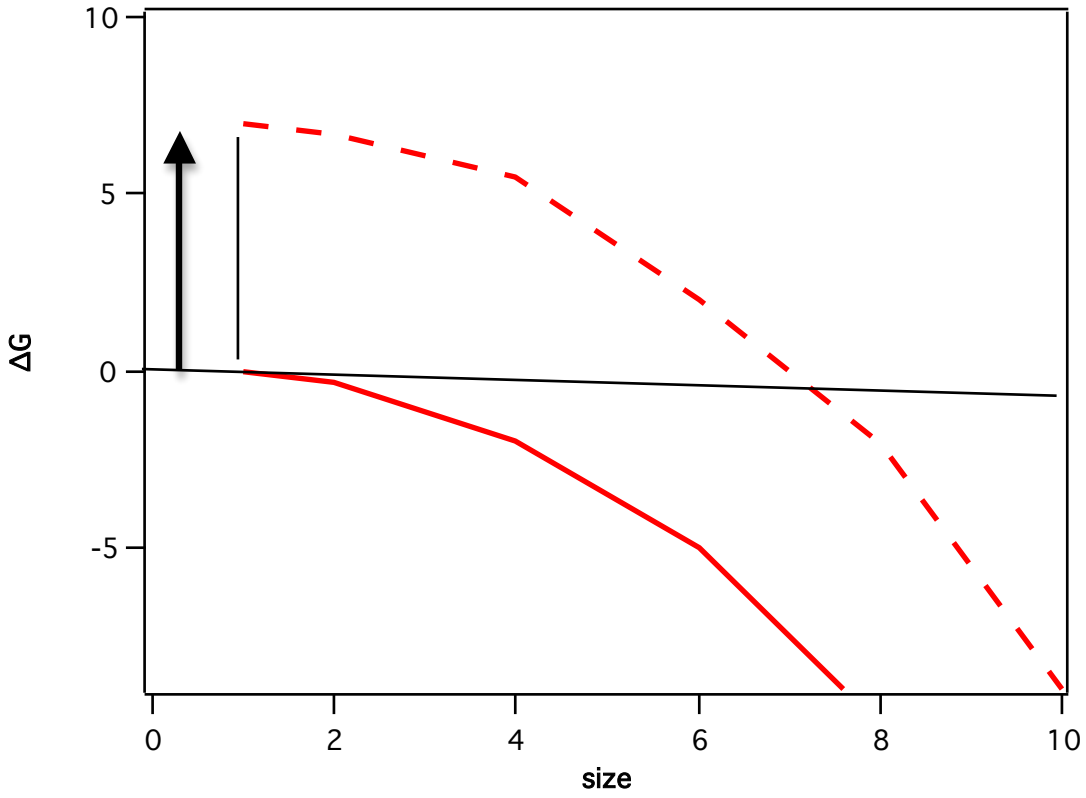


Figure 1.7. Free-energy picture for downhill polymerization. (Adapted from Powers and Ferrone., Kinetic Models for Protein Misfolding and Association, 2010.)

The solid curve demonstrates downhill polymerization in which each next species is more stable than previous ones. If the monomer initially present must undergo an unfavorable conformational change, the shift of the curve to the dashed position is observed. This represents a similar to the nucleation barrier, which a monomer has to overcome before the growth of larger aggregates can be started. Size, as indicated on the X-axis, represents oligomeric state of the formed aggregates. Numbers on the Y-axis are given for visual guidance only.

1.2. Neurodegenerative diseases as models to study protein-misfolding disorders.

1.2.1 Overview

The formation of amyloid aggregates is recognized as a major event of Alzheimer's (AD), Parkinson's (PD), Huntington's (HD), prion's and other neurodegenerative diseases, commonly observed simultaneously with selective loss of neural cells, synaptic alteration and inflammation [57]. However, the affected region of the brain varies in different diseases, which determines the typical clinical symptoms in patients. In AD, the neurodegeneration is observed mainly in the parts of the brain, responsible for long-term memory and abstract thinking, such as amygdalia, hippocampus and areas of entorhinal and neo-cortex [58]. PD is associated with the loss of neuronal cells in the subtantia nigra and dopamine deficiency in the striatum [59]. HD patients demonstrate significant neuronal loss in the areas of neostriatum and cerebral cortex [60].

1.2.2 Mechanisms of neuronal death

Generally, neuronal cell death occurs through programed cell death or apoptosis. There are three major hypotheses on how protein misfolding and aggregation can be correlated with apoptosis: the brain inflammation, the loss-of-function and the gain-of-function hypotheses.

According to the first hypothesis, the formed protein aggregates cause the wide-spread irritation and subsequent chronic inflammatory response in the brain, which

ultimately results in the neuronal death and alterations in the synaptic activity [61, 62]. However, the other research projects have proposed the beneficial nature of inflammation reaction in neurodegenerative diseases [62].

According to the loss-of-function hypothesis, the degeneration of the neural tissue is a consequence of the loss of the native activity of the protein, which is depleted during aggregation process. In Amyotrophic lateral sclerosis (ALS), it was suggested that generation of the reactive oxygen species (ROS) during the process of aggregation could be the possible pathological event, although, how these ROS get generated is not entirely clear [63]. Prion PrP^c was found to be associated with signal transduction pathways, as several of its domains interact with other signal proteins [64]. The binding of a ligand PrR to PrP^c serves as a neuroprotective signal in vitro [65]. Aggregation of this protein was proposed to block its interaction with the members of Bcl2 family of proteins, ultimately enhancing apoptosis [66]. α-synuclein demonstrates similar anti-apoptotic properties in vitro [67], suggesting that the improper balance between pro- and anti-apoptotic activities can be also attributed to other neurodegenerative diseases.

The gain-of-toxic-activity hypothesis proposes that the biggest impact with respect to brain damage in neurodegenerative diseases is produced by a neurotoxic functionality obtained by proteins upon misfolding, oligomerization and/or aggregation. This concept was suggested after observation of the direct induction of apoptotic pathways by introduction of the aggregates of different misfolded proteins in vitro [68-70]. Further support for this idea comes from the studies where the expression of a mutated gene causes neurodegeneration in transgenic animals [71, 72]. It was also shown that proteins oligomerizing to β-sheet conformation become

cytotoxic, indicating that the process of misfolding and aggregation can be inherently toxic [73].

Despite the large amount of evidence supporting a gain-of-function approach in causing cellular dysfunction and death, the pathway of its toxicity is still debated. A series of different molecular mechanisms have been proposed for multiple peptides, and it is possible that some particular pathways predominate depending on if the proteins accumulation occurs intra- or extracellularly [74]. A signal transduction pathway may be activated by extracellular aggregates leading to apoptosis by interacting with specific receptors within a cell. Incorporation of essential intracellular macromolecules, such as transcription factors, cytoskeletal components and chaperones by amyloids into mature aggregates has also been proposed during protein aggregation[75]. The most recent findings provide evidence for the formation of ion channels and subsequent damage to the membrane as a mechanism of amyloid pathology. The “channel hypothesis” on amyloidosis seems to provide an explanation for many physiological manifestations observed in amyloid diseases and follows the known biochemical properties of many amyloid peptides [76]. One of the most studied cases of interaction between amyloid and membrane is Aβ. In 1993, Arispe et al. demonstrated for the first time that this peptide is capable of forming ion channels within lipid bilayer membranes [77]. The formation of pores in the bilayers was also observed in the presence of other disease proteins associated with neurodegenerative diseases [78]. The general channel hypothesis proposes that a population of protofibrils, composed of several globular subunits, is able to penetrate the cell membrane, impairing its specific permeability and altering cell ion content [79-81]. The basis

behind this hypothesis includes several important observations. First, one of the earliest cellular modifications upon introduction of toxic species is the permeabilization of cell membranes, accompanied with nonspecific influx of Ca2+, implying the structural perturbations within the biological bilayers. Second, many proteins and peptides, irrelevant to their association to amyloid diseases, tend to form doughnut-shaped assemblies with hollow core. Third, the formation of amyloid-like pores via protein-peptide oligomerization in the membranes is a common way in which cells kill the foreign or competing cells, a process widely observed in many living organisms, from bacteria to humans [76, 82]. However, the changes in membrane permeabilization have been indicated in cells exposed to oligomeric and protofibrillar intermediates without any porelike appearance. Overall, it is accepted that the interactions of aggregates or their precursors can cause enough change in the structural stability of plasma membrane, creating the areas of disorders with inadequate ionic balance [83, 84].

1.2.3 Role of inclusions in neurodegenerative diseases

The resulting amyloid fibrils represent conformational manifestations of many neurodegenerative diseases, such as Parkinson's, Alzheimer's and, Huntington's. In these disorders, one of the main observations is an accumulation of the fibrils in supramolecular inclusions, generally called inclusion bodies in HD, plaques in Alzheimer's and Lewi's bodies in Parkinson's patients' brains, which are usually located in cytoplasm, nucleus or outside the cell [82, 85-89] (Figure 1.8). Besides the fibrillar aggregates in the amyloid form formed by a disease protein, these inclusions contain

other cellular material specific to a tissue and disease. Although inclusions serve as a common sign of neurodegenerative diseases, there is a still an argument on the role of their formation in the neural toxicity. Many studies have shown that abnormal protein conformations and aggregation are linked to the development of many diseases, however, the latter research efforts suggest that the formation of supramolecular inclusions is not correlated to the toxicity and represents a cellular protective mechanism [88]. These findings introduce a paradox: while aggregation is overall pathological, the formation of the inclusions is inert or even beneficial for a cell. The possible explanation to this contradiction might be the fact that inclusions represent an end-stage step of the aggregation process and the earlier species, such as abnormal monomer or small oligomers, represent the truly toxic species [88].

In Alzheimer's disease, for example, there is a very weak correlation between amount of Aβ amyloid plaques in the post-mortem human brain and the observed severity of the disorder [90]. Also in Parkinson's, only a little connection between presence of Lewy's bodies and neural death in the substatia nigra was observed [91]. Amyloid-like plaques and Lewy bodies can be found in patients without detected loss of neurons and other clinical signs of AD and PD [92].

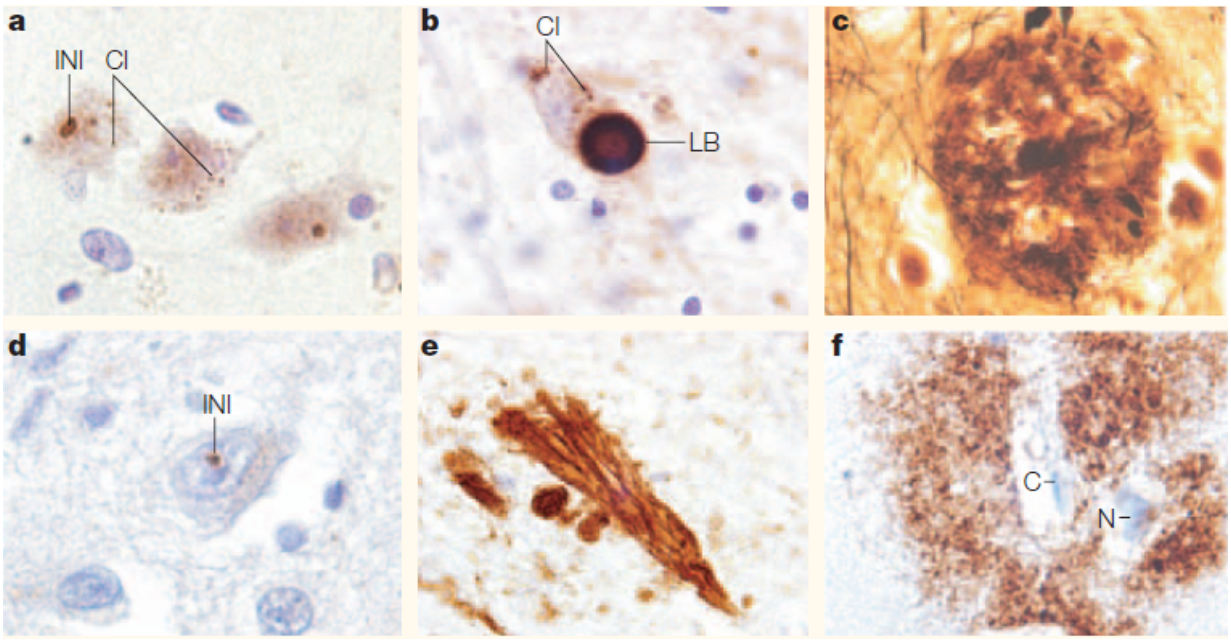


Figure 1.8. Cytoplasmic, intranuclear and extracellular inclusions revealed by electron microscopy. (Adapted from Ross C. A. Nature Review. 2005.)
a.C12 antibody to extended stretch of polyglutamine recognizes intranuclear inclusions (IN) and cytoplasmic inclusions (CIs) in the brain of HD patient. b. Alpha-synuclein specific antibody recognizes Lewis bodies (LB) and other CIs in the substantia nigra of PD brain. c. Intracellular (neuritic) plaque of AD in the cerebral cortex, recognized by hirano silver stain. d. Intranuclear inclusions located in the frontal cortex of HD brain immunostained by anti-ubiquitin antibody. e. Neurofibrillary tangels of AD hippocampus identified by antibody for phosphorylated tau. f. Diffuse plaque in cerebral cortex of AD patient. Extracellular deposits around neuron (N) and capillary (C) recognized by Amyloid-β (Aβ)-specific antibody.

In HD, the moderate neurodegeneration is observed within the cerebral cortex, where the density of the inclusion bodies is the highest, while massive cell death is occurring with the cells of striatum, showing very modest amount of aggregates[93].

Taken together these results suggest that the formation of the inclusions most probably represents a protective cellular response and indicates that neurons have multiple cellular and molecular ways of coping with misfolded proteins and their aggregates.

1.2.4 Cellular protective methods

The intrinsic propensity of proteins to aggregate has led to the development of multiple defense mechanisms against abnormally folded proteins. The first known protective method involves molecular chaperones, many of which were recognized as heat shock proteins. These molecules act on misfolded proteins by refolding them and thus causing them to be non-toxic. Studies involving cell transfection have shown that molecular chaperones increase solubility of the mutant huntingtin fragment and render it less toxic in vitro as well as in invertebrate models in vivo.

A second defense against nonnatively folded proteins includes proteasome as degradation machinery. Eukaryotic UPS (Ubiquitin-Proteasome System) is known as collaboration between complex 26S Proteasome and the ubiquitin system, which tags proteins destined for the further degradation by proteasome with polyubiquitin chains. Recently a link between impaired UPS and the pathogenesis was also found in Parkinson's disease [94].

A third protective method applies several variations of autophagy, including micro- and macroautophagy and chaperone-regulated autophagy. Chaperone-regulated machinery acts on the proteins containing a consensus motif, recognized by chaperon and co-chaperone complexes, which deliver them to transporters within lysosomal membrane[57, 95]. Autophagy was found to be engaged in the degradation of polyglutamine containing proteins, associated with HD.

1.3 Polyglutamine diseases

1.3.1 Overview

Huntington's disease, repeatedly mentioned with other neurodegenerative diseases in previous sections, belongs to polylglutamine (polyQ) human disorders associated with a triplet CAG expansion mutation and can be included together with Friedreich's ataxia, fragile X syndrome, fragile X tremor/ataxia syndrome, and the polyalanine diseases (with the example of oculopharyngeal muscular dystrophy) into the group of the trinucleotide repeat expansion diseases[96]. The genes associated with these disorders contain the trinucleotide expansions, located the regions coding for genes and therefore translated into a mature protein. There are 10 currently known polyQ disorders (Table 1.2).

Table 1.2. The major characteristics of proteins associated with polyQ diseases and the observed clinical manifestations of these disorders

Disease	Protein	CAG Repeat Length		Clinical features upon expansion
		Normal	Pathological	
Spinobulbar muscular atrophy (SBMA) – Kennedy's disease	AR	9-36	38-62	Motor weakness, swallowing, male breast enlargement, decreased fertility
Huntington's disease (HD)	Huntingtin (htt)	6-34	36-121	Chorea, dystonia, cognitive deficits, psychiatric problems
Spinocerebellar ataxia 1 (SCA1)	Ataxin-1	6-44	39-82	Ataxia, slurred speech, spasms, cognitive impairments
SCA2	Ataxin-2	15-31	36-63	Ataxia, polyneuropathy, decreased reflexes, infancy variant with retinopathy
SCA3 – Machado-Joseph Disease (MJD)	Ataxin-3	12-41	62-84	Ataxia, Parkinsonism, spasms
SCA6	Ataxin-6	4-18	21-33	Ataxia, difficulty in articulating, involuntary eye movements, tremor
SCA7	Ataxin-7	4-35	37-306	Ataxia, blindness, infancy variant with cardiac failure
SCA17	TATA-binding protein	25-42	47-63	Ataxia, cognitive decline, seizures, psychiatric problems
SCA 17	Regulatory subunit of protein phosphatase 2A	7-28	66-78	Action tremor, hyperreflexia, dysarthria, gait ataxia, cognitive and psychiatric disorders
Dentatorubral-pallidoluysian atrophy (DRPLA)	Atrophin-1	6-36	49-84	Ataxia, seizures, choreoathetosis, dementia

Despite intensive expression of the disease-associated proteins in many tissues and organs throughout human body, all polyQ disorders are gradually progressive, fatal diseases affecting Central Nervous System (CNS) and thus causing a large spectrum of neurological and psychiatric symptoms (Table 1.2) with no possible cure. Every disease demonstrates a particular pattern of affected brain regions and cell types of CNS (Table 1.3).

As it was described in the part 1.2, neuropathology of polyQ disorders (as well as other chronic neurodegenerative diseases such as AD and PD), is observed in relatively specific brain regions. The proteins involved in polyQ diseases do not demonstrate any relationship, residing at different intracellular locations; some functions of these proteins are well-known and some are still emerging from multiple studies (Table 1.3). Spinal bulbar muscular atrophy (SBMA) is correlated with an expansion of glutamine residues in the androgen receptor (AR), a nuclear hormone receptor participating in G-protein coupled receptor signalling. The AR regulates target gene expression by binding to DNA and modulating chromatin complexes and proteins regulating them. The functionality of AR is important in many different physiological processes including male reproductive functions [97]. Ataxin-1 is associated with spinocerebellar ataxia 1(SCA-1) is found in the nucleus where it is involved in the processing of RNA and transcriptional repression. The main function of Ataxin-2 has been observed in the metabolism of RNA, the control of apoptosis and actin-polymerisation [98-100]. Ataxin-3 appears to act as a polyubiquitin chain-editing enzyme[101] while mediating a degradation of ubiquitinated proteins [102]. Ataxin-6 is translated into a domain of a voltage-dependent calcium channel [103]. Ataxin-7 belongs to the transcription

mediator complex STAGA [104] and therefore contributes to the transcription regulation. The human TATA-binding protein (TBP) holds 25–42 glutamines within its sequence, which upon expansion transforms into a disease agent causing SCA17 [105]. Atrophin-1 has a function of a transcriptional co-repressor and was found, in Drosophila melanogaster, to be associated with Dentatorubral-Pallidoluysian atrophy (DRPLA) [105]. Another proposed role of atrophin-1 is assembling synaptic signalling complexes through interactions with membrane-associated guanylate kinases [106].

Table 1.3. Potential functions and native cellular location of polyglutamine proteins, causing cellular toxicity in the CNS.

Disease	Protein	Potential functions	Predominant localization of wild-type	Location of cellular toxicity in the CNS
Huntington's Disease (HD)	Huntingtin	Transcriptional regulation, scaffolding protein, trafficking of vesicles, axonal transport, postsynaptic and calcium signaling and bioenergetic homeostasis	Cytoplasmic	Caudate nuclei, striatum, cerebral cortex
Spinobulbar muscular atrophy (SBMA) - Kennedy's	Androgen receptor	Nuclear hormone receptor, alternates gene transciption	Cytoplasmic and nuclear	Spinal anterior horn, facial nucleus, hypoglossal nucleus, skeletal muscle
Dentatorubral-pallidoluysian atrophy (DRPLA)	Atrophin-1	Transcriptional co-repressor, assembling signaling complexes at the synapse	Cytoplasmic and nuclear	Globus pallidus, subthalamic nucleus, dentate nucleus, cerebral and cerebellar white matter
Spinocerebellar ataxia 1 (SCA1)	Ataxin-1	Processing of RNA and transcriptional repression, performs protein folding	Nuclear	Cerebellar Purkinje cells, brainstem, spinal cord, globus pallidus
SCA 2	Ataxin-2	RNA metabolism, the apoptosis regulation and polymerization of actin	Cytoplasmic	Cerebellar Purkinje cells and granular cells, spinal cord, thalamus
SCA3 – Machado-Joseph Disease (MJD)	Ataxin-3	Degradation of ubiquitinated proteins	Cytoplasmic	Cerebellum, substantia nigra, brainstem, dentate nucleus, inferior olive
SCA6	Ataxin-6	Domain of voltage-dependent calcium channel	Trans-membrane	Cerebellar Purkinje cells, dentate nucleus

Continued from Table 1.3

Disease	Protein	Potential functions	Predominant localization of wild-type	Location of cellular toxicity in the CNS
SCA 7	Ataxin-7	Regulation of transcription	Nuclear	Cerebellar Purkinje cells, dentate nucleus, inferior olive, spinal cord, retina, brainstem
SCA 12	Regulatory subunit of protein phophatase 2A	Brain-specific regulation of serine/threonine protein phosphatase 2A	Cytoplasmic and mitochondrial	Cerebral cortex and cerebellum
SCA 17	TATA-binding protein	Transcriptional regulation	Nuclear	Cerebral cortex and cerebellum

1.4 Huntington's Disease

1.4.1 Overview

Huntington's disease is an autosomal-dominant, progressive neurodegenerative disorder with a distinct phenotype, including movement, cognitive and neuropsychiatric difficulties, which lead to the dementia and death approximately 15-20 years after the onset of the disease [107]. Usually, the onset of the symptoms occurs in the middle age, however the signs of the disease can be also observed during any stages of the life: from childhood to senescence [108]. As indicated above, HD is polyglutamine disorder caused by the mutation encoding the CAG repeat expansion in the IT15 gene, which produces a long stretch of glutamines proximal to the amino-terminus of Huntingtin (htt) protein. The expansion of polyQ length beyond 35 residues significantly increases the risk of the disease [109]. Overtime, the result of carrying the HD mutation is expressed in the massive degeneration of the neural tissues characterized by the loss of efferent medium spiny neurons in the basal ganglia, which is a primary source for the common HD symptoms [110]. Recently it was also established that even more widespread degenerations occurs in the brain, involving cortical structures [111]. Huntington's disease has a prevalence of around 5 to 10 cases per 100000 in populations of Western European descent, with many other being at risk of inheriting the disorder. This makes HD the most common inherited neurodegenerative disorder.

1.4.2 Historical Background

HD is also referred as Huntington's chorea. Although the first description of epidemic of dancing mania appeared in 1374, it was not until Paracelsus (1493–1541) who first used the term chorea to describe this movement dysfunction, proposed its origin from CNS [108]. In the following decades, until the 17th century, the disorder stayed obscure and its nature had not been comprehended.

Doctors from England, United States, and Norway performed the first attempts for a medical definition of HD as "chronic hereditary chorea" two centuries later, in the 1840s [108]. However, the first vivid description of the disease appears about 3 decades after, in 1872, by a young American doctor, George Huntington, practicing in Long Island, NY, who composed a short, anecdotal, and entirely unreferenced work called "On Chorea", which was published in the Medical and Surgical Reporter of Philadelphia (volume 26, no. 15, April 13, 1872).

In the early 20th century, the American eugenicist Charles B. Davenport identified a large community with inherited disorders, performing first investigation of families with HD [109]. Later, in the 1950s, Dr. Amerigo Negrette proposed that in a large group of people residing around Lake Maracaibo, Venezuela are widely affected by HD. Twenty years later this geographical location became the epicenter of an impressive effort towards the identification of the HD gene.

Starting from 1979, a team of scientists and physicians performed medical records; collected samples of blood and skin, and graphed the transmission of the disorder within members of families living near of the Lake Maracaibo. In 1983 a group

of scientists led by Joseph Martin, Jim Gusella and Nancy Wexler, utilized for the first time the restriction fragment length polymorphism (RFLP) and linkage analyses to discover a polymorphic DNA marker located on the fourth human chromosome, which was predicted to be linked to HD [112]. Ten years later, a team of geneticists and doctors under the guidance of Gusella and several others, collectively named the Huntington Disease Collaborative Research Group, reported the discovery of the HD gene and mutation associated with the disease [113].

1.4.3. Clinical Findings in HD

Individuals affected by Huntington's disease may start demonstrating symptoms at any time point between the ages of 1 and 90 years; and prior to that, they appear healthy and do not show any visible clinical defects [114]. This physically fit period merges unnoticeably into the prediagnostic phase, when patients show minor changes of personality, cognitive alterations, and slightly decreased control of motor function. Both the healthy and prediagnostic stages are often referred as presymptomatic, however in fact during the prediagnostic phase these manifestations become more evident, even though patients may not be attentive to them [115]. The diagnosis for HD is usually established when symptoms obtain a sufficiently developed and specific level [116]. During the prediagnostic stage, affected individuals might show irritability or be disinhibited and unreliable at work. The members of families notice restlessness or fidgeting, which keeps their partners awake for a long time during the nights. Ultimately, this stage fuses with the diagnostic phase, expressed in the detectable signs

of motor impersistence—the incapacity of keeping a voluntary muscle contraction at a constant level, distinctive chorea, loss of coordination, and retardation of saccadic eye movements [117].

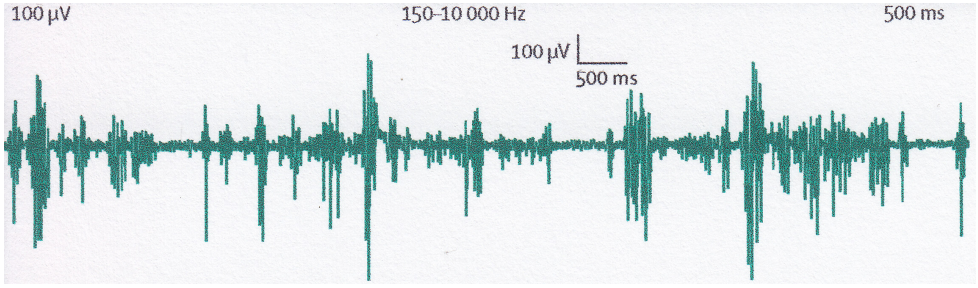
Cognitive impairment in HD, often does not affect long-term memory but weakens executive functions, such as organizing, abstract thinking, or adapting alternative approaches, and impedes the development of new motor skills [118, 119]. These symptoms get worse as the disease progresses; deterioration of speech exceeds comprehension. Unlike cognitive, psychiatric manifestations appear at some frequency but do not follow a continuous progression with disorder severity. Depression is typical consequence, complicated with manic and psychotic symptoms and suicide attempts generally exceed five to ten times that of the general population [118-121].

Chorea, being an intensely used reference point for diagnosis, (figure 1.9,A) represents a poor indicator of disease severity [122]. Patients affected by early-onset Huntington's disease do not necessarily develop chorea, or it might appear only episodically through the illness. Most individuals demonstrate chorea, which initially demonstrates progressive characteristics but later, with the onset of dystonia and rigidity, renders being less prominent [122].

Another manifestation of HD, contributing to patients' overactivity, is motor impersistence (figure 1.9, B) [123]. This inability causes the constant changes in position and often compulsive repositioning. Difficulty maintaining steady pressure during handshake is a feature of Huntington's disease and is called milkmaid's grip. Motor impersistence does not correlate with chorea and usually progresses linearly, which makes it a potential marker for identification of severity of HD [124].

As motor and cognitive dysfunctions obtain more prominent nature, affected individuals eventually die, often from complications of inflammation, injuries due to falls, dysphagia, or aspiration[108]. The standard life expectance from diagnosis to death is 20 years.

A



B

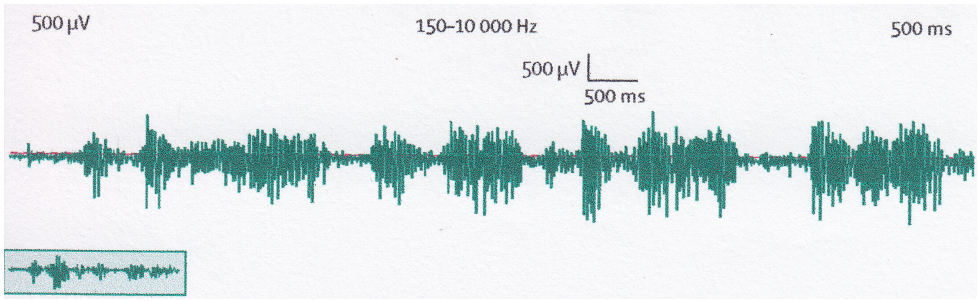


Figure 1.9. EMG recordings of chorea and motor impersistence in HD patients. (Adopted from Walker, 2007, Lancet).

A. Recording of chorea is performed by utilizing a standard belly tendon when surface disc electrodes are fixed over the first dorsal interosseus muscle. The notable irregular pattern of discharges with different amplitude, duration and excitation times of every burst is observed during the period of recording. No EMG activity is observed in healthy resting individuals.

B. B. The patient is asked to abduct the second digit as much as he can and maintain this position. The repeated loss of the motor activity is observed. The inclusion shows the first 400 ms of the chorea recorded at the rest (shown in figure A), which was adjusted for the varying amplitude settings, for comparison. The chorea bursts, on average, exceed the EMG activity from maximum voluntarily effort. Non-affected individuals show the consistent EMG amplitude through the task.

Juvenile form of Huntington's disease, with the usual onset before age 20 and as early as 2 years, often shows symptoms of rigidity without visible signs of chorea[125, 126]. Protracted saccadic eye movements expressed in jerking of the head to glance to the side and seizures are common in young patients and cerebellar dysfunction can potentially appear during this stage of life [126]. A decline in motor skill acquisition and/or school performance represents another set of early signs in children affected with HD[108].

1.4.4. Clinical genetics

The Htt gene is located on the short arm of the fourth chromosome and associated with varying levels of trinucleotide repeat. Normal alleles at this region contain several CAG repeats, however, upon mutation, this stretch exceeds 37 repeats and the disease is expected to be fully penetrant [127-129].

It was shown that some asymptomatic individuals, possessing "intermediate-sized" polyglutamine regions containing from 27 to 35 residues [130, 131] are capable of transmitting the disease to their children, due to the phenomenon known as "genetic anticipation" [132]. The explanation for this phenomenon comes from the fact that the CAG repeats of extended length are not stable and have the propensity to expand further in the next generations especially when the mutated Htt gene is inherited from the father [109]. During mitosis, the possibility of expansion is more common in spermatogenesis, most likely originated from replication slippage, atypical for oogenesis event[133]. Therefore, affected individuals who received the pathological Htt gene from

their fathers may hold a longer CAG repeats stretch and tend to advance in symptoms at earlier stages of life comparing to their fathers [132]. In fact, the amount of present CAG repeats shows an inverse correlation with the age of symptomatic onset and direct connection with the severity of disease (Figure 1.10) [113, 134, 135]**.** Patients with a longer stretch of polyglutamine repeats are predisposed for developing symptoms at earlier stages of life. Particularly extended CAG repeats of 60 or more commonly cause a disease to express itself during childhood or adolescence, referred as a juvenile form of HD. A shorter range of polyQ repeats does not demonstrate such an apparent correlation. Actually, the number of CAG repeats is responsible only for 60% of the variation in the age of onset and severity of the disease, with the rest being represented by environmental conditions and modifying genes [136-139]. The HD patients, despite the similar amount of CAG repeat lengths, may demonstrate a significant difference in the age of onset and major disease manifestations.

Initial experiments revealed that genetic polymorphisms flanking the CAG repeats were able to affect the disease onset [135, 140-143]. Currently, a number of genetic modifiers of Huntington's disease have been identified. All of these modifiers are implicated in a range of mechanisms associated with HD pathology such as protein misfolding, oxidative stress, dopamine toxicity, metabolic deregulation and transcriptional defects. In addition, linkage scans on genome-wide scale (such HD-MAPS study) elucidated possible loci that may include genes that modify HD onset [180-184].

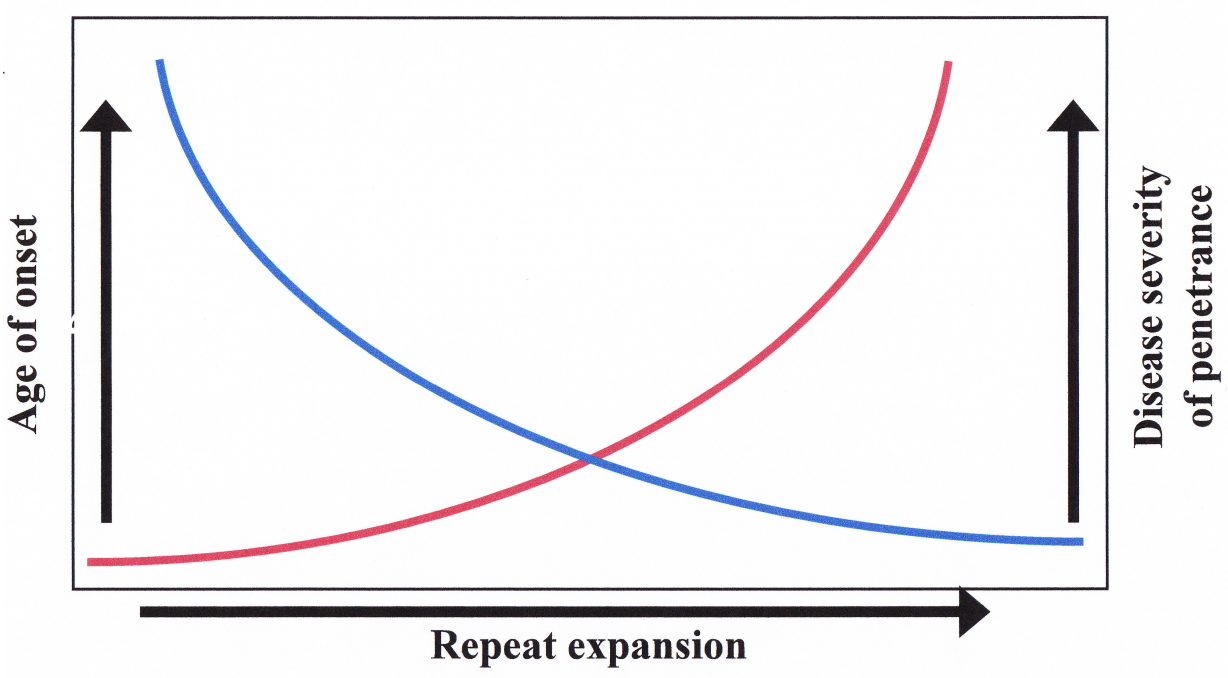


Figure 1.10. Observed correlation between CAG repeat expansion and age of onset and severity of HD. (Adapted from Gusella and MacDonald, Nat Rev Neurosci, 2000)
The longer the polyglutamine repeat the more complete disease severity and penetrance and the earlier the age of onset.

Other polyQ disorders demonstrate related, but somewhat different ranges of CAG repeat length contributing to their onset (Figure 1.11). Additional characteristics of the proteins associated with these diseases, such as the sequences flanking polyQ regions might have a direct affect on aggregation and therefore on the pathology of disease [108].

One of the most recent clinical studies demonstrated that extending CAG repeat size in the normal allele fortifies the association between mutant CAG regions and the severity and progression of HD. In individuals with the lower length of mutant CAG repeat, increasing length of the normal repeat correlates with more acute clinical signs and neuropathology [109]. HD patients with a significantly elongated CAG repeat, rising amount of CAG in the normal allele did not catalyze an increase in the seriousness of the disorder [144]. These findings imply for the first time that the huntingtin allele with normal length of CAG repeats is also able to affect the severity of the disease. One option is that wild-type huntingtin with a long polyQ region could guide to a stronger association with mutant protein fragments, elevating their coaggregation and precluding them from atypically interfering with other proteins [145]. A strong contact between non-pathological and mutant huntingtin could also lead to a substantial decline in normal huntingtin function (see sect. 1.2), causing negative effects.

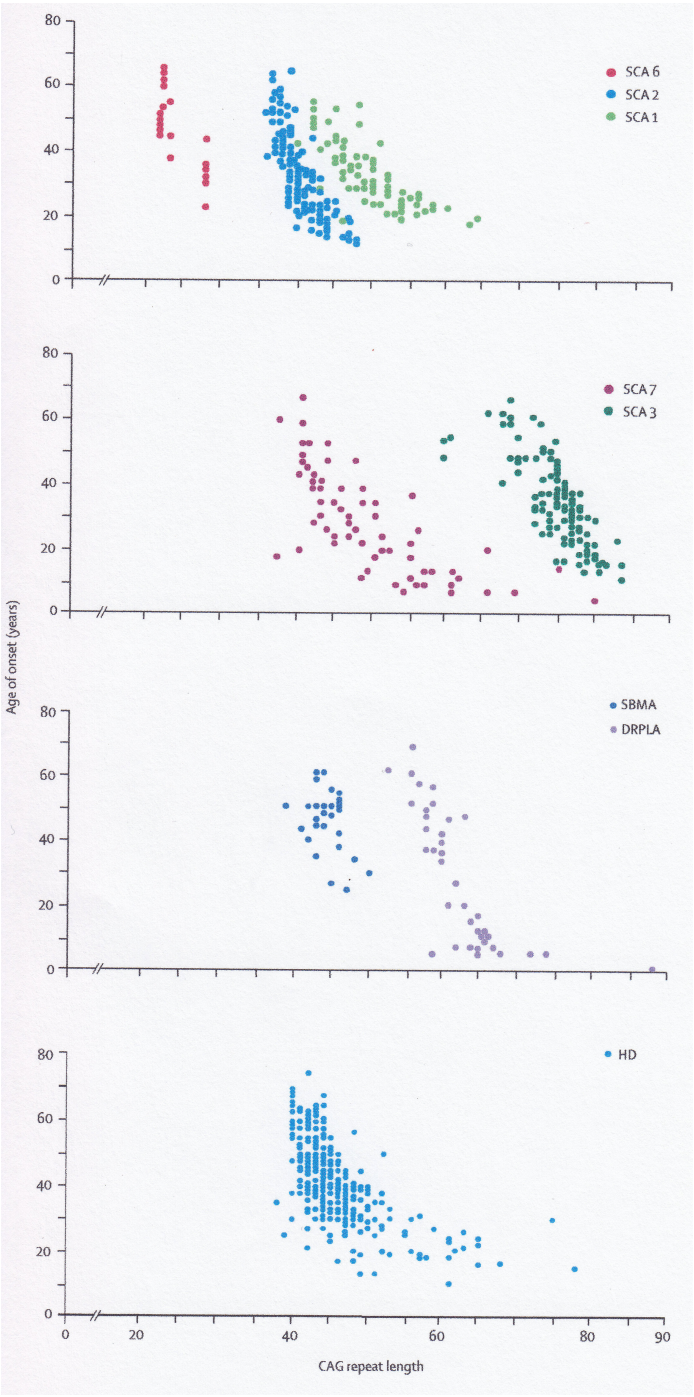


Figure 1.11. Composite graphs illustrating the age of the onset against the number of CAG repeats in 8 polyQ disorders. (Adopted from Walker, The Lancet, 2010)

The strong inverse correlation and the clustering of the number of glutamines are observed in every polyQ disorder.

1.4.5. Neuropathology in Huntington's disease

The Htt gene is universally expressed in many areas of the brain with relative prevalence in cortical pyramidal neurons projecting into striatum but comparatively lower levels in striatal projection neurons [146, 147].

Although some obvious defects are observed in the cerebral cortex and hippocampus, the most distinct neuropathology in Huntington's disease is concentrated in the striatum [110, 148]. The striatum is a major constituent of the basal ganglia, functioning as an interconnected unit of subcortical nuclei associated with the regulation of how to act (and not act) in specific situations [149]. For the most part, the striatum consists of GABAergic projection neurons possessing dendrites that are profoundly covered with spines, accordingly called spiny projection neurons (SPNs): indirect iSPNs, projecting to the external section of the globus pallidus (GPe) and indirectly controling the nuclei that form the border between the basal ganglia and the rest of the brain; and direct dSPNs, projecting to the internal section of the globus pallidus (GPi) and substantia nigra pars.

In general, iSPNs and dSPNs are believed to complement each other in control of movement and action selection. An extended line of experimental evidence suggests that iSPNs facilitate suppression of cortical selection of undesired motions or acts, while dSPNs help to endorse cortical choice of specific actions [149, 150]. As it was described above chorea represents one of the most expressed symptoms during the early phases of HD. The iSPN dysfunction and decline are thought to be in a strong correlation with this observation [110, 151, 152]. A decrease in iSPN GABA release will further disinhibit

autonomously active GPe neurons, consequently leading to reduced basal ganglia inhibition of motor thalamus, hyperkinesia, and disregulated cortical activity. During the later stages of HD, the pathology reaches dSPNs, resulting in struggle to start the movements [148, 151].

Although less remarkable than the degeneration in the striatum, there is clear pathology in other parts of the brain. During the later stages of the disease, neuropathological features have been observed in the following areas: the cerebral cortex, substantia nigra, thalamus, globus pallidus, subthalamic nucleus (STN), hypothalamus, and cerebellum [148, 153, 154].

The development in neuroimaging methods has greatly improved an understanding of HD pathology, elucidating the relations between the morphology of brain changes and the advances in cognitive deficits, such as concentration, short-term memory, and executive functions [155-158]. In 2003, magnetic resonance imaging (MRI)-based morphometric analysis was adopted and verified that individuals affected by HD had considerable volume reductions in nearly every brain region, particularly in the cortex [159]. Further studies also performed by Rosas's group indicated that these alterations happen before detectable symptoms appear [160].

The latest functional MRI (fMRI) studies suggest that functional changes happen during the very early stages of HD, at least close to the same period of time as striatal and cortical atrophy can be noticed. fMRI was shown to be responsive to alteration in striatal function significantly earlier than the appearance of motor symptoms in pre HD, implying that fMRI can be highly potential in identification of early neurodegeneration prior to the onset of clinical indications.

The complete analysis of metabolic imaging data collected during rest has lately been carried out to detect and confirm HD-related spatial covariance patterns (HDRPs) in premanifest individuals [156]. The obtained HDRP depicted decreases in striatal and anterior cingulate metabolic activity, with related metabolic increases in the thalamus, cerebellum nucleus, and in the major motor and visual areas of the cerebral cortex (Figure 1.12).

Longitudinal analysis elucidated a considerable increase in HDRP expression from the start to 18 months and a later turn down from 18 to 44 months. During the study it was determined that HDRP metabolic network and its joints display characteristic temporal profiles (Figure 1.12), which is consistent with the extensive regional variability that has been illustrated in relation to the disease mechanism [158]. Several other lines of evidence support a suggestion that irregular cellular metabolism may be an attribute to the neuropathological progression in Huntington's disease. Nonneuronal tissues show metabolic dysfunctions in HD patients [161, 162]. Mitochondrial toxins are capable of stimulating discriminating striatal neurodegeneration in animal HD models [163]. While the data indicate that the hypermetabolism detected in thalamus and cerebellum is compensated in vivo [164], it is also likely that this event exposes imminent metabolic failure and neural death [165].

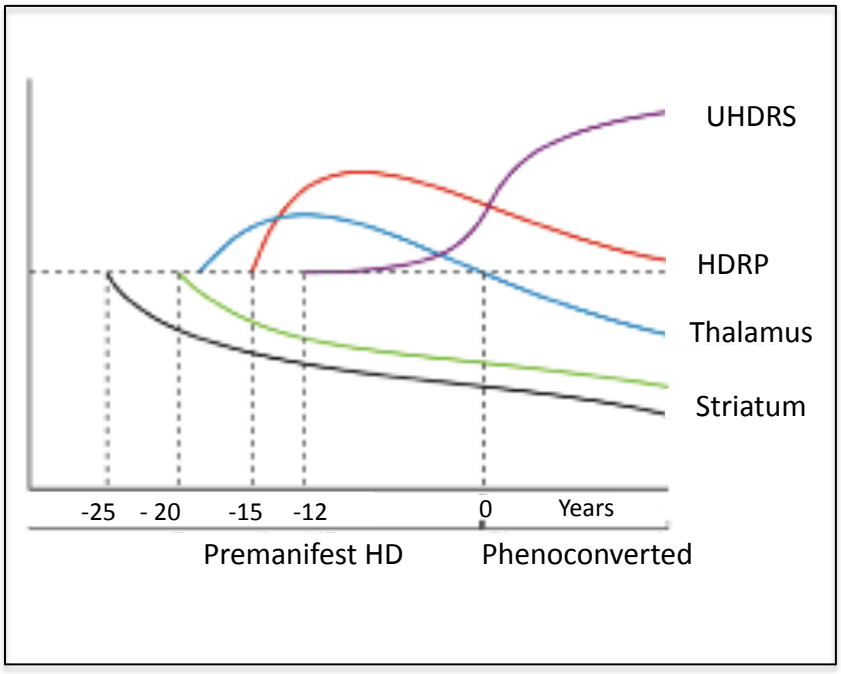


Figure 1.12. Hypothesized time line of alteration in striatal D2 binding, pattern expression and localized metabolism observed in the presymptomatic phase of HD. (Adopted from Eidelberg, J Clin Inves, 2011.)

The individual growth curve model was used for longitudinal imaging data to evaluate the time line of these imaging values in the stage before the established diagnosis. The decrease in the binding of striatal D2 receptor (black) starts around 25 years prior to diagnosis (year 0). Approximately 5 years later, striatal metabolism (green) starts to slow down. The level of thalamic metabolic activity (blue) seems to be increased during the early stages, but regresses to regular level as clinical symptoms (Unified HD Rating Scale [UHDRS]) score, purple) become detectable. HDPR activity is predicted to rise starting around 15 years prior to diagnosis. This network value possibly grows for several years, and then decreases before the onset of clinical signs. The horizontal dash line demonstrates the normal average values.

1.4.6. Modeling HD

Besides the presently viable methods involving post mortem human brain specimen, research on HD mainly relies on animal and cellular models. These models have been effectively utilized to study pathological mechanisms, molecular objects, and therapeutics (Table 1.4).

1. Chemical models

Prior to the discovery of the disease gene, HD animal models were constructed by injecting neurotoxins into the striatal region of the brain. The early publications reporting that intrastriatal injection of kainite, a non-NMDA glutamate agonist, in a direct manner could imitate in rats the axon-sparing striatal lesion detected in the human disease, are the opening point of a extensive literature on the utilization of glutamate analogs to create striatal specific neural degeneration in rodents [166]. Quinolinic acid and kainic acid are the primary agents, which were utilized to generate rodent and non-human primate models of HD, indicating that excitotoxicity could be an important factor in the cell death associated with the disease [167, 168]. These chemical models were also beneficial due to their ability to reproduce the regional specificity of neuropathology in Huntigton's disease. Nevertheless, they are not capable of replicating the pathological pathways stimulated by the mutant gene.

Table 1.4 Rodent genetic models designed to study Huntington's disease

	R6/1 (Mangiarini et al)	**R6/2 (Mangiarini et al)**	**N171-82Q (Schilling et al)**	**YAC128 (Slow et al)**
General feature				
Animal	Transgenic mouse	Transgenic mouse	Transgenic mouse	Transgenic mouse
Construct	1.9-kb fragment from the 5' of human Htt	1.9-kb fragment from the 5' of human Htt	First 171 amino acids of human Htt	Yeast artificial chromosome expressing full-length human Htt gene
Promoter	Human Htt	Human Htt	Human Htt	Human Htt
CAG	113	144	82	128
Onset of symptoms	15–21 wk	5-6 wk	10 wk	8-12 wk
Survival	32–40 wk	12–15 wk	10-24 wk	Normal life span
NII/cell pathology	NIIs and neuropil aggregates throughout the brain	NIIs and neuropil aggregates throughout the brain	NIIs in cortex, hippocampus, amygdala, and striatum	EM48 positive inclusions in striatal cells, no NIIs detected
Brain atrophy and cell loss	Overall brain atrophy, reduced brain volume No evidence of massive cell death	Overall brain atrophy No evidence of massive cell death	Overall brain atrophy	Reduced striatal and cortical volume at 48 wk

Continued from Table 1.4

	R6/1 (Mangiarini et al)	**R6/2 (Mangiarini et al)**	**N17-82Q (Schilling et al)**	**YAC128 (Slow et al)**
Symptoms				
Motor	Clasping behavior (onset 20 wk) Rotarod deficit Gait abnormalities	Clasping behavior (onset 8 wk) Rotarod deficit (5 wk) Hypercinetic movements Resting tremors Circling behavior Increase in limb movements Decrease in grip strength	Clasping behavior (onset 15 wk) Rotarod deficit (15 wk) Hypokinesis Resting tremors Tremors and gait Abnormalities Loss of coordination Muscle weakness	Clasping behavior Rotarod deficit (24 wk) Hyperkinesis (12 wk) and Hypokinesis (24 wk) Gait abnormalities Circling behavior Ataxia
Cognitive	Decrease anxiety	Rigidity in cognitive process Increased exploratory behavior at 4 wk that declines and ends by 8.5 wk	Deficit in working memory Rigidity in cognitive process	Depressive behavior (12 wk)
Others	Progressive weight loss	Progressive weight loss Seizures, diabetes, and cardiac dysfunction	Progressive weight loss	Weight increase

Continued from Table 1.4

	BACHD (Gray et al)	**Transgenic HD rat (Von Horsten et al)**	**HdhQ92-111 (Wheeler et al)**	**HdhQ140 (Menalled et al)**
General feature				
Animal	Transgenic mouse	Transgenic rat	Knock-in mouse	Knock-in mouse
Construct	BAC expressing full-length human Htt gene	1,962-bp rat Htt fragment	Replacing exon 1 of mouse Htt with a mutant human exon 1	polyQ sequence inserted into the endogenous mouse Htt gene
Promoter	Human Htt	Native rat Htt promoter	Mouse Htt promoter	Mouse Htt promotor
CAG	97	51	92-111	140
Onset of symptoms	12 wk (robust by 6 mo of age)	40–50 wk	96 wk	12 mo
Survival	Normal life span	98 wk	Normal life span	Normal life span
NII/cell pathology	mHtt I in neuropil and in cortex and striatum (48-72 wk)	Neuropil aggregates and NI in striatum and less extent in cortex (72 wk)	NIIs and striatal neuropil aggregates (68 wk)	NII and neuropil IB in striatum, cortex, hippo-campus, (16–24 wk)
Brain atrophy and cell loss	Brain atrophy Cortical and striatal volume	Enlarged lateral ventricles Focal lesions in the striatum	No brain atrophy observed No cell loss observed	No brain atrophy observed No cell loss observed

Continued from Table 1.4

	BACHD (Gray et al)	**Transgenic HD rat (Von Horsten et al)**	**HdhQ92-111 (Wheeler et al)**	**HdhQ140 (Menalled et al)**
Symptoms				
Motor	Clasping behavior not reported Rotarod deficit (by 8 wk and progressed by 24 wk)	Progressive impairment of coordination and balance Dyskinesias of the head Gait abnormalities	No clasping behavior No rotarod deficit	Decrease in locomotor activity Hyperactivity and hypoactivity Gait abnormalities
Cognitive		Reduced anxiety-like behavior, emotional and cognitive decline	No symptoms	No symptoms
Others	Weight increase	Weight loss	No abnormal weight loss	No abnormal weight loss

2. Genetic models

HD cell lines are especially useful in the fact that they permit the constant or inducible expression of wild-type or pathological huntingtin, and are essential for the analysis of disease pathways, as well as screening of therapeutics [169-171].

A large number of species, such as the invertebrate *Caenorhabditis elegans* and *Drosophila melanogaster*, nonmammalian *Danio rerio* and some mammals, including rodents, have also been genetically modified to carry a mutation in HD gene. A wide variety of Huntingtin's mouse models demonstrating distinctive levels of resemblance to the human state have been engineered. Models containing either truncated or full-length human or mouse mHtt show substantial phenotypic variations, which may be the consequence of the different protein context, particular mouse strain, or presence of regulatory regions linking the mouse and human huntingtin genes. In order to find a solution to these drawbacks, a number of researchers are proposing also the production of bigger HD models such as minipig, sheep, and the non-human primate.

A). Transgenic mouse. In a breakthrough study, Gillian Bates' group [172, 173] produced the first line of transgenic mouse by introducing a 1.9-kb fragment with the human htt promoter and exon 1 of the human HD gene containing 144 CAGs. This type of mice, referred as the R6/2 mouse line, demonstrate both initial and later acute cognitive, motor and anatomical symptoms [172, 173] (Table 1.4.) Because of the large number of polyQ repeats, the R6 lines are more suitable for representing the juvenile rather than the adult phenotype of HD individual.

The next transgenic line, N171–82Q, was generated by David Borchelt's group and demonstrated a comparable neuropathological and symptomatic phenotype. In contrast to the R6 line, the N171–82Q mice contains fewer polyQ repeats yielding a delayed start of symptoms, which makes it an appealing system for the investigation of premanifest therapies, also permitting an extended experimental window for time-resolved studies.

Transgenic mice carrying full-length htt protein were shown to be more successful than N-terminal exon-1 models due to the neuronal loss and ability to mimic more precisely the succession of events causing HD. Currently, four full-length htt mice lines have been generated (Table 1.4): the first model containing entire length of IT15 gene cDNA clone with an extended CAG repeat tract controlled by the cytomegalovirus (CMV) promoter [174]. The second model is yeast artificial chromosome (YAC) mouse carrying a full-length genomic Huntingtin gene construct [247]. The third model is a mice line, nick-named "Short-Stop", containing a mutant huntingtin gene truncated after intron II. In contrast to YAC128, these mice showed no detectable signs of neuronal abnormalities when evaluated by neural tissue weight, striatal and neuronal count in the striatum in spite of the occurrence of aggregates. This observation proposes that inclusion bodies are not necessarily toxic in vivo and that it is more likely that soluble oligomers of mutant huntingtin may be pathogenic to neural cells. In 2008, Yang's team created the fourth full-length mouse model expressing a bacterial artificial chromosome (BAC) with an entire human HD gene containing 97 CAGs [175]. These mice display advancing motor defects and specific

late pathology in cortex and striatum regions, therefore symbolizing an innovative and vigorous in vivo model for HD studies.

B). Knock-out mice. During the process of generating a classic HD transgenic model, the exogenous Htt gene incorporates randomly, which creates a possibility of unwanted intervention of the activity of other unrelated to the disease genes. In addition, the transgene expression controlled by non-native promoters may yield a phenotype that distantly imitates the disorder, as a result of the expression of Htt gene at concentrations significantly exceeding physiological levels. Genetically accurate models that possess the mutation in the proper genomic and protein framework and at a low concentration have been engineered to more consistently reproduce the pathogenesis of Huntington's disease (Table 1.4).

Knock-in mice have been created by inserting extended polyQ regions into the native mouse Htt gene (Hdh) found in chromosome 5 [176, 177] and/or by substituting mouse htt exon 1 with human one expressing pathological CAG length[178-180].

C). Rat models of Huntington's Disease. Research conducted on rats have the obvious benefits arising from the accessibility of a larger amount of behavioral and imaging experiments that are appropriate for description of neurological defects emerging in many neurodegenerative diseases. As a result, rats are considered to be the best model for the testing of novel therapeutics in longitudinal preclinical studies. The first HD rat models carried mutant huntingtin gene with 44, 66, and 82 CAGs, controlled by either the phosphoglycerate kinase 1 (PGK) or the CMV promoters [181].

This strategy was shown to yield a strong accurate in vivo model for specific degeneration of neural tissue.

D) Large-animal models of HD. Non-human primates represent a great potential for the investigation of human neurological disorders, the area, where the majority of the current experimental models remains to be imperfect. Anthony W. S. Chan's research team [182] created the first huntingtin transgenic rhesus macaques, proving the possibilty of engineering useful transgenic monkey models of HD. This research group created this primate model by inserting changed forms of the huntingtin gene into monkey eggs. The transgenic monkeys demonstrate essential clinical characteristics of HD, such as dystonia and chorea [182]. Other large-animal models, progressively acquiring importance, are miniature pigs and sheep. Their increased size, organ capacity, and physiology make them particularly suitable to resemble of humans, making them appropriate for preclinical studies and safety controls.

E). Non-mammalian models of HD. Genetic operations in the mice, rats, sheep, and monkeys are expensive and time-taking procedures. Besides that, ethical problems commonly appear with the studies involving these larger animal models. This triggered a requirement for more trivial, faster (in terms of reproduction), and more affordable models.

The first model worth mentioning is yeast. Even though yeast represents a unicellular organism, the straightforwardness and genetic pliability of the system offered the mechanistic details [183, 184] along with an essential scientific base for drug discovery. HD is widely investigated using *C. elegans* as a model organism; where

neurodegeneration was shown to be a function of both age and length of polyglutamine repeat [185, 186]. The fly, especially *Drosophila melanogaster*, is an ideal option for modeling HD due to its entirely functional nervous system. Furthermore, in *Drosophila*, foreign genes can be created and expressed in particular tissues and point of time. HD has also been modeled in *Danio rerio* (zebrafish) which possesses some obvious benefits for modeling polyglutamine disorders: they demonstrate a rapid external growth, reproduce fast and in large quantities, and are transparent, allowing a straight examination of fluorescently tagged proteins, tissues and entire organs[109].

1.5 Huntingtin Protein

1.5.1 Structure

As huntingtin has an impressive size of 348-kDa and highly flexible structure, it hampers the growth of crystals to reveal its conformation in an unbound state. As a result, 17 years following the discovery IT15 gene, there is no clear understanding on the structure of its translated protein.

The famous signature of the huntingtin is the polyglutamine stretch at its NH2 end. Htt also contains multiple consensus regions, such as huntingtin, elongation factor 3, protein phosphatase 2A, and TOR 1 (HEAT) repeats that represent domains essential for protein-protein associations[187]. This protein undergoes multiple

posttranslational modifications and is enriched in various consensus sites for proteolytic enzymes (Figure 1.13).

PolyQ region

The polyglutamine region in human htt starts at the 18th amino acid and, in the case of a wild-type, includes up to 35 glutamines [113]. The initial htt multiple sequence alignment demonstrated that the polyQ stretch represents an early acquirement of the protein [188]. Currently it is known that the polyglutamine region is a major regulator of htt association with its partners and that huntingtin protein is capable of binding to a large variety of biomolecules [189]. Htt protein possesses flexible or multifunctional domains, which allow to obtain particular conformations and activities as required by its partners and cellular localization. In order to elucidate the role of the polyQ stretch in normal huntingtin function, a particular mouse line [190] with a deletion of the small CAGs repeat coding for 7 glutamine residues in the mouse HD gene was created. These animals were born showing a standard Mendelian frequency and demonstrated only slight deficits, i.e., deficiency in learning and memorizing abilities. It was proposed that the polyQ stretch is not necessary for any important function of htt but alternatively may perhaps adjust a native function of the protein. Experiments on Hdh fibroblasts, performed *in vivo*, implied that the polyQ is involved in altering durability and energy level of the cells.

PolyPro region

In higher organisms, and particularly in mammals, the polyQ tract is followed

by a polyproline (polyP) region, which is believed to be a new and random acquirement in huntingtin evolution. It was proposed that the polyP role may be expressed in the stabilization of the polyQ repeat by maintaining it soluble, and interestingly, it has co-occurred with longer polyQ stretches [191].

The HEAT repeats

The so-called HEAT repeats are located downstream of the polyQ region. They represent 40-amino acid-long stretches that appear several times within a given protein and are known to participate in protein-protein interactions [192]. The presence of these repeats implies that huntingtin may execute its physiological function by utilizing various protein partners.

1.5.2. Huntingtin Cleavage

In 1996, one of the major discoveries in HD was performed by Michael Hayden's laboratory, when htt was shown to undergo proteolytic cleavage by caspases [193]. Upon the observation that neurotoxicity was elevated in cells containing truncated huntingtin, the elucidation of htt as a caspase substrate led to the development of a theory correlating the cleavage of huntingtin by caspase to the further intensification of caspase activity in a malicious cycle.

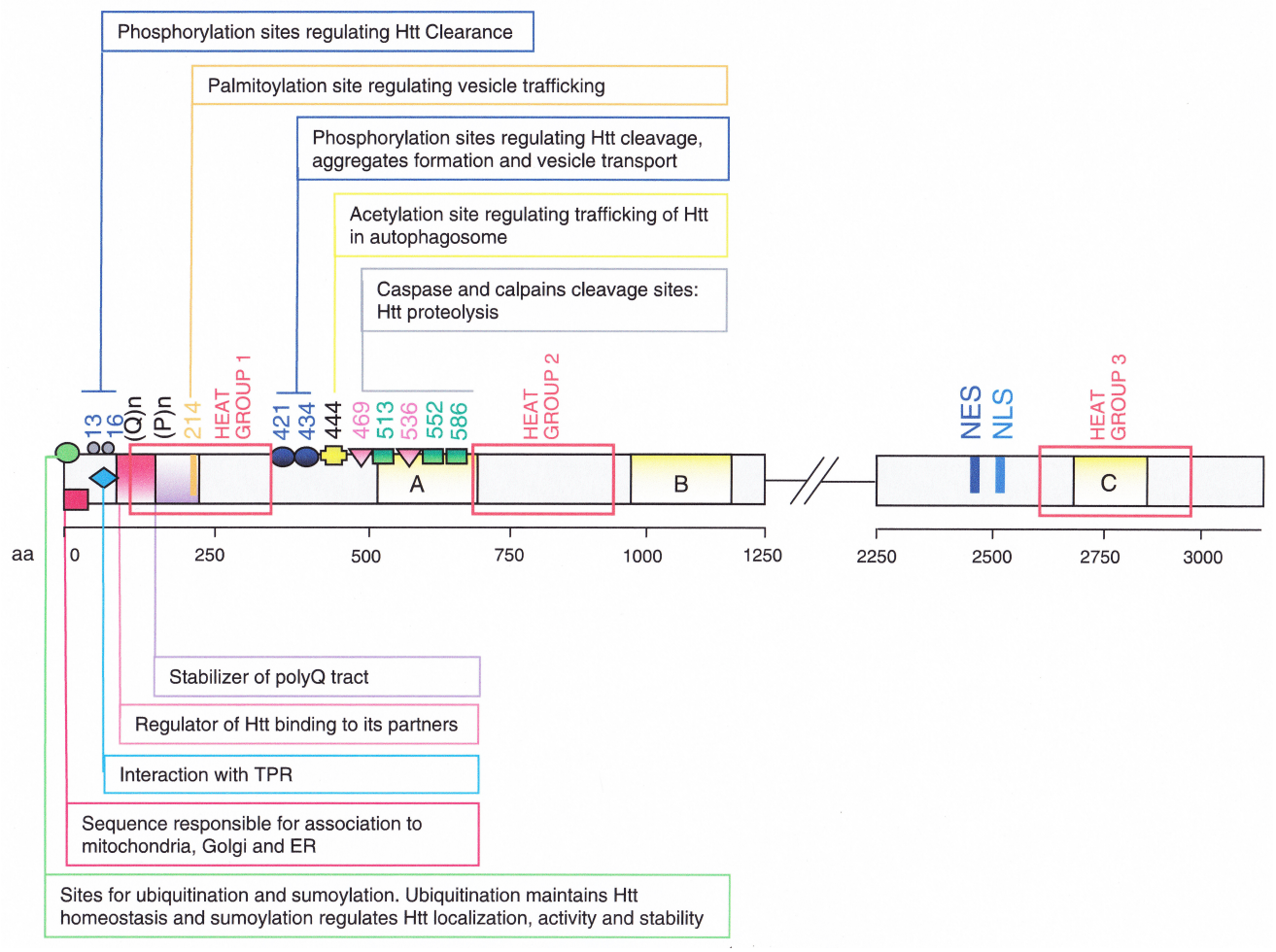


Figure 1.13. **Schematic illustration of the huntingtin amino acid sequence**. (Adapted from Zuccato, Physiol Rev, 2010)

(Q)n describes the polyglutamine stretch followed by polyproline region (P)n. The red rectangles represent three major HEAT repeats. The smaller green rectangles show the caspase cleavage sites and the according amino acid position, whereas pink rectangles demonstrate calpain cleavage sites and their amino acid numbers. Yellow large boxes: regions specifically cleaved in cerebral cortex.

This theory, referred as the toxic fragment hypothesis, states that proteolytic cleavage of huntingtin releases toxic constructs including the elongated polyglutamines region and that their buildup may cause the recruitment of even more proteolytic caspases and the ultimate death of a cell. The further finding from this group points toward caspase-6, as a major executor of a critical proteolytic cleavage in HD. In fact, the prevention of caspase-6-performed htt cleavage decreases the disease symptoms in HD mouse model [194].

Besides caspases [193, 195, 196], calpains were shown to also truncate huntingtin [197-199]. Htt protein is also vulnerable to aspartyl proteases, which similar to caspases and calpains, cut the protein proximal to the NH2-terminal region [200]. The truncation takes place both in native and mutant huntingtin, but the latter is more prone to cleavage and contributes NH2-terminal sections that are observed in the cytoplasm and nucleus of different types of cells [85, 172, 196, 200, 201].

Alternatively, one study performed on human post mortem specimen stated that mutant htt is actually more resistant to truncation than native protein [202]. It was proposed that, besides proteolytic enzymes, other aspects could influence the cleavage of htt. One of the proposed regulators of htt truncation is posttranslational modification.

1.5.3.Posttranslational modification of htt protein

The early research demonstrated that htt is ubiquitinated at lysines K6, K9,

and K15 and directed to the proteasome [85, 203]. Ubiquitination levels significantly contribute to sustaining of huntingtin homeostasis in the organism through regulation of its structural permanence, function, and localization within a cell. When the polyglutamine region reaches pathological length, this course of action is interrupted leading to proteosomal malfunctioning and accretion of truncated huntingtin that renders toxicity to neuronal cells.

Several studies demonstrated that htt can undergo phosphorylation on Ser-421 performed by protein kinase B or Akt [204-206] and at Ser-434, Ser-1181, and Ser-1201 by the cyclin- dependent kinase 5, and this phosphorylation limits caspase-regulated huntingtin truncation at residue 513 and decreases the formation of aggregates and neurotoxicity [207]. Recently, the phosphorylation sites of entire htt protein were localized by mass spectrometry, yielding six new serine phosphorylation positions. The mutation on Ser-536, belonging to one of these target proteolytic regions, prevented cleavage executed by calpains and decreased the toxic impact produced by mutant huntingtin [208]. Phosphorylation of particular residues appears to introduce neuroprotective nature to native huntingtin and is involved in neuronal vesicles' transport, regulated by htt. The recent findings showed that the IkappaB kinase (IKK), earlier indicated to be associated with htt [209], is capable of phosphorylation of Ser-13 and may trigger phosphorylation of Ser-16. Post-translational modification of these amino acids advances alteration of the neighboring lysines and labels native htt for proteasomal and lysosomal clearance [210]. It was also demonstrated that the occurrence of a polyQ elongated region limits the frequency of this

phosphorylation and degradation of mutant htt protein.

Sumoylation on the first 17 N-terminal amino acids of native htt regulates its localization within a cell, level of activity, and resistance against degradataion [191, 211] (Figure 1.13). Another kind of post-translation modification affecting htt is palmitoylation at Cys-214 by huntingtin- interacting protein 14 [212]. The fact that htt represents a target for palmitoylation is consistent with its suggested involvement in controlling vesicular trafficking, as the palmitoylated proteins frequently play role in the active organization of the machinery that regulates the transport of vesicles and vesicle function at neuronal synapses. Elongation of the polyQ region in htt protein causes diminished palmitoylation, providing additional input to the elevated neurotoxicity [212].

1.5.4. Huntingtin Functions

Wild-type huntingtin protein is engaged in the range of arch-important cellular processes and mechanisms. Htt was found to be important in embryonic development since its absolute deletion in HD knock-out mice results in the death of embryos during the period preceding gastrulation and the development of the CNS [213, 214]. Further findings imply that htt is necessary at various stages of embryonic growth and that its compete inactivation or reduced functional activity produces an exceptionally premature phenotype in mouse models [215].

Several *in vitro* and *in vivo* studies revealed another important cellular function of huntingtin -being an anti-apoptotic factor. Experiments on

immortalized cell lines showed that overexpression of normal htt in neural cells shielded them from toxic stimuli [216, 217]. The later research supported these findings and demonstrated that native htt can prevent neuroblastoma and kidney cells loss prompted by the pathological form of protein [218].

An extra evidence for an antiapoptotic function of huntingtin arose from studies on YAC18 transgenic mice contianing full-length normal htt. It was observed that the primary striatal neurons in this mouse line were confined from apoptosis in contrast to YAC72 model with a mutant huntingtin [218]. One of the proposed mechanism on htt prosurvival activity states that the normal htt interferes with apoptosis by directly binding to active caspase-3 and blocking its proteolytic action [219].

Huntingtin is also believed to play an active role in axonal and vesicle transport processes. Htt is localized mostly in the cytoplasm of neurons and present in elevated numbers in compartments enclosing vesicle-regulating proteins [170]. In *Drosophila*, it was found that the decline in htt expression interrupts axonal transport [220]. The native huntingtin also participates in rapid mitochondrial axonal trafficking in mammalian neural cells [221]. Several publications specify that normal htt mediates axonal transport by contributing to the association of the motor machinery on microtubules [222].

Active messaging between neurons is controlled by a series of proteins located in the synapse [109]. Wild-type huntingtin is in constant interaction with the proteins associated with cytoskeleton and synaptic vesicles, which are important in both exo- and endocytosis at the neuronal terminals. This allows htt

to be involved in the regulation of synaptic action in neurons [223]. In addition, htt is capable of binding molecules playing a major role in transmission through neuronal synapses, such as endophilin A and syntaxin. The strength of the binding correlates well with the length of the polyglutamine region and is boosted by the occurrence of an elongated CAG repeat, causing interruption of synaptic communication in Huntington's disease [223].

1.5.5. Studies on PolyQ aggregation in vitro

As it was mentioned in the previous sections, in HD, a polyQ tract containing 36 residues in the Htt exon1 fragment is benign, while the addition of only 2–3 extra glutamines is correlated with disease risk. Upon reaching this threshold and above it, longer polyQ stretches are linked with earlier ages of symptomatic onset. A series of simple (naked) polyQ-containing synthetic peptides were utilized in the initial attempts to investigate the biophysical basis of polyQ repeat length dependence, by studying their aggregation kinetics in vitro [224]. Chen et al. found that that $K_2Q_nK_2$ peptides, containing 28, 36 and 47 glutamines go through a random coil to β-sheet transition with kinetics related to their aggregation kinetics, implying the absence of off-pathway species in their aggregation pathway. Detailed studies on the time dependence of aggregate growth verified that polyQ aggregation follows nucleated polymerization assumption.

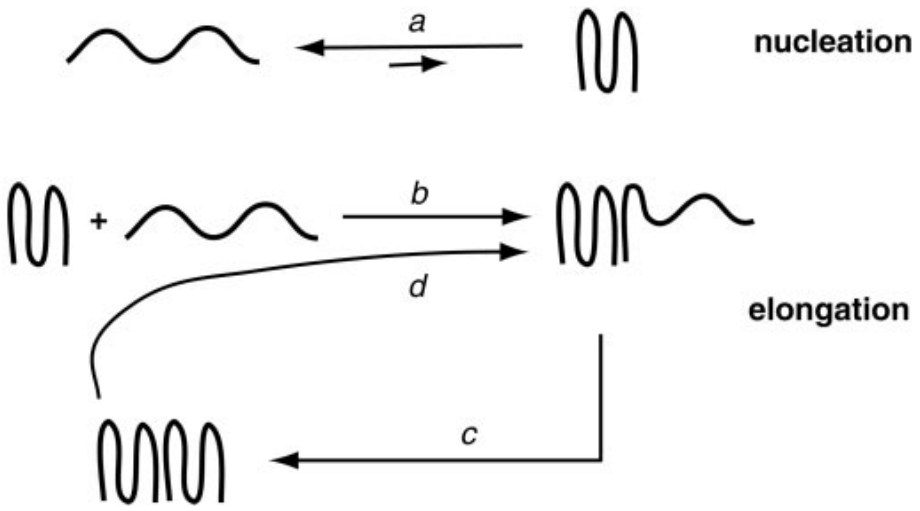


Figure 1.14 **A model for simple polyQ aggregate nucleation and elongation.** (Adapted from Chen, Proc. Natl. Acad. Sci., 2002)

Step a: nucleation is occurring through an unfavorable transition from an extended random coil conformation to a compact conformation representing the aggregation nucleus. The elongation part includes an initial binding (described in b) of the nucleus to an extended state monomer and following consolidation of structure (shown in c) that creates a novel binding site for the next monomer. The consequential dimer binds another extended monomer (d) to carry on the process.

However, comparing to other models of nucleated protein polymerization, it was suggested that the aggregation of all simple polyQ peptides occurs through monomeric nucleus (Figure 1.14). It was also shown that in the case of simple polyQ system, higher aggregation rates for peptides containing longer polyQ regions are linked to more favorable equilibrium constants for nucleation process.

The finding that the proteolytic cleavage of an entire htt exon1 is required for neural toxicity gave a rise to an idea of an important role that polyQ flanking sequences have in htt aggregation. It was demonstrated that the proline-rich region (polyPro) located on the C- terminal side of the polyQ in exon 1 decreases aggregation kinetics and stability of aggregates but does not have any effect on the aggregation mechanism [225].

The role of first 17-amino-acid located on the N-terminus of Htt exon1, was studied using $HTT_{NT}Q_nP_6$ (abbreviation used in the study) peptides. Wetzel' team has shown that addition of flanking sequence of 17 N-terminal residues to the polyQ sequence determines a complex alternative aggregation mechanism of the peptides [226]. According to the proposed model, the N-terminal 17-amino-acid region has a propensity to obtain an aggregation-challenging compact coil conformation, but upon its addition to polyQ sequence, this segment obtains a more extended structure. In its extended state, it tends to aggregate into a semi-stable, micelle-like species, with the N-terminal 17-residues segment forming the freely packed core and the adjusting polyQ sequences located externally. It was proposed that this initial step of aggregation reaction is non-nucleated, but the consecutive step gives a rise to amyloid nuclei formed from these initial, micellular

aggregates. The resulting nucleus and subsequent fibrillar aggregates were suggested to have a core composed of both N-terminal 17-residues and polyQ segments, with former located on the internal side of the aggregates. After amyloid fibrillar aggregates have been formed, the remaining amount of monomer (the majority of the molecules present during the nucleation stage) supplies the progression of the fibril-elongation process. Surprisingly, the later model for Htt exon1 fibrils formed from $htt^{NT}Q_{30}P_{10}K_2$ fragment proposed by Wetzel and others suggest that N-terminal 17-residues fragment remains water-exposed and flexible in its α-helical conformation, comparing to highly rigid and dehydrated core formed by the polyQ sequence [227]. The potential way of resolving distribution of the flanking htt region within the structure of the mature fibrils is presented in chapter 4.

Another revision recently proposed by Wetzel concerns the size of the nucleus formed during aggregation of simple polyQ peptides. It was shown that all these peptides still aggregate through nucleated growth polymerization, however, with the decrease in the amount of glutamine residues from Q26 to Q23, the oligomeric state of the critical nucleus for aggregation changes from monomeric to dimeric to tetrameric. This difference in nucleus size proposes a general duplex antiparallel β-sheet scaffold for the nucleus and furthers supports a hypothesis that aggregation of the peptides with longer stretch of polyQ repeat follows nucleation mechanism with a nucleus = 1. Kar et al suggests that the found variation in the size of the critical nucleus might have a reflection in the toxicity of the longer polyQ repeat length fragments, though the exact reasoning behind this

assumption still requires clarification.

In the current work we investigate the aggregation pathways of htt exon1 peptides containing 42 and 22 glutamine residues. Our findings agree with proposed non-nucleated initial step for aggregation of htt exon1 peptides, containing longer stretch of polyQ ($NtQ_{42}P_{10}K_2$), as proposed for $HTT_{NT}Q_{30}P_6K_2$ fragment [226]. However, aggregation of the shorter $NtQ_{22}P_{10}K_2$ follows an alternative, nucleation-dependent mechanism with a critical 8-meric nucleus. These results are presented and discussed in detail in chapter 3 of this document.

CHAPTER 2

The major methodology: Small Angle Neutron Scattering

2.1 SANS as a powerful tool to study protein aggregation and amyloid formation

2.1.1 Overview

Current challenges in neurodegenerative disease research are to map out the polyGln aggregation pathway by identifying the various precursor structures and establishing their pathological role. As of now the causal link between the particular structures of huntingtin protein aggregates and their toxicity is unclear. Assuming that an Htt aggregates through a specific pathway, and not a funnel system, identifying and characterizing the structural details of intermediate species is the essential first step, potentially pointing the direction for designing therapeutic agents that alter this aggregation pathway [228]. However, there are some major concerns that need to be taken into consideration while pursuing this goal: aggregate structures change over time and exist as an ensemble, having polydispersity in both size and shape distribution. This imposes limitations on many techniques for their appropriate study.

Despite an increasing amount of evidence supporting the toxic nature of several intermediates, such as Aβ and α-synuclein more information is still required on their structure and place in the aggregation pathway. Potentially, information on the intermediate species can be obtained by using a technique utilizing high hydrostatic pressure (HHP) to follow the protein aggregation [229]. It was stated that HHP allows stabilization and isolation of the formed intermediates [229]. Time-resolved methods for monitoring the evolution of the various intermediate species would be an another alternative for providing critical insight on the mechanism of protein aggregation [230, 231]. Small-angle neutron scattering (SANS) is particularly well suited for structural characterization of protein aggregates and is capable of combining solution environments with time-resolution. Thus SANS is potentially a powerful method to gain new insights into the aggregation process.

Previous applications have shown SANS to be promising for studying fibrils of amyloid peptide Aβ [232, 233] and α-chymotrypsin [234]. Small angle experiments were applied to study the fibrillization pathway of insulin and revealed the three major participants in this process: monomers, final fibrils and intermediate oligomers [235]. Through identification of these components as well as a description of their low-resolution three-dimensional structures, the elongation pathway of insulin amyloid fibrils has been proposed (Figure 2.1).

We have used SANS to study the process of polyGln aggregation and to identify early oligomeric intermediate structures. In this work I describe how we optimize SANS technique and data analysis to obtain a unique insight into polyGln

self-assembly as it relates to protein misfolding and disease, and specifically to obtain the knowledge about the size and shape of the small mutant htt-like intermediates and the internal composition of mature fibrils.

Small angle scattering techniques complement classically used methods for studying aggregation mechanisms of amyloids [235]. As it was mentioned in the previous sections, one of the existing hypotheses on the formation of amyloid fibrils suggests that it follows a nucleation- dependent process [236-238]. In general, fibrillization starts with monomeric species, which proceeds to form a nucleus (in monomeric or oligomeric form) [239] prior to elongation of protofilaments. In this context, SANS for example, can probe the oligomeric state of the nucleation complex. A deeper understanding on the structure and mechanism behind the association of amyloid fibrils in general, and on the early stages of nucleus formation in particular, is crucial for the elucidation of the progression of amyloidosis. The knowledge on the structure of nucleus will provide an insight on the intermolecular forces driving the formation of the amyloid at the very initial stages (nucleation) as well as later intramolecular associations, taking place during fibrillization, with a nucleus functioning as a scaffold.

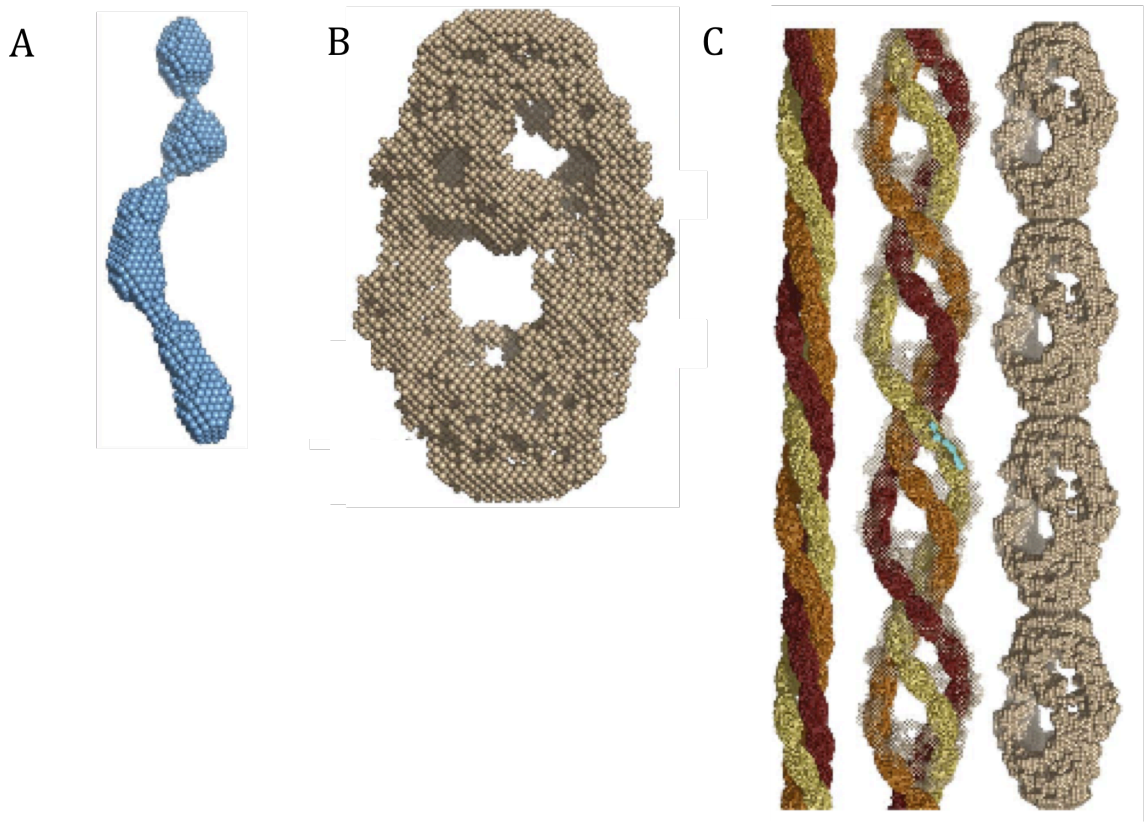


Figure 2.1. The association pathway of insulin fibrils. (Vestergaard et al., PLoS Bio, 2007).

A. The ab initio shape of insulin oligomer containing 5-6 monomers with a maximum dimension of 200 Å. B. The repeating helical subunit of the insulin fibrils calculated using ab initio algorithm. This subunit demonstrates a 690 x 390 x 160 $A^{\circ 3}$ overall dimensions and contains roughly 200 oligomeric species depicted in A. C. Three protofibrils (depicted in orange, yellow, and red) interweave to make up the mature insulin fibril. Four repeats of the ab initio reconstructed SAXS model of the coherently scattering subunit of the fibril (as depicted in B) are shown to form the final structure of these amyloid fibrils.

2.1.2 Investigation of protein aggregation mechanisms by SANS

Providing structural parameters for the species nucleating fibrillization is complicated due to the fact that at low concentrations of proteins this nucleus represents the thermodynamically least-favorable intermediate [240]. However, a mathematical model by Powers et al. of the nucleation-dependent aggregation proposes that the nucleus can become thermodynamically stable and prone to accumulation above the supercritical concentration [238]. It was shown that the stability of the nucleus increases with an elevating concentration of the monomer, and finally, if concentration is sufficiently high, the nucleus becomes more stable comparing to the monomer. The concentration at which this phenomena can be observed is referred to as the "supercritical concentration". When the used concentration is above the supercritical concentration, the monomer, but not the nucleus, is the species of the highest energy status present on the fibrillization pathway, and the aggregation reaction initially follows an irreversible-like polymerization. The observed early species, therefore, do not correspond to a thermodynamic nucleus, but are referred as a structural nucleus. Following this line of reasoning, it is possible to assume that the three-dimensional structure of the structural nucleus is thus the same as the thermodynamic nucleus existing at low concentrations [238]. Nevertheless, the structural nucleus is still intrinsically difficult to isolate and describe. The majority of currently used techniques for characterization of aggregation mechanisms are not ideal as they either upset the

equilibrium within the fibrillation process (in the case of x-ray crystallography or NMR), or samples are exposed to surface-mediated defects (potentially introduced by scanning tunneling microscopy, AFM, and cryo-EM). On the other hand, small-angle neutron scattering permits the assessment of the aggregation process directly in solution, so that the separate species existing during the evolution of equilibrium can be examined. Within this work, I describe our SANS experiments on fibrillization of normal and pathological-like htt peptides and presented the low-resolution structures of various intermediates formed during the process of their aggregation determined by ab initio reconstruction algorithms. Moreover, in the case of normal htt peptide, which demonstrates a nucleation-dependent mechanism, the structure of a 8-mer, potentially representing the htt structural nucleus, is revealed, yielding novel details on its conformation and further extending our current knowledge on the initial stages of amyloid formation.

2.2 The Theoretical Background on Small-Angle Scattering

2.2.1. Underlying physics of SANS

In order to have a good understanding of SANS data presented within the scope of this work, a short description of the essential elements of the nature and underlying physics of the technique is presented below.

Neutron particles in a beam possess the properties of plane waves, and because neutrons fundamentally interact with the nucleus of atoms they are able to penetrate significantly through most matter. Upon collisions with atoms in a

protein, neutrons scatter radially from each scattering center (the atoms in the molecule). Due to their wave like nature coherently scattered neutrons are capable of both constructive and destructive interference [241]. Ultimately the intensity of the scattered neutron beam as a function of scattering angle (or momentum transfer) depends on the shape and conformation of the molecule as well as the organization of these molecules in solution (or other phases). Even though small-angle solution scattering is frequently referred to as a low-resolution technique (since it does not yield any insight on atomic coordinates), it would be more suitable to identify it as a powerful technique in describing the structures of biological molecules, in terms of size and shape, with a high precision [242]. Ultimately, it is not due to the resolution limits of the technique, but mostly to the rotational averaging of the molecules freely moving in solution that constraint the informational content of SANS. For crystallographers, resolution is defined as the highest angle data measured, because it states for a value of the minimum distances between scattering points that is possible to distinguish. However for SANS professionals, resolution limits are described more frequently as the smallest angles at which data can be collected, since it is this parameter that establishes the largest distances that can be described by the data [241]. This observation is a consequence of the reciprocal features of the scattering data in reference with real space dimensions, and the requirement for the minimum angle necessary to describe a scattering particle with a particlar size is provided in detail below.

Although there are technical specifics associated with each small-angle scattering instrument, schematically the basic experimental setup is comparatively

straightforward (Figure 2.2). A highly collimated neutron beam is directed onto the sample, commonly containing protein or macromolecular complexes in solution with a relatively high concentration (>3 mg/mL). The scattered neutrons are captured on a detector, while the direct beam (unscattered neutrons) is commonly taken up by a beam stop; it is the size and position of the direct beam that play the major role in specifying the minimum angle measurable in a particular experiment.

Conventionally, the scattering profile is usually described by intensity (I) as a function of the amplitude of the scattering vector or momentum transfer, q:

$$q = \frac{4\pi \sin\theta}{\lambda}, \qquad (6)$$

where λ represents the wavelength of neutrons, and θ is half the angle between the incident and scattered neutrons (Figure). For fixed λ, I(q) vs. q represents the intensity as a function of scattering angle. Mathematically, the relationship between the shape of a scattering object and its resultant scattering pattern can be articulated as follows:

$$I(q) = \left\langle \int \left| \rho(\vec{r}) - \overline{\rho}_s) e^{i\vec{q}\cdot\vec{r}} d\vec{r} \right|^2 \right\rangle, \quad (7)$$

where $\langle\ \rangle$ describes the rotational average and describes the scattering length density (SLD) distribution of the biomolecule in Cartesian space or the shape of the molecule as seen by SANS. Therefore $\rho(\vec{r}) - \overline{\rho}_s$, which represents the variation in scattering length density between the element located at $\vec{r}$ within the biomolecule and SLD of the surrounding solvent, is termed as the contrast.

The direct determination of the shape profile, $\rho(r)$ from the measured SANS data is not possible as can be seen mathematically by the rotational averaging term in equation 7. However the Fourier transform of the I(q) data does provide real space structural information in terms of the interatomic pair distribution function, P(r), of the scattering particle (Figure 2.2, C).

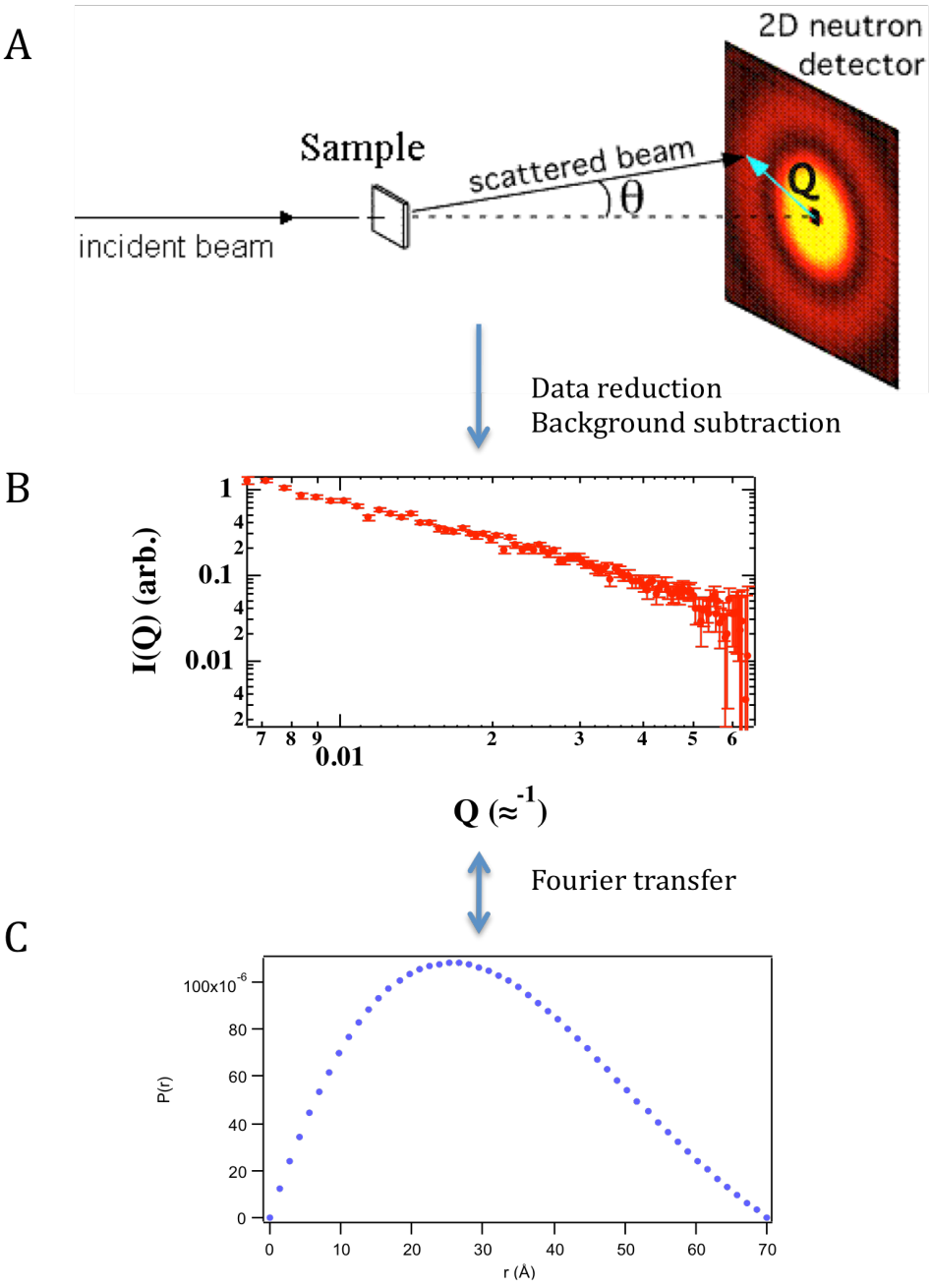


Figure 2.2. Small-angle neutron scattering experimental set-up and data analysis.

A. The scheme of experimental setup, showing the collimated neutron beam, used to illuminate the sample, containing a solution of biomolecules. The scattered neutrons are captured by 2-D detector. B. The scattering pattern is expressed in terms of scattering intensity I(q) as a function of angle q. C. The shape of biomolecule can be deduced from the Fourier transform of I(q) vs. q profile which yields pair distribution function (Pr).

P(r) defines the frequency of interatomic distances (r) within a structure of a protein and reflects the symmetry shape and size of a single protein or multi-component protein complexes [243].

From a SANS profile, the two major parameters linking to the overall dimensions and shape of the proteins can be reliably determined: the forward scattered intensity, which is also referred to as zero-angle scattering, I(0), and the radius of gyration, Rg. I(0) is not measured directly, as it is impossible to differentiate it from the neutrons that penetrate through the sample unscattered, but can be calculated by extrapolation. Given N particles within a set volume (V) and a SLD contrast between the protein particles and solvent of $\Delta\rho$, I(0) is directly related to N and to the square of the scattering volume V (see eq. 8). On the other hand, by replacing the sample volume with the partial specific volume (υ) of the biomolecule, equivalent expression can be written in terms of the concentration (C), molecular weight (M) and Avogadro's number N_A (eq.8 right hand expression):

$$I(0) = N(\Delta\rho V)^2 = \frac{C\Delta\rho^2\upsilon^2 M}{N_A}, \quad (8)$$

Thus one can derive the molecular weight of the scattering species through experimental determination of I(0), thereby providing insight into the oligomeric state of the proteins in solution. The radius of gyration refers to the root-mean-squared distance of all basic scattering volumes from their centre of mass corrected for their SLDs. Objects possessing the same volume but varying shapes will demonstrate distinctive Rg values. Therefore, Rg describes the distribution of

mass in the biomolecule.

A very valuable method for prompt estimation of Rg and I(0) comes from the Guinier approximation [244, 245]. Guinier demonstrated that for adequately small angles, q:

$$I(q) = I(0)e^{\frac{-q^2R_g^2}{3}}. \qquad (9)$$

From this equation, a linear fit of ln[I(q)] vs. q^2 provides the values of Rg and I(0) from the slope and intercept with y axis, correspondingly. In the case of globular proteins, this relationship is valid only when $q*R_g < 1.3$, whereas for molecules demonstrating elongated shapes this limit needs to be set even smaller, following the rule: the more extended the molecule, the lower the limit [246]. The requirement for a linearity of a Guinier plot when studying globular proteins is a quick and strong quality control tool [241].

2.2.2 Contrast Variation

In order to obtain information on the molecular shape from the SANS data, one can assume a protein to be an object with homogeneous scattering length density (SLD) of a mean contrast $\Delta\rho = \overline{\rho}_p - \overline{\rho}_s$, with the underlying assumption that the internal density variations inside a protein are negligible compared to the divergence between the average scattering density of the protein $\left(\overline{\rho}_p\right)$ and that of its solvent $\left(\overline{\rho}_s\right)$. Equations (7) and (8) in that case entail that the intensity of the scattering from a protein in solution is correlated to the square of the contrast, so

that:

$$I(q) \propto \Delta\rho^2, \quad (10)$$

where $\Delta\rho$ can be directly obtained from the atomic organization of the protein and solvent it is in. If the average SLD of the protein is equal to the one from the solvent $\left(\overline{\rho}_p - \overline{\rho}_s = 0\right)$, practically no small-angle scattering from the protein is detected, since the input to the scattering coming from the internal density variations are normally minor. Contrast variation (CV) studies including the systematic changes of protein and solvent SLDs can be utilized to significantly assist in investigation of protein complexes or protein hydration properties [247-250].

Contrast variation is a particularly useful concept in neutron scattering. Because neutrons interact only with the atomic nuclei, the strength of the scattering interaction can depend greatly on the isotopic state of the atom. Fortunate for biological samples, one of the major variations in an extent of neutron scattering comes from highly abundant hydrogen (^{1}H) and its stable isotope deuterium (^{2}H or D). As a result, it is feasible during a SANS experiment to create separate contrasts for two associating proteins if one of them is in deuterated form [251]. For such a protein system different mixtures of H_2O and D_2O in the aqueous buffer can be used to "blank" out one of the proteins. For the purpose of this work, the difference in SLD between hydrogen and deuterium was used to conduct the contrast variation experiments aimed in determination of the hydration profiles of htt mature fibrils and verification of their proposed internal

structure (Appendix).

There are several factors that make SANS quite beneficial compared to other small-angle scattering techniques. In contrast to X-ray scattering, where X-rays are scattered by electrons, neutrons are scattered by atomic nuclei, which is non-destructive, non-ionizing penetration for biomolecules. This factor makes SANS very useful in a long-term kinetic reaction, in which samples are exposed to the beam for an extended period of time. In the present work, the kinetics of htt peptide aggregation were followed for about 60 hours, which would be unacceptable for x-ray experiments due to inevitable degradation and significant denaturation of the sample [252]. An additional advantage of using neutron scattering experiments is that data can be easily placed on an absolute scale, which is necessary for quantifying aggregate molecular weight and mass per length. This is not standard and more prone to large errors with X-rays. Lastly, SANS allows investigation of structures ranging in dimensions from 10-1000 Angstroms, following the growth of earliest smaller particles (such as initial intermediates) into mature htt fibrils and subsequent fibril-fibril interactions.

2.2.3 *Ab initio* shape reconstruction

Ab initio shape reconstruction represents one of the methods that can be used in SANS data analysis. This shape description algorithm utilizes Monte-Carlo search to create an outer envelope of the macromolecule of interest [253]. The software by Svergun et al. is currently the most employed tool for this type of analysis, which executes a method to rebuild an ab initio low resolution profile of

randomly oriented biological molecules in a solution from its SANS data [254]. A search volume, containing the macromolecule, which is simply a sphere of adequately large radius, is packed with the number, N, of densely arranged spheres (traditionally called dummy atoms) of a particular radius [255]. Taken into consideration the fixed spatial locations, the envelope of the dummy atom model is entirely expressed in term of a vector X containing N components which allocates each dummy atom either to the protein or to the solvent segment. In order to describe any given structure reliably, the required amount of dummy atoms is usually considered to reach several thousands. The purpose of shape restoration from the scattering data is therefore shifted toward the problem of determining a configuration X where a target function f(X) is minimized [253]. The target function takes into consideration the inconsistency between the experimentally obtained data and the calculated scattering profile of the ab initio model [256], besides a few other features of the model measured as penalties. A solid and inter-connected model is provided by the introduction of the penalties for the looseness and disconnectivity, respectively. The presence of peripheral penalty guarantees that the model is located proximal to the middle of the search volume [253]. The input from the penalties to the target function is estimated to range within 10-50% for the absolute model. Simulated annealing (SA) is utilized to achieve the overall minimization of the goal function [257].

When it comes to all ab initio methods, a distribution of shapes is found depending on the starting random models, that is why numerous runs are frequently carried out and analyzed to elucidate the most constant parameters of

the model. A program DAMAVER [258] executes the fast superposition of a series of models and estimates the level of the shape similarity. Subsequently, it detects the most representative ensemble member of the highest probability and creates the mean model. The utilization of these methods permits the evaluation of the uniqueness of the *ab initio* models calculated from scattering data [259].

Ab initio reconstruction was performed to obtain the building blocks of insulin and actin fibrils from the conventional three-dimensional small-angle scattering data [235, 260]. A 2-D dummy atom modeling program was developed and implemented to specify the cross-sectional organization of protein components in the actin filaments assembly [260]. In this work, we utilize the *ab initio* reconstruction algorithm to describe the shape of the earliest intermediates formed during the aggregation of normal and pathological-like htt peptide. While the performed 3-D analysis for shape determination of the early species was very informative, the attempted two-dimensional reconstruction of htt fibrils' cross-section was not found to be suitable for this particular system.

CHAPTER 3.

Results and Conclusions

3.1 Structural Formation of Huntingtin Exon 1 Aggregates Probed by Small-Angle Neutron Scattering

(presented as a paper published in Biophysical Journal (2011), 100(10), 2504-2512)

3.1.1 Introduction

Considering the vast amount of knowledge currently present in the field of conformational diseases, it is unpleasantly surprising that the majority of them remain incurable. One of the possible explanations to this phenomenon comes from the lack of information on the correlation between the evolution of the forming aggregates and the mechanism of neural toxicity. The clear understanding on the pathology of HD requires an insight into the structural dynamics of htt aggregation, as well as identification of the species responsible for cellular death. As indicated in previous chapters, it is currently known that the pathological events are initiated by the proteolytic cleavage of htt protein. This truncated fragment containing the expanded polyQ region has a strong tendency to misfold toward a stable three-dimensional conformation enriched in β-sheet structure with self-assembly resulting in fibrillar aggregates. Although the link between polyQ length, aggregation, and disease is inescapable – the longer the polyQ repeat, the earlier the age of onset and the more severe the symptoms – numerous studies

in HD cell and animal models show a strong dichotomy between aggregate presence and neuronal toxicity[261-263]. Some results even imply that aggregates are in fact beneficial for neuronal survival. In order to address the question "what makes an aggregate toxic?" an increasing number of reports in the amyloid field have appeared over the last decade focusing on the mechanism of aggregation implicated in a variety of neurodegenerative diseases, including Alzheimer's and Parkinson's diseases, and in which the importance of protein oligomerization in pathogenesis has been emphasized[4].

Various biophysical methods, including AFM [264, 265] and analytical ultracentrifugation [266] have been applied toward the identification and structural characterization of early oligomers formed during Htt fibrillization with the smallest structures detected being spherical oligomers of 50 Å diameter [265]. Small-angle neutron and X-ray scattering have been efficiently applied to various aggregation systems including amyloid-β fibrils[232, 234, 235, 267]. Here, SANS is utilized in a time-resolved fashion to pinpoint the various transient intermediates that form along the aggregation pathway for a truncated mutant Htt exon 1 peptide. This peptide contains the 17 N-terminal residues of Htt followed by a 42 glutamine repeat, a 10 proline stretch, and 2 lysines on the C-terminus for enhanced solubility ($NtQ_{42}P_{10}$ for short). With SANS, the structures of biological macromolecules can be directly measured in solution without radiation degradation. Within the scope of this chapter, new quantitative and detailed structural information is reported on the aggregation of $NtQ_{42}P_{10}$ by monitoring the growth of early-formed species into eventual fibrils.

3.1.2 Materials and Methods

Materials

The peptide with sequence MATLEKLMKAFESLKSFQ$_{42}$P$_{10}$K$_2$ (NtQ$_{42}$P$_{10}$) and molecular mass 8583 Da was synthesized by solid-phase FMOC chemistry (Keck-Yale facility) and purified by reverse phase HPLC. Mass spectrometry confirmed the purified peptide contained 42 glutamines. Trifluoracetic acid (TFA) (>99.5%, Thermo Scientific), 1,1,1,3,3,3-hexafluoro-2-isopropanol (HFIP) (99.5+%, Acros Organics), 10X PBS powder (Fisher Scientific), and D_2O (99.9% D, Cambridge Isotope Laboratories, Inc., Andover, MA) were used without further purification.

Peptide solubilization and disaggregation

The NtQ$_{42}$P$_{10}$ peptide was solubilized and disaggregated similar to established procedures [268]. Briefly, the purified peptide was dissolved in 50:50 TFA:HFIP, incubated for 3 hr, and then dried under Argon gas. For HPLC kinetics, the peptide was first solubilized into 0.05% TFA D_2O, ultracentrifuged at 356k × *g* at 4 ^oC for 2 hr, and the supernatant brought up to 1X PBS D_2O to initiate the reaction.

High performance liquid chromatography

An Agilent 1100 series HPLC equipped with a C3 column was used with a reverse phase water (+ 0.05% TFA) / acetonitrile (+ 0.05% TFA) gradient. The peptide concentration was determined by HPLC using a standard curve generated

from $NtQ_{22}P_{10}$ peptide (same sequence as $NtQ_{42}P_{10}$ but with 22 glutamines) and amino acid analysis. To overcome the 40 min HPLC runtime per sample, $NtQ_{42}P_{10}$ aliquots removed from the reaction vessel were quench-frozen in liquid nitrogen and stored at –80 °C. The time point aliquots were then individually thawed, ultracentrifuged at 356k × *g* at 4 °C for 20 min, and supernatant injected onto the HPLC for monomer concentration measurement.

Thioflavin T fluorescence

The fluorescence of Thioflavin T (ThT) was monitored using a Perkin Elmer LS 55 fluorescence spectrometer with excitation and emission wavelengths, λ_{ex} = 450 nm and λ_{em} = 482 nm, respectively. Sample aliquots of 20 μl were removed from the reaction vessel, diluted with 130 ml 1X PBS, and then 5 μl of 2 mM ThT solution was added to give 64.5 μM ThT final. Equivalent background runs of ThT in 1X PBS were performed and subtracted from the sample signal.

Congo red birefringence

An aliquot of the $NtQ_{42}P_{10}$ fibril sample grown in 1X PBS D_2O and measured by SANS was placed on a microscope slide and dried at 37 °C overnight. The sample was fixed with 10% formalin, rinsed with deionized water, stained with 1% Congo red solution for 30 min at room temperature, rinsed again with deionized water, and then mounted with a coverslip after drying. Images were collected using a Leica microscope (40×) and cross-polarizers.

Transmission electron microscopy

TEM was performed using a Hitachi H-600 electron microscope in the Biological Sciences Division, University of Tennessee. Aliquots of the $NtQ_{42}P_{10}$ fibril sample grown in 1X PBS D_2O and measured by SANS were fixed to collodion/carbon grids and stained with a 1% uranyl acetate solution for imaging. Single fibril diameter was measured at 21 different positions on the TEM image using the software program ImageJ [269] and converted from pixels to nanometers using the image scale bar. The fibril radius in Angstroms was then calculated to directly compare with the SANS results.

Small-angle neutron scattering

SANS experiments were performed on the CG3 BioSANS instrument at the High Flux Isotope Reactor (HFIR) at Oak Ridge National Laboratory. Neutrons of wavelength, λ = 6.0 Å with a wavelength spread, $\Delta\lambda/\lambda$ = 0.14 were used. Time-resolved measurements were made with a 6.9 m sample-to-detector distance with static sample measurements also made at 1.2 and 15.4 m to obtain data at higher and lower respective wavevector transfer, $Q = 4\pi\sin(\theta)/\lambda$, where 2θ is the scattering angle. Samples were loaded into 2 mm pathlength rectangular-shaped quartz cuvettes (Hellma USA, Plainville, NY). For the time-resolved measurements, 15 min neutron exposure times were collected. Samples with sedimenting aggregates were dispersed by inverting the cuvette several times to ensure sample homogeneity and then measured for 15 min. It was found that no significant

sample settling occurred over this timeframe as gauged by comparing the total integrated intensity on the detector for consecutive 5 min exposures. For static measurements on $NtQ_{42}P_{10}$ fibrils, 0.5, 1, and 2 hr exposure times were collected at 1.2, 6.9, and 15.4 m, respectively. These longer exposure times were achieved by binning multiple 15 min exposures obtained upon first dispersing the sample to homogeneity. Scattered neutrons were detected with a 1 × 1 m two-dimensional position sensitive detector with 192 × 192 pixels.

$NtQ_{42}P_{10}$ was solubilized and disaggregated following the above procedures except the Argon-dried film was resuspended first in 0.05% TFA D_2O to dissolve the peptide, brought up in PBS D_2O, and then centrifuged at 16.1k × *g* for 2 min to remove any particulates. The $NtQ_{42}P_{10}$ fibril sample was prepared in the same manner and allowed to aggregate for >500 hr at 22 °C before measurement. The final $NtQ_{42}P_{10}$ concentration was determined by HPLC to be 0.93 mg/mL (108 mM) in PBS D_2O for the kinetics sample and 0.76 mg/mL (89 mM) for the fibril sample. All neutron scattering experiments were conducted at 20 °C.

Data reduction was accomplished using Igor Pro software (WaveMetrics, Lake Oswego, OR) with SANS macros developed by Dr. Ken Littrell (ORNL). The total two-dimensional scattering was corrected for the scattering from the quartz cell and from stray neutrons in the room and for non-uniform detector response. Then, the scattering was normalized by the incident beam flux, radially averaged to obtain $I(Q)$ vs. Q, and then $I(Q)$ was placed on absolute scale intensity using porasil B as a calibration standard with known scattering cross-section. Background scattering from the buffer, $I_{bg}(Q)$, and the Q-independent excess incoherent

scattering, $I_{inc}(0)$, were subtracted from the total scattering, $I_{tot}(Q)$, to obtain the protein sample scattering, $I_{sam}(Q) = I_{tot}(Q) - I_{bg}(Q) - I_{inc}(0)$, with $I_{inc}(0)$ determined from Q > 0.35 Å^{-1} assuming the residual difference in intensity is completely incoherent scattering.

Fitting procedures for the SANS data followed a Guinier analysis [245], as described in the previous section. The peptide scattering length density, ρ_{prot} = 3.564 × 10^{10} cm^{-2} was calculated from the atomic composition using the program CRYSON [270] that accounts for H/D exchange. The D_2O scattering length density used was ρ_{D2O} = 6.404 × 10^{10} cm^{-2}.

Modified Guinier fits were made by plotting ln[$Q \cdot I(Q)$] vs. Q^2 to obtain the $slope = -R_c^2/2$, where R_c is the cross-sectional radius of gyration and related to the cylindrical radius by $R = \sqrt{2}R_c$. Then, from the intercept ($I_c(0)$), the mass per length was calculated

$$M_L = \frac{I_c(0)N_A}{\pi c(\Delta\rho)^2 \bar{\upsilon}^2} \qquad (11)$$

Values for ρ_{prot}, ρ_{buf}, and $\bar{\upsilon}$ were the same as above.

3.1.3 Results and Discussion

$NtQ_{42}P_{10}$ aggregates formed at high concentration are amyloid-like

Using equivalent conditions to our SANS experiments, which require peptide concentrations in the mg/mL range, we performed kinetic studies on $NtQ_{42}P_{10}$ aggregation with 93 μM monomer in phosphate buffered saline (PBS)

D_2O at 22 ºC. Monomer disappearance was followed by reverse phase HPLC and fibril production monitored by the fluorescence of ThT, an amyloidophilic dye. Fig. 1A shows that the aggregation reaction proceeds rapidly with the amount of soluble monomer decreasing in the absence of a noticeable lag phase and a single exponential rate constant k_{mon} = 1.28 ± 0.09 hr^{-1}. Only 10% monomer remained in solution after 2 hr and was completely consumed by 6.5 hr. Concurrently, the ThT fluorescence signal, indicative of the presence of amyloid-like structures, also lacks an observable lag phase and reaches a plateau at approximately 2 hr with an exponential rate constant k_{fib} = 1.1 ± 0.3 hr^{-1} similar to k_{mon} (Figure 3.1A). The inverse dependence of the lag phase duration on monomer concentration is a phenomenon already known from other amyloid proteins [271], and an observed lag phase recently was reported for a F17W mutant $NtQ_{30}P_6$ peptide [226]. Once formed, the $NtQ_{42}P_{10}$ aggregates were characterized by Congo red birefringence and TEM. Aggregates stained with Congo red and viewed by polarized light microscopy display a gold/apple-green birefringence characteristic of amyloid-like fibrils (Figure 3.1B).

Likewise, TEM shows the fibrillar nature of the $NtQ_{42}P_{10}$ aggregates (Fig. 3.1C), where both individual fibrils and fibril bundles were observed. The radii of individual fibrils range from 28 to 56 Å with an average radius, R = 40 ± 8 Å. In summary, we find that $NtQ_{42}P_{10}$ aggregation at high monomer concentrations in D_2O buffer does not drastically affect the amyloid properties of the resulting fibrils but only the kinetics of formation – faster aggregation rates and shortened lag phase times in the ThT assay.

Quantifying the $NtQ_{42}P_{10}$ aggregates by SANS

With the scattering vector range chosen for the time-resolved SANS measurements, $0.01 < Q < 0.1$ Å^{-1}, we monitored size changes that occur from hundreds down to tens of Angstroms, respectively. To optimize the time-resolved measurements for sample kinetic rate, data binning, and instrumental signal-to-noise, we first performed several experiments at different starting monomer concentrations and varied the neutron exposure time binning. As a result, the kinetic run described here is the best representative, where we found that 15 min exposures and 50 to 100 mM $NtQ_{42}P_{10}$ provide a sufficient balance between time resolution and SANS signal-to-noise ratio for adequately following structural changes in time. These initial experiments also confirmed reproducibility with comparable results obtained and an expected direct dependence of the kinetic rate on peptide concentration. For the results described below, each of the time-resolved SANS profiles represent 15 min exposure intervals where structural changes taking place over this time were integrated to reflect the average structure. Select SANS profiles acquired are shown in Fig. 3.2 and illustrate how the scattering intensity ($I(Q)$) increases and changes shape over the time-course of the aggregation reaction.

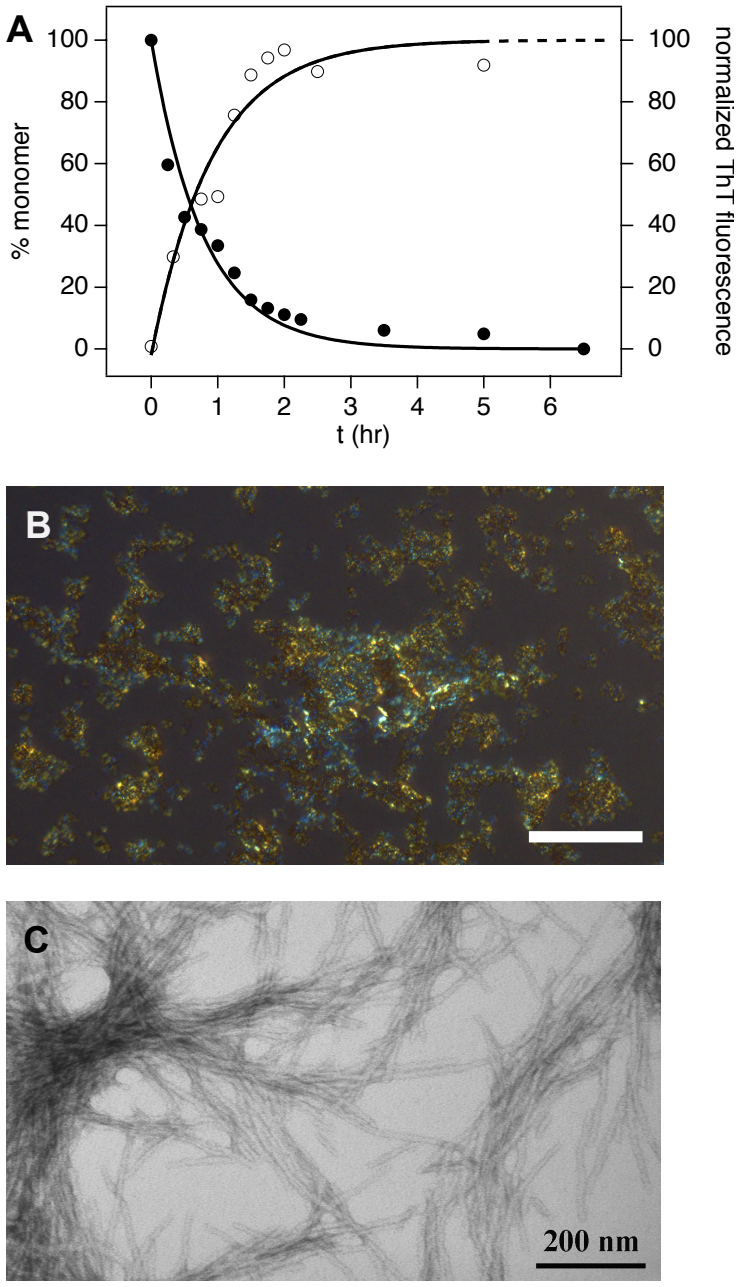


Figure 3.1 Aggregation kinetics of $NtQ_{42}P_{10}$ and the resulting fibrils. (*A*) Disappearance of monomeric $NtQ_{42}P_{10}$ (93 μM) from the supernatant solution (solid circles) and increasing ThT fluorescence (open circles) corresponding to the formation of ThT-binding amyloid-like aggregates. Solid lines are single-exponential fits to the data with an extrapolation of the ThT fluorescence fit (dashed line). (*B*) $NtQ_{42}P_{10}$ fibrils stained with Congo red exhibit birefringence by polarizing optical microscopy (scale bar = 50 mm), which is indicative of amyloid-like aggregates. (*C*) TEM image of the fibrils formed by $NtQ_{42}P_{10}$ with a calculated fibril radius, R = 40 ± 8 Å and range of 28 to 56 Å.

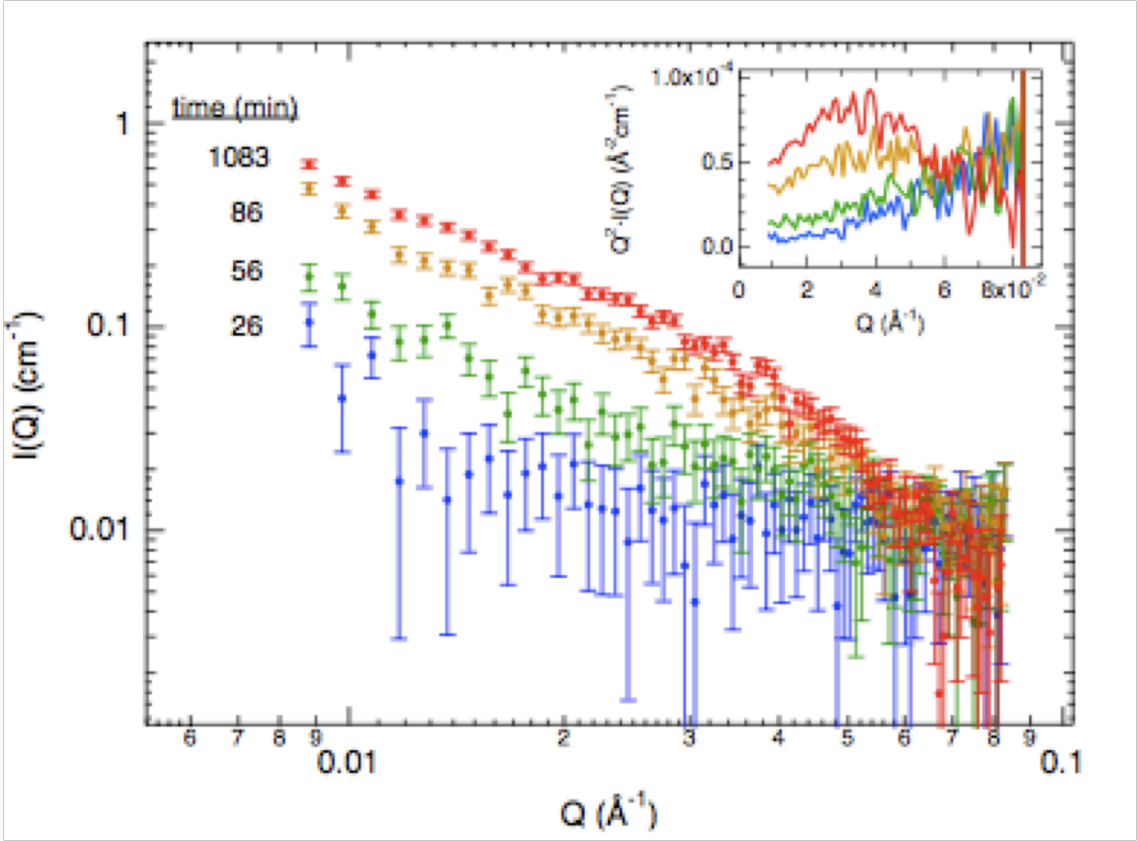


Figure 3.2. Time-resolved SANS on $NtQ_{42}P_{10}$.
Select time-resolved SANS profiles, $I(Q)$ vs. Q, of $NtQ_{42}P_{10}$ at 26 (blue), 56 (green), 86 (yellow), and 1083 min (red). Kratky plots, $Q^2 \cdot I(Q)$ vs. Q, (inset, colors correspond to the same times in the main plot, error bars omitted for clarity) show the formation of a maximum about $Q \approx 0.03$ Å^{-1} with time, which is indicative of an increase in folding and structural rigidity. Peptide concentration was 108 mM.

Changes in the shape of the scattering profiles are emphasized in the Kratky plot representation, $Q^2 \cdot I(Q)$ vs. Q, (Figure 3.2 inset) with the formation of a maximum at $Q \approx 0.03$ Å^{-1} revealing a conformational transition over time toward increased folding with greater structural rigidity. To acquire quantitative information on the size and oligomerization state of the $NtQ_{42}P_{10}$ aggregates, we applied Guinier analysis (see Section 2.2 and 3.2) [245] on these SANS profiles and determined the radius of gyration (R_g) and apparent molecular mass (M). Figure 3.3A displays a representative Guinier fit to the SANS curve obtained at 26 min. We found two linear fit regions, a proper one ($Q_{max} \cdot R_{g,1} = 1.06$) corresponding to a distinct population of small species ($R_{g,1} = 16.1 \pm 5.9$ Å) and a second approximate one ($Q_{max} \cdot R_{g,2} = 1.46$) that is indicative of larger species with $R_{g,2} = 57 \pm 24$ Å. The dimensions obtained for the larger aggregates were affected by additional influences including polydispersity and favorable interactions between smaller species. The continuous upturn in the Guinier plot at small Q also indicates the presence of aggregates larger than the Q-range of the instrument configuration.

Eventually, during the course of aggregation, the increasing anisotropy of the forming fibrils resulted in the fibril length exceeding our SANS measurements such that only the two-dimensional fibril cross-section was monitored. The advantage is that cross-sectional measurements were unaffected by fibril length polydispersity. At this point, the conventional Guinier analysis was no longer valid and the data were treated by modified Guinier plots as previously performed for SANS on amyloid-β fibrils [232]. The cross-sectional radius of gyration (R_c) and mass per

length (M_L) were obtained from the fits (see Section 3.2). Fitting errors (~5%) were the major source of error for M_L and taken into consideration. Slight variations in the partial specific volume ($\bar{v}$), which depend on protein type and folding state [272], can contribute an additional ~3% error. Including other errors conservatively results in ~10% total error propagated into the determination of M_L, which underscores the strength of the technique and allows us to interpret these values with high confidence. Figure 3.3B shows a representative modified Guinier fit to the SANS curve for 86 min, where two distinct linear fit regions were observed. The first one reveals a size of $R_{c,1}$ = 25.6 ± 1.8 Å ($Q_{max} \cdot R_{c,1}$ = 1.31), which corresponds to growing individual fibrils and closely matches the TEM calculated radius. The second linear region with a larger radius of cross-section, $R_{c,2}$ =59.8 ± 4.2 Å ($Q_{max} \cdot R_{c,2}$ = 1.18) correlates to fibril bundles formed by attractive fibril-fibril interactions.

Early-stage aggregate structural growth

As stated previously, SANS data for $NtQ_{42}P_{10}$ early aggregates forming within 71 min were analyzed using Guinier fits to obtain R_g and M, and we focus on the smaller aggregates having a true Guinier regime. When plotted as a function of time, $R_{g,1}$ and M_1 (Figure 3.4 A, B) show a progressive increase in growth. Dividing M by the peptide monomer molecular mass yields the average number of peptides (n) incorporated into an aggregate species (Figure 3.4 B right axis). We identified the earliest species to have dimensions intermediate between a dimer and trimer (n = 2.6 ± 0.4 peptides) having R_g = 16.1 ± 5.9 Å.

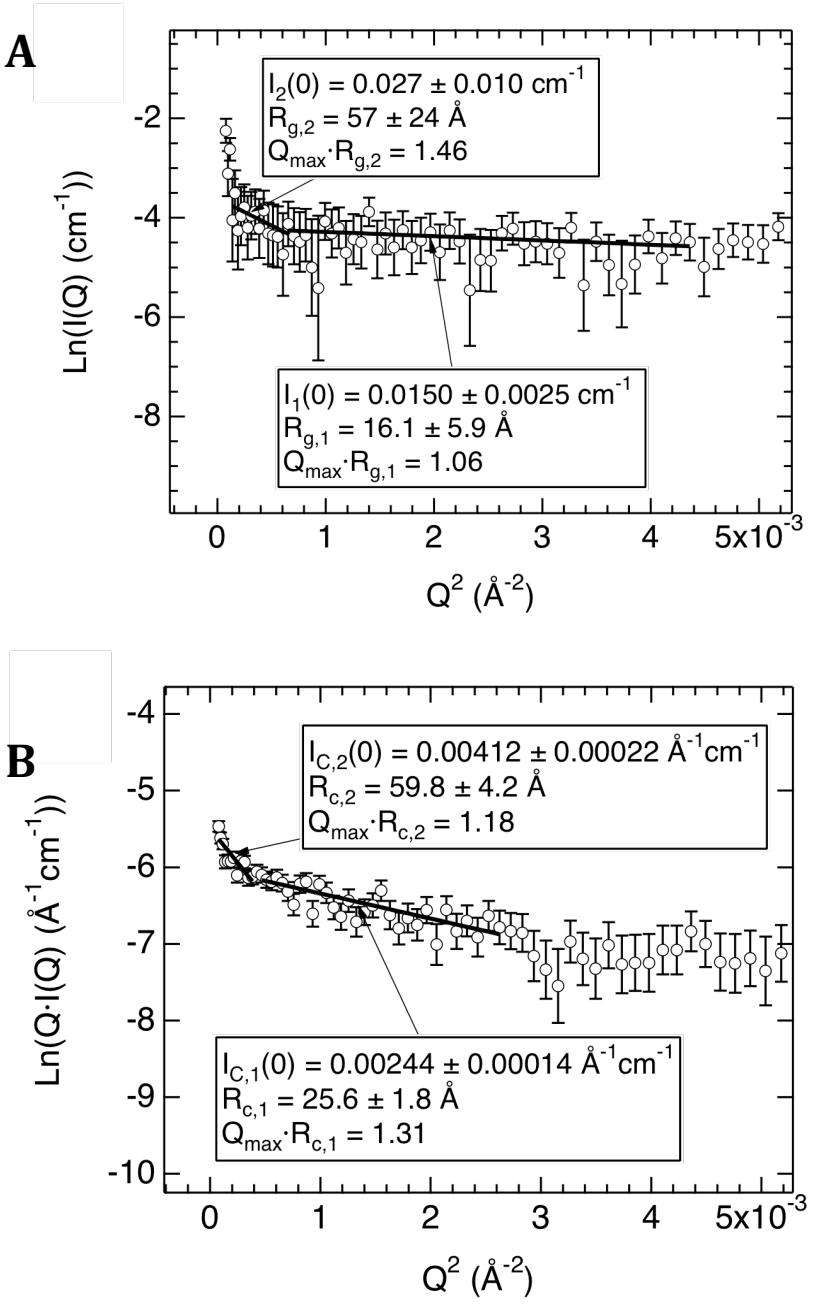


Figure 3.3 Time-resolved SANS data analysis on $NtQ_{42}P_{10}$ aggregation.
A. A representative Guinier plot and fits to the SANS curve for 26 min. B. A representative modified Guinier plot and fits to the SANS curve for 86 min. Error bars are standard errors from the radial average of the intensity data on the neutron area detector.

While no distinct scattering signal from the monomer itself was detected, the HPLC data in Figure 3. 1A show the monomer amount decreased from 60 to 25% over 30 to 75 min of the aggregation reaction. Since scattering is weight averaged, the oligomer scattering signal dominates and obscures the monomer scattering contribution.

For comparison, the radius of gyration of a monomer with a random coil conformation, $R_{g,rc}$, can be predicted from the relation

$$R_{g,rc} = 1.98N^{0.602} \qquad (12)$$

where N = number of residues [273]. In our case, the 71 amino acid $NtQ_{42}P_{10}$ has a predicted $R_{g,rc}$ = 26 Å, pointing out that the earliest species we observe here are smaller and more compact relative to a monomer random coil. These initial dimer and trimer species grow into oligomers of n = 14 ± 2 peptides with R_g = 46.4 ± 4.2 Å by 71 min of the aggregation reaction. It is worth noting that small mutant Htt oligomers composed of 5 to 15 proteins recently have been observed in a HD cell model [274]. From a linear fit to the change in the number of peptides with time (dn/dt), we found that 6 ± 2 peptides/hr are incorporated into the growing oligomers, a rate which then rapidly accelerates after 56 min.

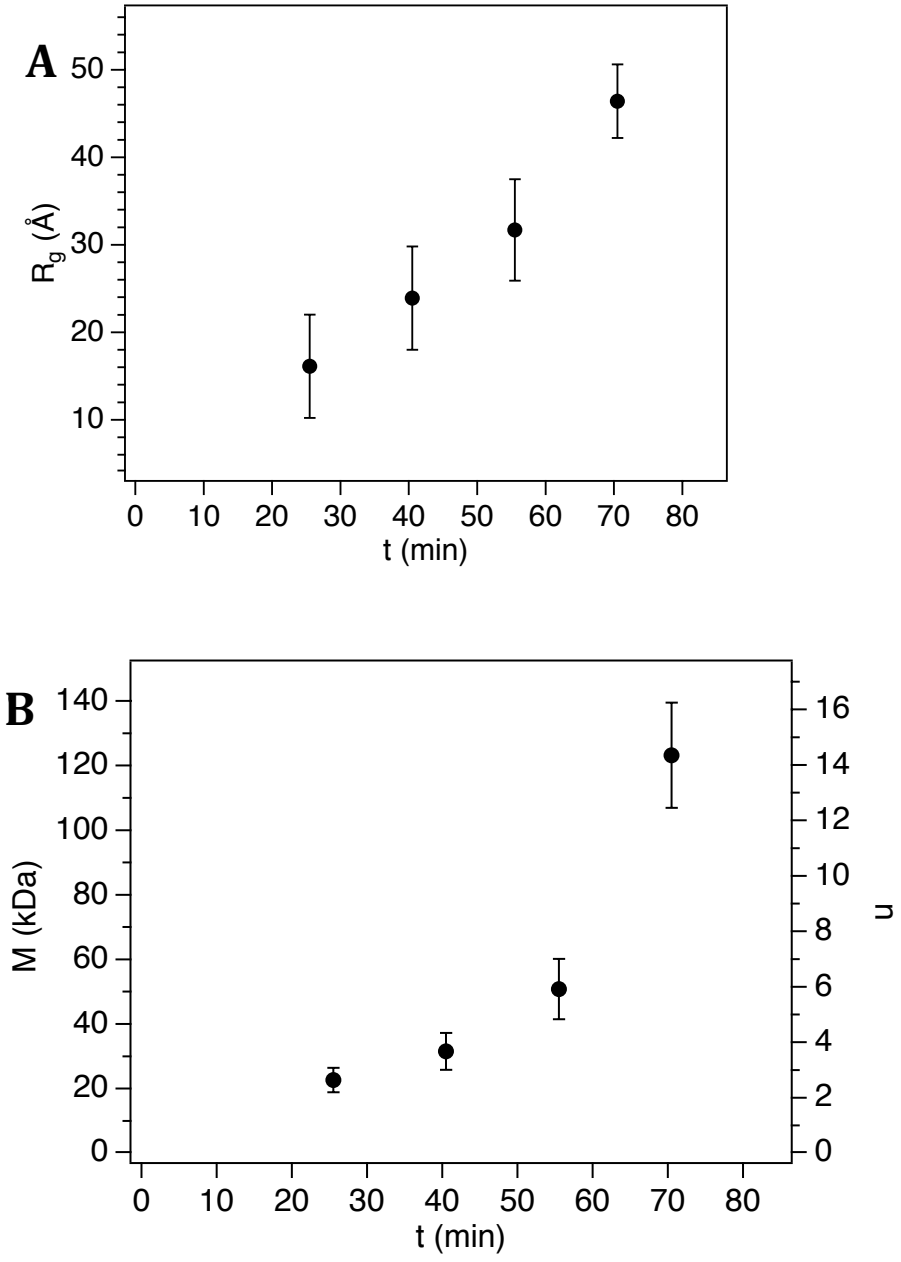


Figure 3.4. Early aggregates of $NtQ_{42}P_{10}$ formed within the first 71 min.
A. The smallest species progressively grow from an initial $R_{g,1}$ of 16 to 46 Å over this time. B. The corresponding change in molecular mass with time for the smaller aggregates, M_1 = 22.6 to 123 kDa, with the number of $NtQ_{42}P_{10}$ peptides per aggregate, n, (right axis) calculated from M_1.

Again, these growing oligomers are in the presence of larger structures, which we approximated from Guinier fits to obtain $R_{g,2}$ and M_2 ($1.46 \leq Q_{max} \cdot R_{g,2} \leq 2.21$). During the initial 71 min, $R_{g,2}$ increases from 60 to 90 Å with $M_2 \approx 41$ to 370 kDa ($n \approx 5$ to 43), respectively. For comparison, TEM on Htt exon 1 Q44 identified 40 to 50 Å diameter oligomers after 30 min aggregation [275], and AFM on Htt exon 1 Q53 showed spherical oligomers with diameters of 50 to 650 Å within 1 hr with fibrils present by this time [265]. Spherical oligomers of similar size were also observed by TEM for $NtQ_{30}P_6$ by 15 min of the aggregation reaction [226]. These reported spherical oligomers range from comparable in size up to much larger than the oligomers we detected by SANS. This is consistent with our analysis that aggregates larger than the size window of our time-resolved SANS configuration are simultaneously present.

By plotting R_g vs. M (Figure 3.5) to remove the time component of growth, correlations between the two parameters were discerned. To compare with the experimental results, two different cases were calculated and plotted: 1) an equivalent random coil and 2) a solid sphere. The equivalent random coil $R_{g,rc}$ was calculated using Eq. 12 with $N = M/120.9$, where 120.9 Da is the average molecular mass per residue of $NtQ_{42}P_{10}$. The solid sphere $R_{g,sphere}$ was calculated from M and $\bar{\upsilon}$ = 0.73 mL/g by

$$R_{g,sphere} = \sqrt{\frac{3}{5}}\left(\frac{3M\bar{\upsilon}}{4\pi N_A}\right)^{1/3} \qquad (13)$$

The initial $NtQ_{42}P_{10}$ dimer and trimer structures closely agree with the sphere model M and R_g. Since a sphere is the shape of minimal volume, this indicates

that these earliest and smallest species are spherical and compact. As the aggregates become larger in mass and radius, they slightly diverge from the sphere model but are always distant from the random coil model. The increasing anisotropy of the aggregates during growth can account for this divergence from the sphere model.

While formation of fibrils by monomer addition driven by aggregation of the polyQ region has been demonstrated and recognized for some time now [224, 276], recent work by Thakur *et al.* proposed an aggregation model where the 17 residue N-terminus of Htt fragments guides the initial assembly into oligomers [226]. This is consistent with AFM studies by Muchowski and colleagues that showed mutant Htt fragments first form oligomers before fibrils [264]. Our results above on $NtQ_{42}P_{10}$ illuminate the structural details of the oligomerization process with evidence for the initial formation of compact spherical dimers and trimers.

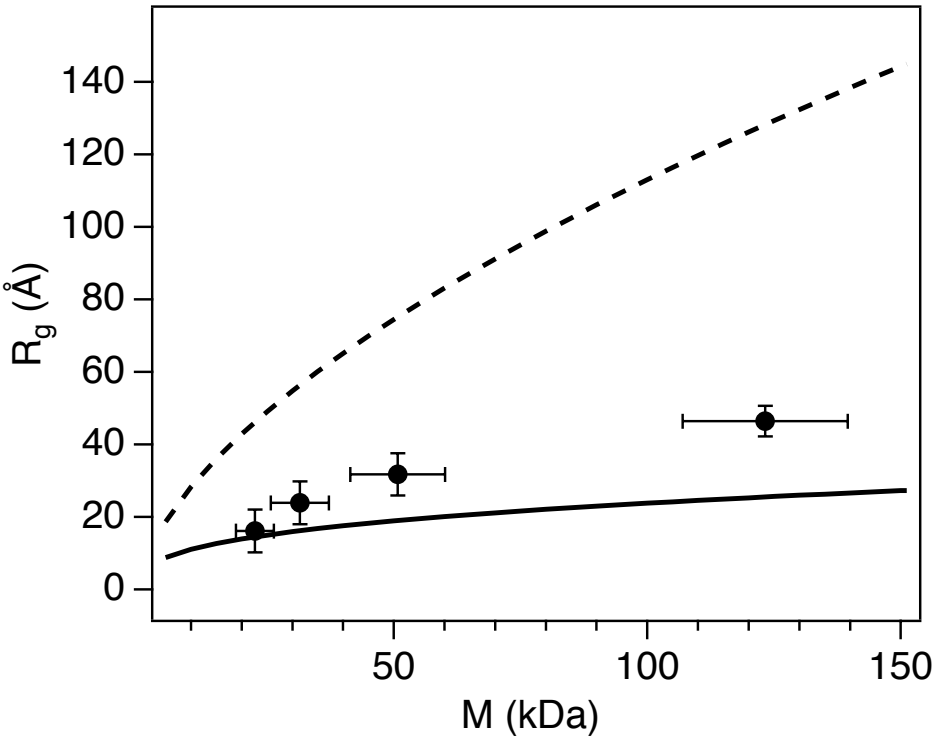


Figure 3.5 Structural insights on $NtQ_{42}P_{10}$ aggregation.
A plot of $R_{g,1}$ vs. M_1 compared to model calculations for a random coil (dashed line) using Eq. 4 and a solid sphere (solid line) using Eq. 5. Error bars are the standard error from the Guinier fits.

These directly identified early intermediates progressively grow in mass and size anisotropy toward becoming predominantly fibrillar structures. At this point, fibril growth dominates and is now discussed.

Later-stage two-dimensional fibril growth

SANS data for $NtQ_{42}P_{10}$ fibril growth were fit using modified Guinier plots (see Figure 3.3 B) to obtain the R_c and M_L as a function of time (Figure 3.6 A, B). Both the single fibril, with $R_{c,1}$ and $M_{L,1}$, and fibril-fibril aggregates, with $R_{c,2}$ and $M_{L,2}$, were properly treated by modified Guinier analysis. The open symbols in Figure 3.6 A were from measurements with aggregate sedimentation, which affected accurate determination of M_L (omitted from Fig. 3.6B and 3.7) for these times but did not appreciably affect $R_{c,1}$ and $R_{c,2}$ as seen by their close agreement with the encompassing data. Samples were re-dispersed to homogeneity just before measurement of the last three time points to ensure accurate M_L values. During the time-course for fibrillization, the larger aggregates were about twice the size and mass of the individual fibrils ($R_{c,2} \approx 2R_{c,1}$ and $M_{L,2} \approx 1.5M_{L,1}$), suggesting the larger aggregates are two aligned and favorably interacting fibrils.

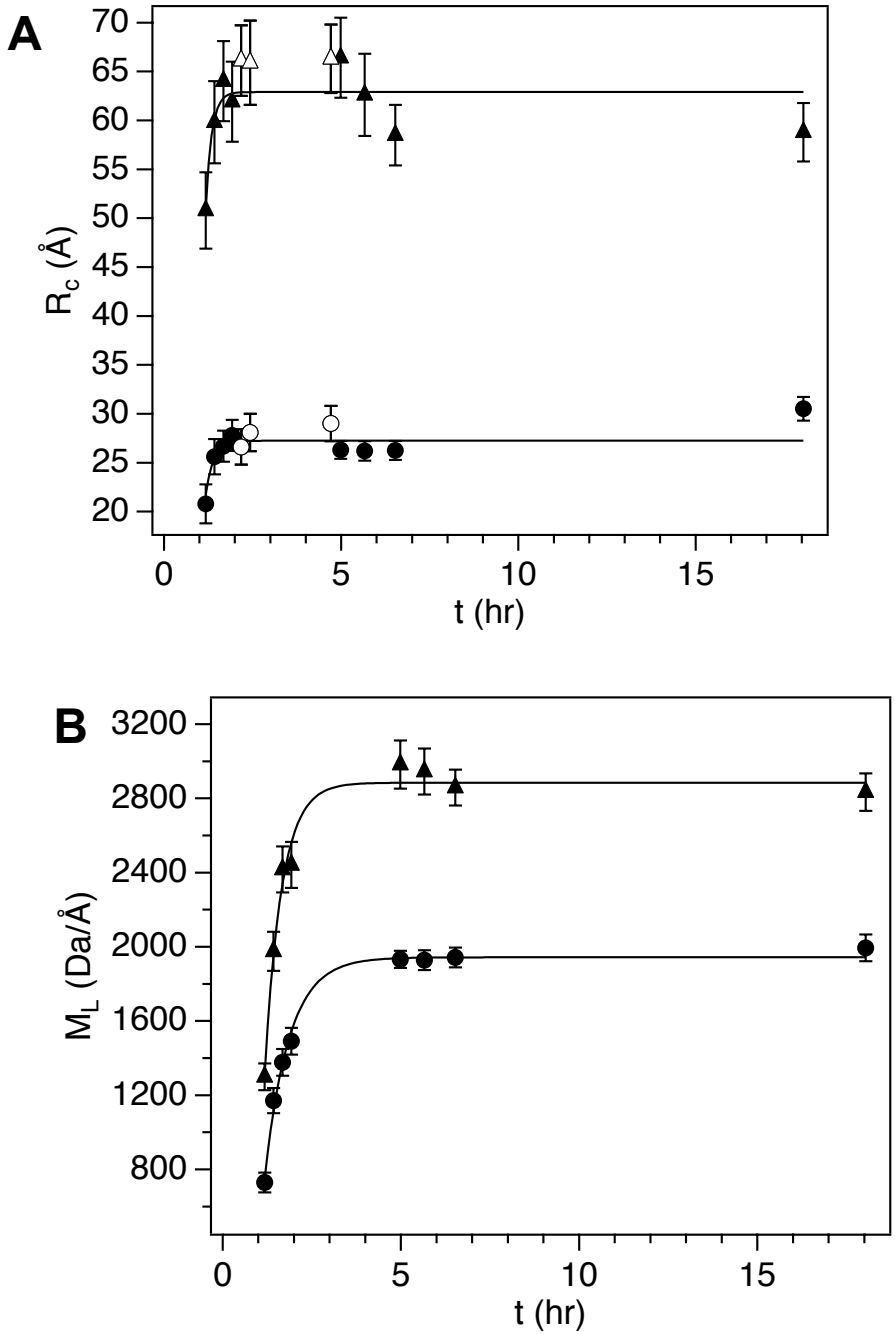


Figure 3.6.
Aggregates of $NtQ_{42}P_{10}$ beyond 1 hr are predominately fibrillar and their two-dimensional cross-section is analyzed. A. Change in the two characteristic length-scales, $R_{c,1}$ (circles) and $R_{c,2}$ (triangles), with time. Open symbols are R_c measurements where sample sedimentation was observed. Solid lines are single-exponential fits to the data. B. Corresponding changes in $M_{L,1}$ (circles) and $M_{L,2}$ (triangles). Solid lines are single-exponential fits to the data.

Single-exponential fits to R_c growth with time (Figure 3.6 A) yield saturating cross-sectional radius of gyration values: $R_{c,1}(\infty)$ = 27.2 ± 0.4 Å and $R_{c,2}(\infty)$ = 63 ± 1 Å. Single-exponential fits to the corresponding M_L changes with time (Figure 3.6 B) yield reliable rates: $k_{ML,1}$ = 1.5 ± 0.2 hr^{-1} and $k_{ML,2}$ = 2.1 ± 0.3 hr^{-1}, with saturating mass per length values: $M_{L,1}(\infty)$ = 1940 ± 30 Da/Å and $M_{L,2}(\infty)$ = 2880 ± 50 Da/Å. Most notably, the single fibril $k_{ML,1}$ closely matches the kinetic rates for monomer loss (k_{mon} = 1.28 ± 0.09 hr^{-1}) and fibril formation (k_{fib} = 1.1 ± 0.3 hr^{-1}) obtained from the HPLC and ThT assays, respectively.

Indeed, the $NtQ_{42}P_{10}$ aggregation kinetics describing monomer loss, formation of ThT-positive aggregates, and structural growth in R_c and M_L are in agreement with ~2 hr to reach saturation. This importantly demonstrates that the peptide monomer pool is depleted by about the same time as structural changes in R_c and M_L reach a plateau. Since SANS is insensitive to the fibril length, it is unknown from the SANS data alone if fibril elongation via monomer addition to the growing fibril ends continues well beyond the observed plateau in structural changes. Monomer loss followed by HPLC confirms this is not the case.

By plotting $R_{c,1}$ vs. $M_{L,1}$ (Figure 3.7 A) we removed the time component in order to assess how these two parameters are related for the single $NtQ_{42}P_{10}$ fibril. For comparison, we calculated from M_L and $\bar{\upsilon}$ = 0.73 mL/g the cross-sectional radius of gyration for a solid cylinder

$$R_{c,cyl} = \left(\frac{M_L \bar{\upsilon}}{2\pi N_A} \right)^{1/2} \qquad (14)$$

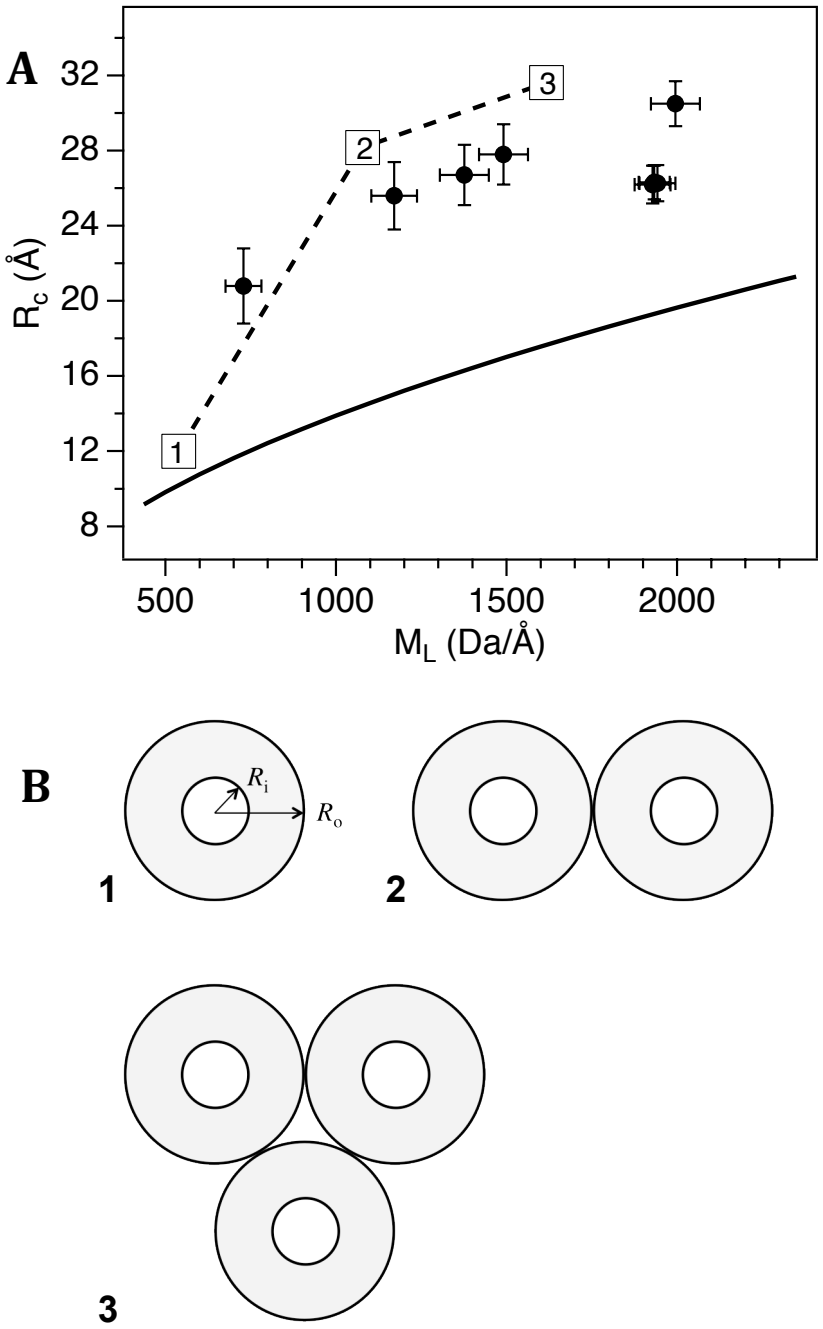


Figure 3.7. Prediction of the internal organization of $NtQ_{42}P_{10}$ fibril.
(A) A plot of $R_{c,1}$ vs. $M_{L,1}$ compared to model calculations for a solid cylinder cross-section (solid line) using Eq. 6 and the Perutz hollow cylinder b-helix cross-section for 1, 2, and 3 filaments per fibril (numbered squares) using Eqs. 7a, 7b, and 7c, respectively. The dashed line is a guide. Error bars are the standard error from the modified Guinier fits. (*B)* Schematics of the hollow cylinder models for 1, 2, and 3 filaments per fibril used to calculate R_c in (*A*). The inner and outer radii of a single filament, R_i and R_o, respectively are indicated on model 1.

We also calculated the cross-sectional radius of gyration for three hollow cylinder models representing 1, 2, and 3 filaments of the Perutz β-helix model [41] (Figure 3.7B) from the mass moment of inertia and parallel axis theorem

$$R_{c,f1} = \left(\frac{R_i^2 + R_o^2}{2} \right)^{1/2} \quad (15a)$$

$$R_{c,f2} = \left(R_i^2 + 3R_o^2 \right)^{1/2} \quad (15b)$$

$$R_{c,f3} = \left(\frac{3}{2} R_i^2 + \frac{15}{4} R_o^2 \right)^{1/2} \quad (15c)$$

In the Perutz model, each filament has an internal and external radius of R_i = 5.9 Å and R_o = 15.9 Å, respectively. Also, each filament consists of 20 glutamines per 4.75 Å b-sheet repeat cross-section [41] yielding $M_{L,f1}$ = 538.95 Da/Å with $M_{L,f2} = 2M_{L,f1}$ and $M_{L,f3} = 3M_{L,f1}$. The discrepancy observed between the experimental data and the solid cylindrical cross-section model suggests that the peptide is not optimally packed within the growing $NtQ_{42}P_{10}$ fibrils – $R_{c,1}$ is always $\sim 1.6R_{c,cyl}$. In fact, we found excellent agreement with the hollow cylinder models, where even the M_L for $NtQ_{42}P_{10}$ contains an additional 364.5 Da/Å compared to the calculated Q20 hollow cylinder M_L values that would make the two and three filament models superimpose with the experimental data. Although this does not directly prove that $NtQ_{42}P_{10}$ fibrils are β-helix structures, the model is consistent with our results. The implications for the internal fibril structure are discussed below within the context of the formed mature fibrils.

Fibril structure determined from SANS

A more complete scattering vector range ($0.004 < Q < 0.18$ Å^{-1}) was utilized to carefully probe the final structure of the $NtQ_{42}P_{10}$ mature fibrils by SANS (Figure 3.8). The fibril sample prepared at 89 mM was allowed to aggregate for more than 500 hr at 22 ^{o}C before taking measurements. The resulting scattering profile was fit with a long-cylinder form factor shape [277], which treats the single fibril as a simple cylindrical cross-section of infinite length to reveal a radius of 47 ± 1 Å. The observed scattering at $Q < 0.02$ Å^{-1} is due to attractive fibril-fibril interactions, where deviations are seen between the experimental data and the extrapolated fit.

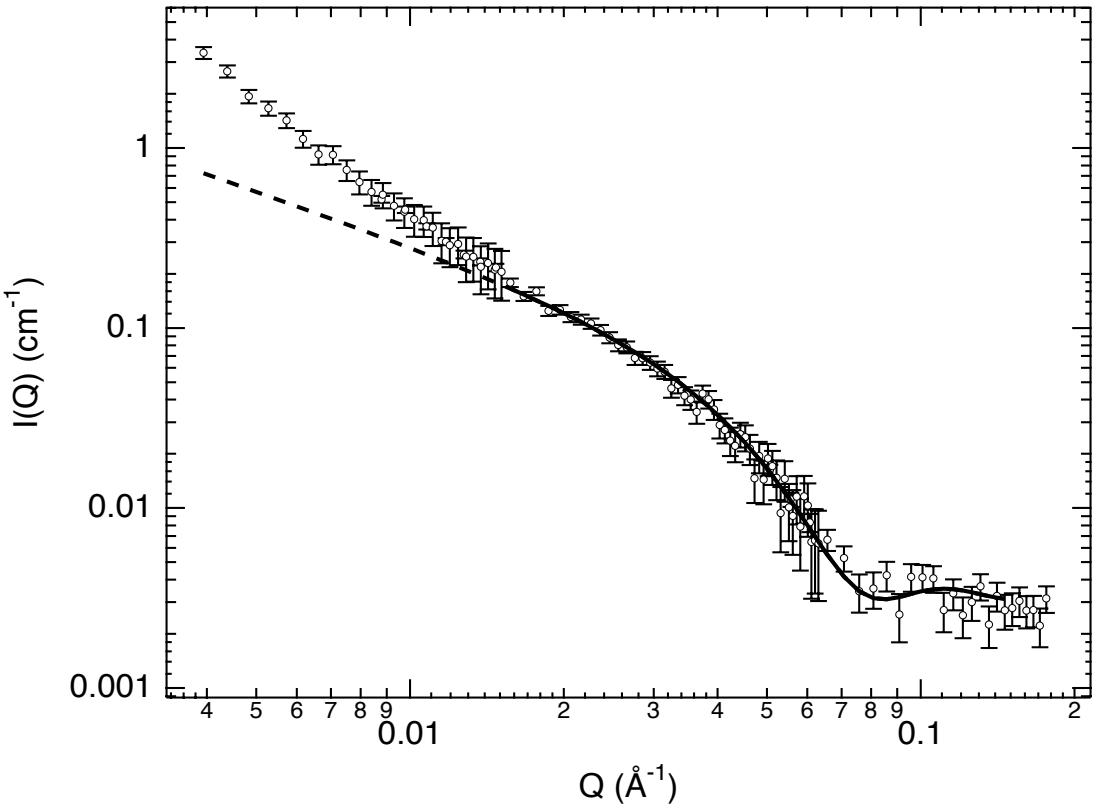


Figure 3.8 SANS data obtained over a larger *Q*-range allow the $NtQ_{42}P_{10}$ fibril structure to be more carefully probed.
Peptide concentration was 89 mM. A long-cylinder form factor fit (solid line) accounts for the SANS profile at high and intermediate *Q* values. Deviations observed between the experimental scattering data and the extrapolated fit (dashed line) at $Q < 0.02$ Å^{-1} are due to favorable fibril-fibril interactions. Error bars are standard errors from the radial average of the intensity data on the neutron area detector.

From the modified Guinier analysis, we found $I_{c,1}(0)$ = 0.00308 ± (9 × 10^{-5}) Å^{-1} cm^{-1} and $R_{c,1}$ = 33 ± 1 Å, which yielded $M_{L,1}$ = 1800 ± 50 Da/Å and $\sqrt{2}R_{c,1} = R_1$ = 47 ± 1 Å, respectively. The obtained fibril radius R_1 matches both the cylinder form factor fit radius and the average single fibril radius determined by TEM (R = 40 ± 8 Å). Also, from the modified Guinier analysis, we found $I_{c,2}(0)$ = 0.0057 ± (3 × 10^{-4}) Å^{-1} cm^{-1}, which yielded $M_{L,2}$ = 3300 ± 200 Da/Å, and $R_{c,2}$ = 77 ± 4 Å. All of these values are also comparable to those from the $NtQ_{42}P_{10}$ kinetic end-state fibrils, demonstrating reproducibility.

Measurements of both the $NtQ_{42}P_{10}$ fibril radius and mass per length provide constraints on the internal structural arrangements of the peptides within the fibril. Thus, in order to determine the number of peptides per fibril, we constructed a simple volume element to describe the fibril cross-section and compare the maximum and actual number of peptides filling this volume. With a β-sheet structure having a 4.75 Å repeat distance and the peptide chains oriented normal to the fibril direction, we calculated the maximum number of peptides per β-sheet repeat distance ($n_{L,max}$) from the fibril radius as follows

$$n_{L,max} = \frac{\pi R^2 L N_A}{\bar{\upsilon} M_{pep}} \qquad (16)$$

where L = 4.75 Å, $\bar{\upsilon}$ = 0.73 mL/g, and M_{pep} = peptide molecular mass (= 8583 Da). The actual number of peptides per 4.75 Å (n_L) was directly calculated from M_L

$$n_L = \frac{M_L L}{M_{pep}} \qquad (17)$$

Interestingly, the mature fibrils have three times less peptide than expected for a laminated β-sheet structure with $n_{L,max}$ = 3.2 ± 0.1 peptides and n_L = 1.00 ± 0.03 peptides. From the time-resolved SANS data with a saturating radius $R_1(\infty)$ = 38.5 ± 0.6 Å and mass per length $M_{L,1}(\infty)$ = 1940 ± 30 Da/Å, we found $n_{L,max}$ = 2.13 ± 0.07 and actual n_L = 1.07 ± 0.02, respectively. Again, our results are quantitatively consistent with the Perutz β-helix structural model. Since 40 glutamines are required to stabilize the β-helix structure with 2 full turns to form a filament[41], only 0.5 peptide / 4.75 Å for $NtQ_{42}P_{10}$ is needed. Thus, a single fibril of 2 to 3 filaments would have 1 to 1.5 peptides / 4.75 Å. In addition, the hollow cylinder models above have $R_{c,f2}$ = 28 Å and $R_{c,f3}$ = 32 Å, values matching the measured fibril $R_{c,1}$. Extending upon the idea of 2 to 3 filaments composing a single fibril, we noted the smallest detected species are dimers and trimers and could serve as the basic building blocks for the growing fibrils, although this could be fortuitous.

Evidence against the β-helix model for polyQ has been reported [278, 279]. However, we emphasize that the sequence context of $NtQ_{42}P_{10}$ appears to play an important role in the resulting fibril structure. Also, polymorphisms in fibril structure are a common characteristic, as demonstrated for amyloid-β[280]. Different peptide concentrations and growth conditions can result in structural variations and if any structural heterogeneity is present, SANS will give an average structure. One consideration that may help to satisfy both R_c and M_L values is that either, or both the 17 residue N-terminal and P_{10} regions of $NtQ_{42}P_{10}$ maintain extended or disordered conformations while the Q_{42} region forms the β-sheet core of the fibril. X-ray crystallography experiments on Htt exon 1 Q17 show the N-

terminal region is capable of adopting an α-helix conformation while the proline-rich C-terminus forms a polyproline helix [281]. X-ray fiber diffraction on Htt exon 1 Q51 also suggests that the flanking polyproline region is not part of the ordered fibril structure [282]. Our results on $NtQ_{42}P_{10}$ fibrils underscore the need to further investigate the effects of sequence context, polyQ length, and growth conditions on Htt fibrils. More detailed knowledge concerning the fibrils would aid in determining the aggregation pathways undertaken by monomeric and oligomeric species to construct the resulting fibrils.

3.1.4 Conclusions

With time-resolved SANS, we are capable of pinpointing and characterizing the smallest of mutant Htt species as they exist in solution at a given time during the aggregation process. These species can be distinguished from the presence of the larger aggregates that are simultaneously present. Our results provide a detailed view where aggregation of the mutant Htt exon 1 peptide, $NtQ_{42}P_{10}$, begins with compact spherical dimers and trimers coexisting with larger, quickly forming oligomer species. Monomer depletion and ThT-positive aggregate formation ensues with no observed lag phase. The dimers and trimers grow into 14-mers within 71 min, which matches the 5- to 15-mer oligomers recently observed in a HD cell model[274]. Beyond this point, the aggregates in solution become predominately fibrillar. The late-stage fibrillization process reaches a plateau in cross-sectional radius and mass per length within ~2 hr, concurrent with full monomer depletion and saturation of ThT-positive aggregate development. These formed fibrils are

amyloid-like but interestingly, the peptide is assembled within these fibrils at only about 31 to 50% of the capable packing efficiency. This packing arrangement is quantitatively consistent with the Perutz β-helix structural model as opposed to any compact β-sheet rich structural arrangements [232, 278, 279, 283].

The detailed structural information provided here for the early-stage oligomers, forming fibrils, and end-state fibrils should be of great value in more carefully discerning their respective structures and roles in the aggregation pathway. Implications for the toxicity of early oligomers to neuronal cells now can be discussed within the context of the structural details we provide. This time-resolved SANS approach to capture the structures of transient amyloid-like aggregates in solution establishes a strong method that should prove useful in the further study of aggregates pertaining to HD as well as the study of other protein aggregation processes like those involved in Alzheimer's and Parkinson's diseases.

3.2 Glutamine Expansion Length Dictates Structural Properties of Huntingtin Exon 1 Assembly

(presented as a paper submitted to Journal of Biological Chemistry)

3.2.1 Introduction

The CAG-repeat length located in the first exon of the IT15 gene is a key component in the onset of Huntington's disease (HD), an incurable neurodegenerative disorder. If longer than ~36 repeats, the translated 350-kDa huntingtin (Htt) protein is cleaved into a smaller proteolytic fragment, Htt exon1, that still contains the abnormal expanded polyglutamine (polyQ) sequence. The polyQ sequence, which confers upon the Htt exon 1 the capacity to aggregate and form amyloid-like fibrils [284], is flanked on its N-terminal side by a 17-residue amphipathic a-helix and by a proline-rich region of 50 amino acids on its C-terminal side. The critical release of Htt exon1 to initiate HD pathology is well documented [194], however, how this toxic fragment leads to neuronal cell death is still unknown.

Intensive studies have been conducted to investigate the potential role of the N-terminal 17-residues since their ability to be post-translationally modified can, in turn, modulates Htt exon 1 aggregation [191, 285]. Very recent work by Truant and collaborators showed that prevention of phosphorylation of Ser13 and Ser16 resulted in greater cell toxicity [286]. In addition, the past thirteen years have seen accumulation of evidence indicating that the presence of aggregates

does not correlate with neuronal cell death. Since other amyloid neurodegenerative diseases including Alzheimer's and Parkinson's showed the same trend [90-92], this has led the amyloid field to investigate thoroughly the aggregation process with a focus on the oligomerization phase.

Acquiring structural information on the early-formed oligomers and possibly pathogenic agents responsible for HD can be quite challenging as these structures are highly unstable and difficult to isolate. Various biophysical methods such as atomic force microscopy, analytical ultracentrifugation and confocal microscopy combined with the number and brightness method have been applied toward the structural characterization and the mechanism of assembly of these early formed oligomers [274]. Lately, using small-angle neutron scattering (SANS), a powerful tool in uncovering the nano-scale changes that an amyloid protein undergoes during aggregation [267, 287], we followed the *in vitro* fibrillization of $NtQ_{42}P_{10}$ – a peptide mimicking Htt exon 1 with the N-terminal 17-residues, a 42-glutamine repeat and a 10-proline stretch. The earliest structures monitored were a 16.1Å dimer/ trimer that grew into a β-helix fibril packed with ~1peptide per 4.75 Å β-sheet repeat distance [287].

Aggregation kinetic studies on various Htt exon 1 peptides including the ones with naked polyQ sequences have revealed a strong Q-repeat length dependence. Therefore, we questioned if the variations in the aggregation rates could in fact mirror important structural and mechanistic differences in the fibrillization pathway. To track them along with structural changes in the early oligomers, we applied time-resolved SANS combined with *ab-initio* shape

modeling experiments on the non-pathologic peptide $NtQ_{22}P_{10}$. We found that the fibrillization pathway of $NtQ_{22}P_{10}$ is different from $NtQ_{42}P_{10}$, as the initial steps require the formation of a stable and ThT positive 8-mer. We also found structural differences in the early-formed oligomers as well as in the mature fibrils. While our previous findings showed that $NtQ_{42}P_{10}$ fibrils possess a packing arrangement consistent with the Perutz β-helix structural model, we found here that this same model does not apply to $NtQ_{22}P_{10}$. Our SANS experiments reveal that normal Htt-exon1 fibrils present a better fit to sterical zipper model.

3.2.2 Experimental procedures

Preparation of the peptide – The peptide $NtQ_{22}P_{10}$ (MW=6026 Da) containing the sequence $MATLEKLMKAFESLKSFQ_{22}P_{10}K_2$ was handled as previously described [268, 287]. Briefly, once synthesized by solid-phase FMOC chemistry (Keck-Yale facility, New Haven, CT), purified by reverse phase HPLC (Biologic DuoFlow Chromatography system, Bio-Rad) and analyzed by mass spectrometry, the peptide went through a disaggregation and solubilization step. All the experiments described here were conducted at 20º C, except when mentioned, and used the following chemicals: trifluoracetic acid (TFA) (>99.5%, Thermo Scientific), 1,1,1,3,3,3-hexafluoro-2-isopropanol (HFIP) (99.5+%, Acros Organics), 10X PBS powder (Fisher Scientific), and D_2O (99.9% D, Cambridge Isotope Laboratories, Inc., Andover, MA).

$NtQ_{22}P_{10}$ was dissolved in 50:50 TFA: HFIP, incubated for 3 hrs, and then dried under Argon. Then, the peptide was solubilized into 0.05 % TFA D_2O,

ultracentrifuged for 2 hrs at 356k x *g*, 4° C. An aliquot of the supernatant was injected into an analytical reverse phase HPLC (model No 1100 series HPLC, Agilent Technologies, Santa Rosa, CA) and the peptide concentration determined by using a standard curve generated from $NtQ_{22}P_{10}$ and amino acid analysis. The aggregation reaction was then initiated by bringing the supernatant up to the desired concentration in 1X PBS D_2O and monitored by reverse phase HPLC (RP-HPLC), Thioflavine T (ThT) fluorescence and time-resolved small angle neutron scattering (TR-SANS) combined with *ab-initio* shape reconstruction experiments. In addition, the mature formed fibrils were submitted to SANS, Congo red birefringence and transmission electron microscopy (TEM).

RP-HPLC kinetics – At various time points, an aliquot of the aggregation reaction was taken, snap-frozen in liquid nitrogen to overcome the 40-min HPLC runtime and stored at -80° C. Each aliquot was then, separately thawed, spun at 356,000 x g at 4° C for 20 min and injected onto HPLC for determination of the remaining monomer concentration.

ThT Fluorescence assays – Formation of amyloid-like fibrils was followed by measuring an increase in ThT fluorescence (Perkin Elmer LS 55 fluorescence spectrometer) with excitation and emission wavelengths set at λ_{ex} = 450 nm and λ_{em} = 482 nm, respectively. The same time points as for the RP-HPLC were observed with 20 μl of the on-going aggregation reaction diluted 7.5 times in 1X PBS and mixed with 5 μl of a 2 μM ThT solution. A negative control was also run

with equivalent concentration of ThT and the obtained signal subtracted from the sample signal.

TR-SANS and SANS measurements – The protocol used for $NtQ_{42}P_{10}$ and recently published [287] was applied to $NtQ_{22}P_{10}$. Once disaggregated, the Argon-dry $NtQ_{22}P_{10}$ was dissolved in 0.05% TFA D_2O then, brought up in PBS D_2O, and centrifuged at 16,100 x g for 2 min to remove any large particulates. The fibril sample was prepared in the same manner and allowed to fully aggregate before measurement. The final peptide concentration was determined by RP-HPLC to be 0.9 mg/ml (150 mM) in PBS D_2O. Samples were placed into 2 mm pathlength rectangular-shaped quartz cuvettes (Hellma USA, Plainville, NY).

For the time-resolved measurements, samples were placed at 6.82 m from the detector and 15 min neutron exposure times were collected. To ensure sufficient signal to noise ratio at the very early stages of aggregation, longer exposure times were used to obtain the data on aggregates at 2, 4, 6 and 8 hours and were achieved by binning multiple 15 min exposures. Samples with visual sedimentation were mixed by inverting the cuvette several times to ensure sample homogeneity and then measured for a time period of 1 hr. No significant sample settling during this period was found as evaluated by comparing the total integrated intensity on the detector for consecutive 15 min exposures. For static measurements on $NtQ_{22}P_{10}$ fibrils, 0.5, 1, and 2 hrs exposure times were collected at 0.3, 6.82, and 15.4 m, respectively. Scattered neutrons were detected with a 1 x 1 m two-dimensional position sensitive detector with 192 x 192 pixels.

The experiments were performed on the CG3 BioSANS instrument at HFIR, Oak Ridge (ORNL, Oak Ridge, TN). The neutron wavelength was set to 6.08 Å with a wavelength spread $\Delta\lambda/\lambda$ = 0.14 to maximize the neutron flux. As mentioned above for the TR-SANS, sample and background intensities were collected at a detector distance of 6.82 m providing the magnitude of the Q-vector in the range of 0.007 to 0.15 $Å^{-1}$. For the static experiments, additional measurements at 0.3 m and 15.4 m were also conducted (Q-vector ranging from 0.03 to 0.7 $Å^{-1}$ and 0.003 to 0.065 $Å^{-1}$, respectively). Data reduction was carried out [IgorPro software (WaveMetrics, Lake Oswego, OR) implemented with SANS macros by Dr. Ken Littrell, ORNL] in order to correct for detector sensitivity and background scattering from quartz cell and stray neutrons. The resulting protein scattering signal was normalized by the incident beam flux and placed on absolute scale intensity using a porasil B as a calibration standard. Additional backgrounds scattering were also subtracted including the buffer and the Q-independent excess incoherent scattering signal. The data sets from the three distances were merged using the routines included with the data reduction software.

TR-SANS analysis – At the early stages of aggregation, SANS experimental data were subjected to Guinier analysis to derive the radii of gyration, Rg, and forward scattering intensity I(O) for each formed particles from the plots of Ln(I(Q) vs Q^2. Given the identified proportionality of I(0)/c to MW, the values of molecular weight for every detected intermediates were obtained. At the fibrillization stages of the reaction the longitudinal parameters of the formed fibrils exceeded the resolution limit of Q_{min} data, making the described above

standard Guinier method not suitable for the analysis. The scattering profile was multiplied by q to allow a removal of information on the length of an object. The modified Guinier analysis was applied on the resulting I(q)q vs q data yielding the radius of gyration of cross-sections (R_c) and after additional calculation, the mass per length, ML of the mature fibrils.

Three dimensional modeling – The 3-D models of the early intermediates were reconstructed *ab initio* from SANS data using a simulated annealing algorithm program DAMMIN. At least 10 independent runs of DAMMIN were performed for each observed particle. The obtained results were compared and the most probable models were found by Damsel. The program Damsup was subsequently used to align the most feasible representations and the filtered average model was yielded by Damfilt. This procedure was performed 4 to 10 times and obtained averaged shapes were pair-wise superimposed and compared using SUPCOMB to ensure a single structural solution for every presented model.

TEM – Transmission electron microscopy images were collected using a Hitachi H-600 electron microscope in the Biological Sciences Division, University of Tennessee. Aliquots with mature fibrils of $NtQ_{22}P_{10}$ grown in 1 X PBS D_2O were fixed to collodion/carbon grids and stained with a 1% uranyl acetate solution for further imaging. The diameter of a single was measured at multiple different positions on the obtained TEM image using the software program ImageJ [288] and converted from pixels to nanometers using the image scale bar.

3.2.3 Results

The use of time-resolved SANS allows the detection of sub-nanoscale structural changes that occur during protein aggregation. In order to determine the role of the polyQ-length repeat in the fibrillization process, we ensured that the $NtQ_{22}P_{10}$ experiments were conducted similarly as the ones we reported for $NtQ_{42}P_{10}$. Therefore, a 0.9 mg/ml (150 μM) solution of $NtQ_{22}P_{10}$ in 1X PBS D_2O was prepared and immediately transferred into a 2 mm pathlength-Banjo cell. The aggregation reaction was then monitored by time-resolved SANS for 55 hrs until no more changes in the sample scattering intensity could be detected. The results of the SANS measurements performed at HFIR Oak Ridge National Laboratory are shown in Figure 3.9. The observed increase and shape changing of the scattering intensity are indicative of overall aggregate structural modifications with drastic changes at 17 and 36 hrs.

$NtQ_{22}P_{10}$ does not form dimer/trimer intermediates – When exposing the $NtQ_{22}P_{10}$ sample to the neutron beam, a snapshot of the scattering profile was taken every 15 min during eight hours. This time frame in combination of a high concentration of peptide allows us to witness any early structural changes that can occur.

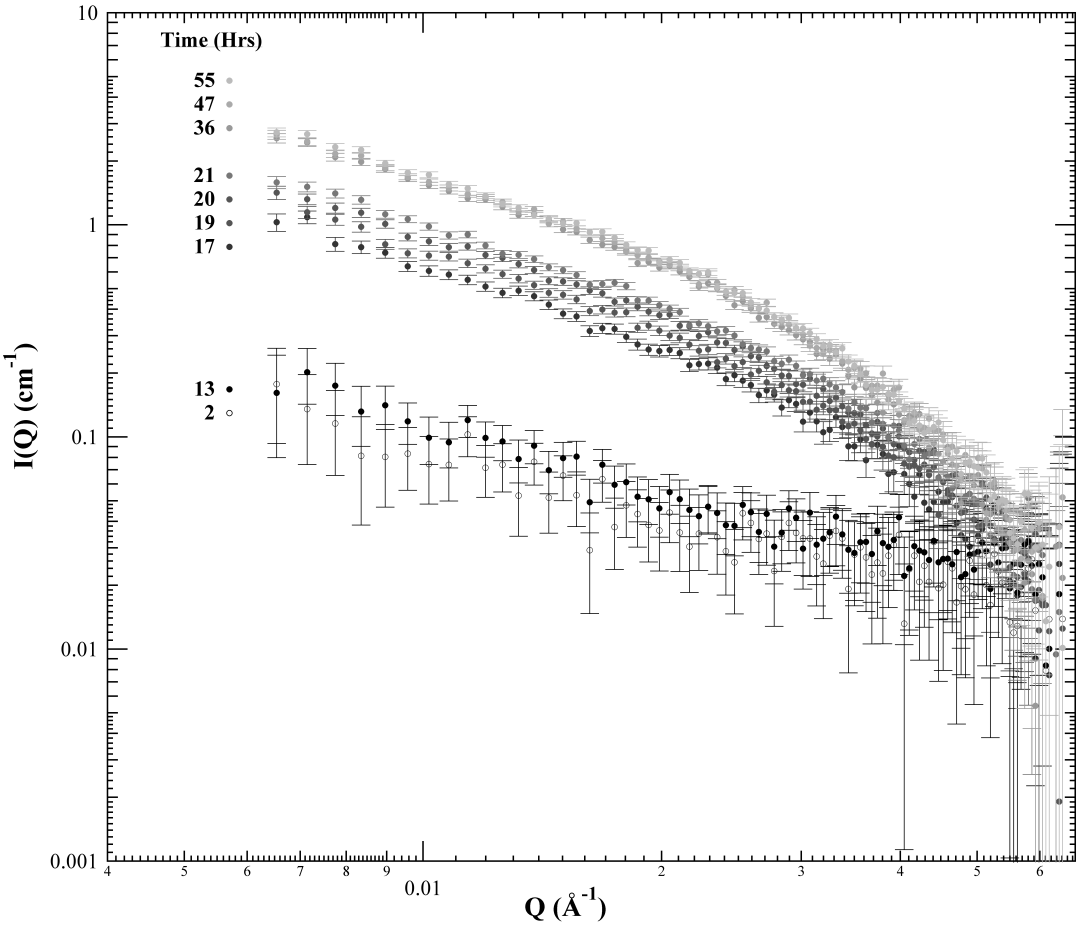


Figure 3.9. Time-resolved SANS on aggregation of $NtQ_{22}P_{10}$.
Experimental data shows the increase in the scattering intensity as $NtQ_{22}P_{10}$ (concentration 150 uM) aggregation reaction proceeds to the completion. The numbers on the graph indicate the time within at which a particular measurement was taken (grey-scale color coded). The change in the shape and slope of each curve reflect the structural development within the formed aggregates.

Once we obtained the 32 scattering profiles of $NtQ_{22}P_{10}$, we performed Guinier analysis to determine the radius of gyration, R_g, as well as the molecular weight, MW, of the early-formed intermediates. We also analyzed how the R_g and MW values were evolving other time to determine the aggregation kinetics. As expected, $NtQ_{22}P_{10}$ is aggregating much slower than $NtQ_{42}P_{10}$ (Table 3.1). The first structural changes appeared after 15 min with the formation of the first intermediates; one type at 30 min, two types at 45 min and more thereafter (Figure 3.10 A and 3.10 B). The two early types were stably present through the 8 hrs and up to 13 hrs. As the scattering signal corresponding to the smaller type was continuously present in each 15 min snapshots for 8 hrs, we improved our analysis by binning these consecutive 15 min exposures. The result was refined Rg and MW values for this first intermediate. With an Rg of 20.2 ± 6.2 Å and a MW of 48.5 ± 8 kDa (Figure 3.10 C), we found this intermediate to be made of 8.0 ± 1.4 peptides. Surprisingly, even with slower kinetics, the initial steps of $NtQ_{22}P_{10}$ fibrillization do not involve the formation of a dimer/trimer as previously observed for $NtQ_{42}P_{10}$ [287].

As 3.6 ± 0.7 % of ThT fluorescence was measured at 30 min – and only 96.9 ± 0.3 % of monomer detected – the $NtQ_{22}P_{10}$ 8-mer appears to possess an amyloid-like structure (Figure 3.11) with an R_g of dimensions similar to the $NtQ_{42}P_{10}$ trimer/tetramer (Table 3.1). The second observed species, as indicated by the second Guinier fit in Figure 3.10 B and 3.10 C, is the result of 8-mer/8-mer interactions ($R_{g,2}$ = 43.3 ± 9.9 Å and n = 13.2 ± 2.2).

Table 3. 1. Aggregation kinetics and structural parameters of $NtQ_{22}P_{10}$ and $NtQ_{42}P_{10}$ peptides obtained from Guinier analysis of the SANS profiles.

Time (min)	$NtQ_{22}P_{10}$		$NtQ_{42}P_{10}$	
	$R_{g,1}$ (Å)	N peptides	$R_{g,1}$ (Å)	N peptides
25.5			16.1 ± 5.9	2.6 ± 0.4
40.5	23.4 ± 13.4	8.0 ± 0.1	23.9 ± 5.9	3.7 ± 0.7
55.5			31.7±5.8	5.9 ± 1.1
70.5	22.5 ± 4.8	8.1 ± 1.0	46.4±4.2	14.4 ± 1.9
120	24.4 ± 6.2	7.8 ± 1.2	Regular Guinier not applicable	

Table 3. 2. Structural parameters of $NtQ_{22}P_{10}$ and $NtQ_{42}P_{10}$ peptides derived from P(r) function.

Time (min)	$NtQ_{22}P_{10}$		$NtQ_{42}P_{10}$	
	$R_{g,1}$ (Å)	N peptides	$R_{g,1}$ (Å)	N peptides
25.5			20.48 ± 0.5	2.5 ± 0.2
40.5			23.11 ± 0.7	2.9 ± 0.2
55.5	26.9 ± 2.4	8.0 ± 0.6	33.59 ± 1.6	5.5 ± 0.5
70.5			47.01 ± 2.4	13.0 ± 1.5
120			Regular Guinier not applicable	

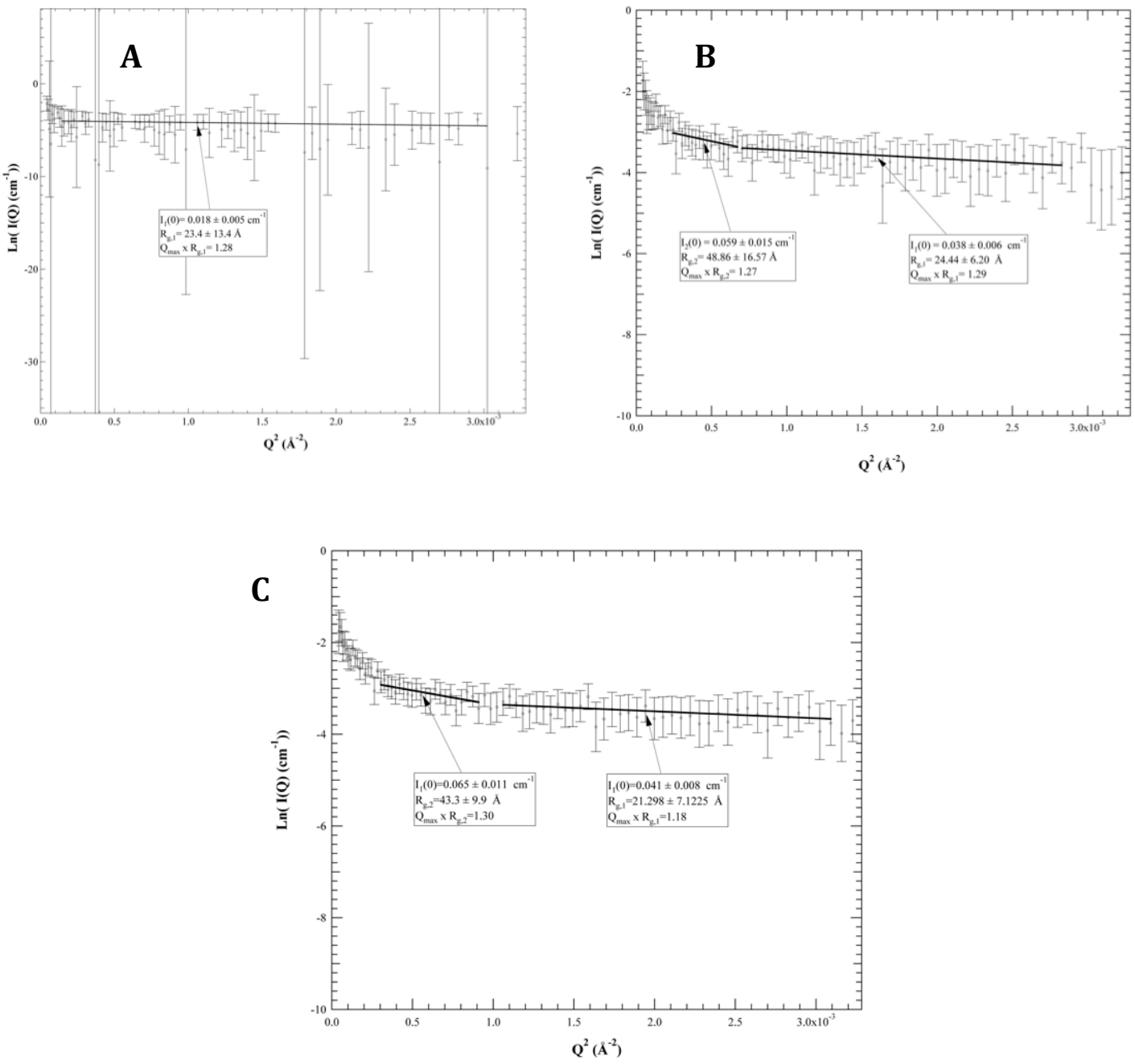


Figure 3.10 Guinier analysis on the early stages of $NtQ_{22}P_{10}$ aggregation observed by TR-SANS. A. Guinier plot for the earliest (30 min) observable aggregates shows the presence of single species with R_g of 23 Å. The large error bars are the consequence of a low signal to noise ratio. B. The representative Guinier plot for a data binned from 0 to 2 hours, showing the consistent presence of the smaller 8-meric intermediates and the emergence of the larger species, resulting from interspecies interaction. C. The occurrence of the $NtQ_{22}P_{10}$ 8-mer is exhibited within the first 8 hours of aggregation reaction.

Up to 13 hrs these two species are the only early-stage aggregates we could characterized as no additional true Guinier regime (Qmax.Rg ≤1.3) could be detected.

The GNOM-calculated Pair-distribution functions confirm the previous structural results for $NtQ_{22}P_{10}$ and $NtQ_{42}P_{10}$ early-oligomers – To investigate the role of the Q-repeat length in the morphology of $NtQ_{22/42}P_{10}$ oligomers, frequency distribution functions of the interatomic distances, P(r), were calculated for each oligomer. As an inverse Fourier transformation of I (Q) *vs.* Q, the pair distribution P(r) yields the shape and the diameter information of the scattered particle. During the mathematical transformation (GNOM, [289]), we did not include the scattering data obtained at low Q as they pertain to large particles including the second species. By doing so, we exclusively accessed to the structural details of the smallest species. Therefore, since the 8-mer is stably present up to 13 hrs, we only determined one P(r) profile for $NtQ_{22}P_{10}$ with a maximum linear dimension, r_{max}, of 86 Å (Figure 3.12 A). Figure 3.12 B shows the P(r) profiles obtained for each $NtQ_{42}P_{10}$ oligomers indicative of their structural evolution with shape change and size growth. Importantly, each pair-distribution function shows asymmetry implying that each of the various intermediates possesses an elongated shape starting at 60 Å at 25.5 min and reaching 130 Å at 70.5 min. By 70.5 min, the P(r) profile exhibits two peak maxima at ~ 30 and 80 Å suggesting that the larger species is unavoidably being detected within the analysis for this point in time.

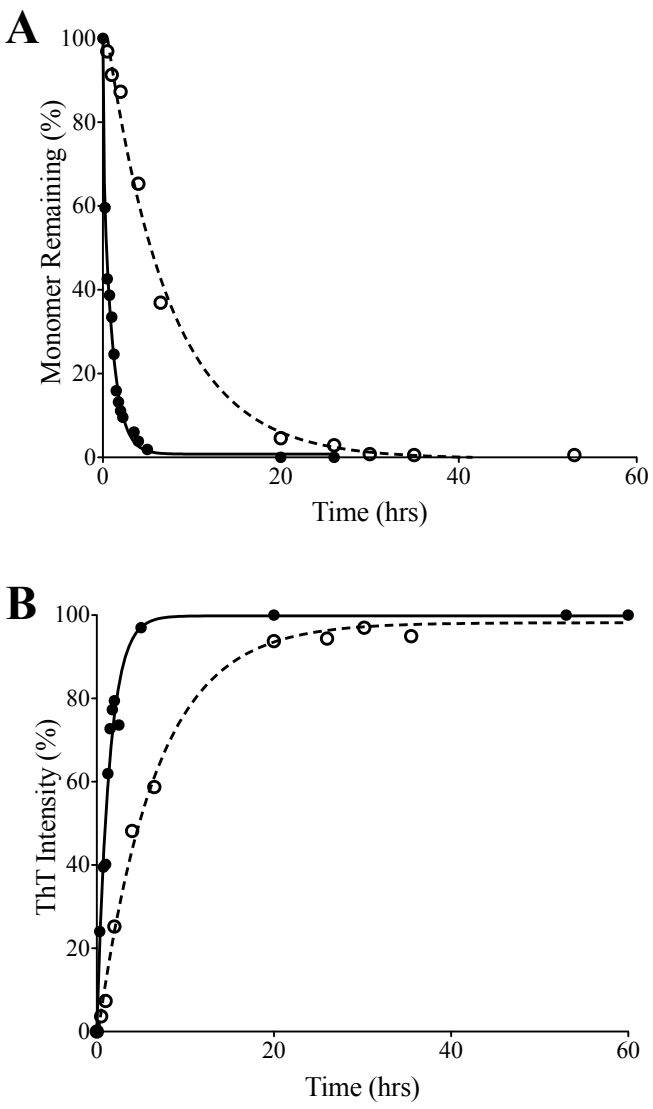


Figure 3.11 Dynamics of monomer loss and amyloid aggregate formation for normal and pathological-like htt exon-1 peptides. A. Aggregation reaction observed by HPLC sedimentation assays, monitoring a disappearance of -o- 130 uM $NtQ_{22}P_{10}$ and -•- 93 uM $NtQ_{42}P_{10}$ monomers. B. The binding of ThT dye to the formed amyloid aggregates (concentrations as indicated above) followed by the increase of fluorescence. At 30 min time point $NtQ_{22}P_{10}$ species demonstrate ~ 4% fluorescence, signifying the amyloid-like characteristics of the earlier intermediates. Both, (A) the monomer disappearance and (B) ThT fluorescence intensity increase were fit by single exponential functions.

In Table 3.2, we reported the R_g values and numbers of peptides derived from the P(r) profiles. When comparing with Table 1, these values are similar, strengthening our previous conclusions.

Ab initio experiments allow shape reconstruction of $NtQ_{22}P_{10}$ and $NtQ_{42}P_{10}$ early-oligomers – Different shapes of htt-exon 1 observed by AFM have been previously reported including spherical oligomers with diameters of 4 to 5 nm [264]. In order to determine the effect of the Q-repeat length toward the shape of the observed intermediates, we conducted *ab initio* shape reconstruction experiments using the P(r) output data obtained for each intermediate into DAMMIN (simulated annealing algorithm program, [290]). Figure 3.13 represents the resulting structure for each intermediate. As anticipated by our previous results, the *ab initio* models for $NtQ_{42}P_{10}$ early species show elongated shapes. The dimer/trimer, which measures ~33 x 33 x 58 Å, resembles the "spherical oligomers" detected by AFM [264]. At 40.5 min, the trimer/tetramers, revealed originally by Guinier fits, have now been refined as a trimer with only an increased length of ~ 65 Å. The last possible shape reconstruction that we determined was for 55.5 min as no unique P(r) existed for 70.5 min. Remarkably, the $NtQ_{42}P_{10}$ 6-mer demonstrates not only a longer and larger shape but also a twisted silhouette, that is generally visible by AFM and EM on mature fibrils [280]. In the case of the $NtQ_{22}P_{10}$ 8-mer (Figure 3.13B), this stable oligomer not only possesses an elongated shape but also is more dense then the NtQ42P10 6-mer as indicated by its measurements (~ 43 x 43 X 86 Å).

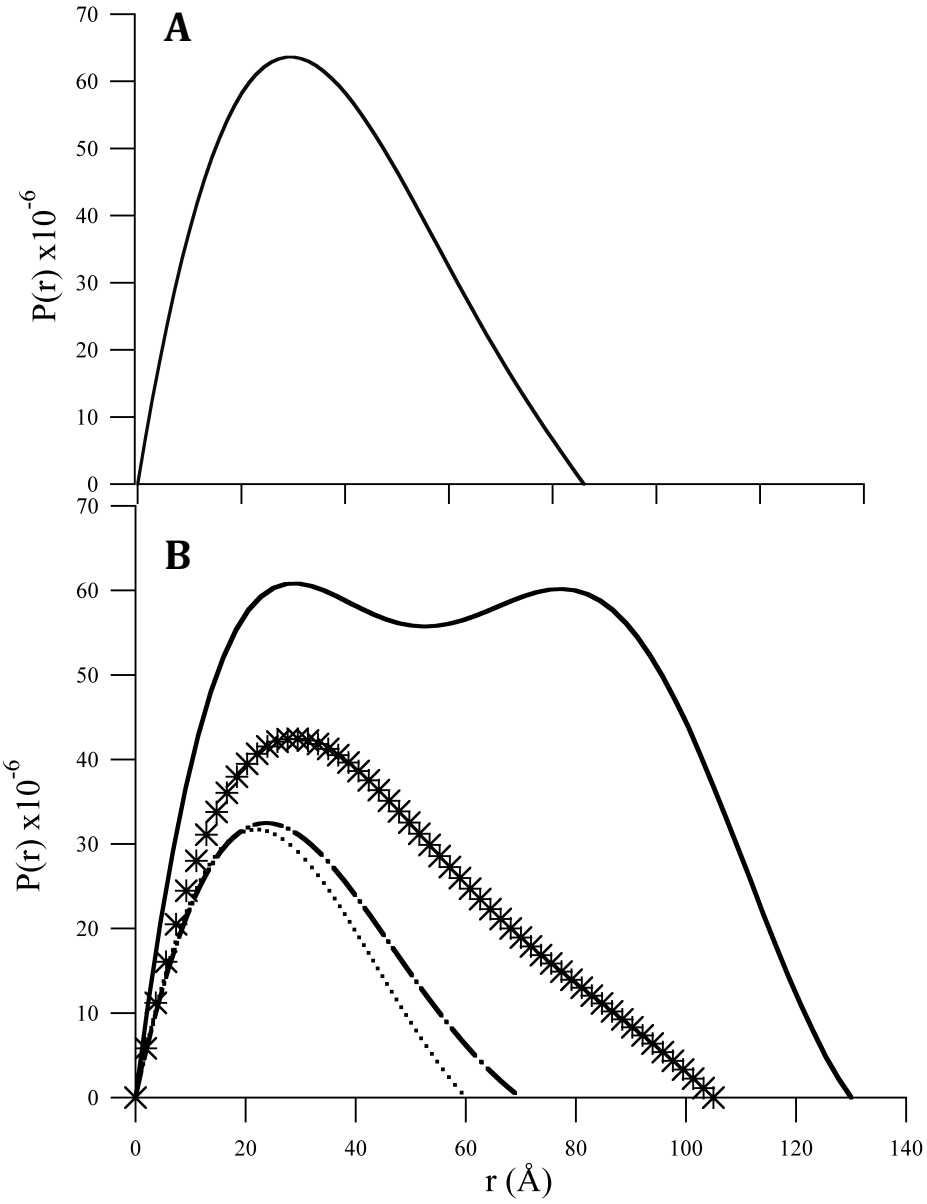


Figure 3.12 P(r) analysis of scattering data on the early $NtQ_{22}P_{10}$ and $Nt_{42}P_{10}$ intermediates.

A. P(r) functions of $NtQ_{22}P_{10}$ 8-mer, present within first 13 hours of reaction, indicating the shape and maximum dimensions of the observed particle. B. Structural evolution of $NtQ_{42}P_{10}$ aggregates represented by their P(r) graphs for 25.5 min (·····), 40.5 min (—·—), 55.5 min (✳✳✳✳) and 70.5 min (——). The structure obtained for aggregates observed at 70.5 min shows the presence (association) of two different species.

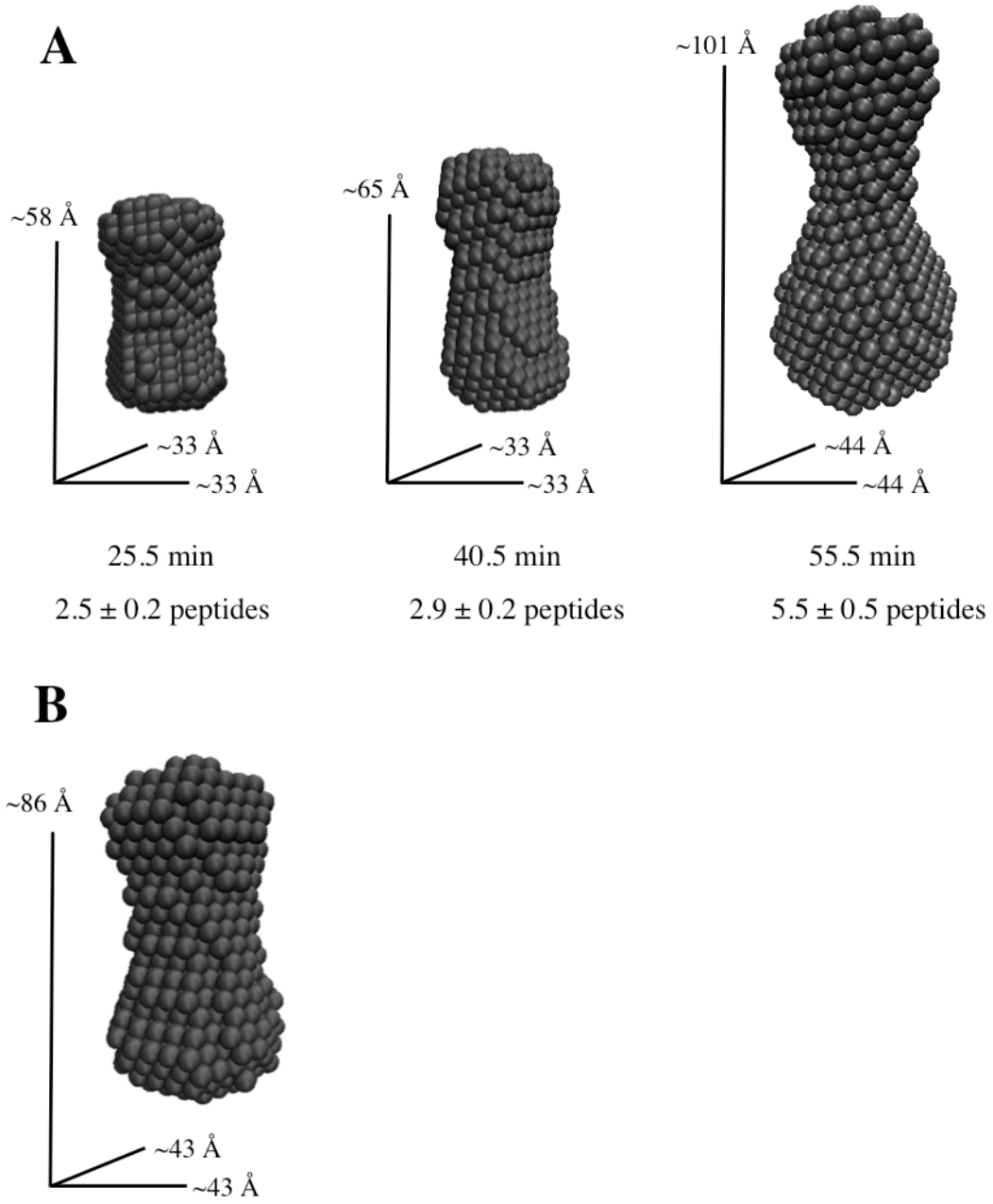


Figure 3.13 3-D models of the early intermediates formed at the initial stages of $NtQ_{42}P_{10}$ and $NtQ_{22}P_{10}$aggregation. A. Early aggregates formed by $NtQ_{42}P_{10}$ at 25.5 and 40.5 min show an elongated shape with dimensions of 58 and 65 Å, respectively. By the time of 50.5 min, the predominant structure of aggregates formed by pathological htt exon1 peptide demonstrate a twisted morphology, shown for amyloid protofibrils. B. The dummy atom model for the earliest $NtQ_{22}P_{10}$ precursors consistently present during the first 13 hours of aggregation reaction. The reconstructed structure shows anisotropy within its shape with 86 Å length of longer and 43 Å length of shorter axis.

Although our Guinier analyses were a good indicator of the size and weight of the various early oligomers, the *ab initio* shape reconstruction experiments enabled us to tune more precisely their dimensions. This said, it is important to note that the reconstructed figures presented here do not demonstrate a unique solution but rather highlight the most predominant conformations within a given stage of the aggregation reaction.

$NtQ_{22}P_{10}$ fibrils demonstrate distinctive structural features. At 17 hrs into the aggregation reaction, the $NtQ_{22}P_{10}$ 8-mer intermediate progressively grows in mass and size anisotropy toward a more fibrillar structure. As the longitudinal parameter of the growing fibrils exceed the resolution limit that can be detected by SANS, only the two-dimensional fibril cross section can be monitored. The implication is that conventional Guinier analysis is no longer valid to treat the data, and modified Guinier analysis has to be applied. By doing so, we obtained the following information including the radius of gyration of cross-section, Rc, and the mass per length, M_L. Figure 3.14 A shows the continuous growth of the 8-mer toward a fibril shape with increasing values of Rc and n peptides/4.75 Å – the latter being proportional to M_L with n = (M_L x 4.75)/ $MW_{NtQ22P10}$. Starting at 17 hrs with an Rc of 23.32 ± 1.17 Å and a number of peptides of 1.9 ± 0.15 per 4.75 Å, the short fibrils pursue their size augmentation for another 29 hrs where they finally reach a stable size and conformation (Rc = 33.24 ± 1.12 Å and n = 5.2 ± 0.4 peptides per 4.75 Å) – the dimensions of the final fibrils being of 36.25 ± 0.46 Å with 5.3 ± 0.4 peptides/4.75 Å.

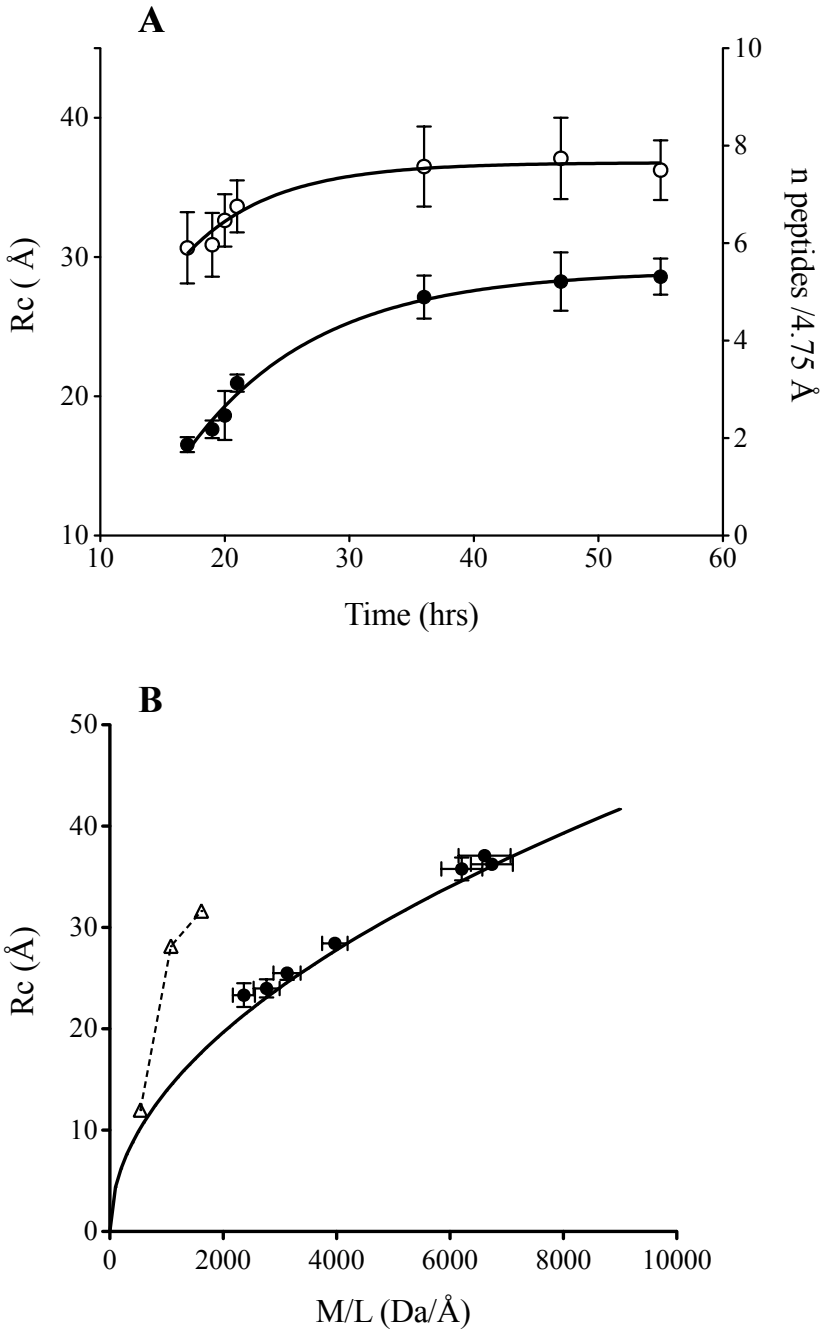


Figure 3.14 SANS data for $NtQ_{22}P_{10}$ fibril growth. A. The kinetics of formation of mature fibrils followed by growth of radius of gyration of cross-section (R_c)(left y-axis) and the increase in the amount of peptides present in 4.75 Å of the fibrils cross-section (right y-axis.). Mature fibrils demonstrate the packing of 5.3 ± 0.4 peptides/4.75 Å and Rc = 36.25 ± 0.46 Å by the completion of aggregation reaction, at time= 55hrs. B. The fibrillization pathway of $NtQ_{22}P_{10}$ observed in terms of the increase in Rc as a function of M_L. The growth of $NtQ_{22}P_{10}$ fibrils follows a solid cylinder model (solid line) and significantly diverges from the association of two-three hollow filaments model (dotted line), proposed for $NtQ_{42}P_{10}$.

Once the various sizing parameters were characterized, we then investigated the possible structural organization of the $NtQ_{22}P_{10}$ fibrils i.e. assessed if $NtQ_{22}P_{10}$ packing followed the Perutz model with formation of a hollow cylinder as we previously reported for the pathological Q length, $NtQ_{42}P_{10}$.

By plotting Rc *vs.* M_L, we found that our experimental data superimposed perfectly with those calculated for a solid cylinder model and not for a fibrils, consisting of hollow filaments (Figure 3.14 B) (see ref. [287] for model equations).

3.2.4 Conclusion

The importance of elucidating the aggregation pathway of Htt-exon1 has been recognized for some time, with a special emphasis toward the oligomerization steps as the early-formed Htt-exon1 species are the primary suspects for causing neuronal cell death in HD [265, 275]. Thus, thorough structural and mechanical knowledge of the process prior to amyloid fibrillization is critical – including nucleus formation, as the nucleus is the fibrillization trigger. So far, little experimental data regarding the structural changes of Htt-exon1 before nucleation can be found in the literature.

By combining time-resolved SANS with *ab initio* shape reconstruction we were able to not only follow the Htt-exon1 assembly pathway but also to obtain the 3D-shape envelope of the early-formed oligomers. Indeed, refining the shape of the Htt-exon1 early intermediates revealed that the $NtQ_{42}P_{10}$ oligomers, including the dimer/trimer, possess an elongated form that has been previously reported as spherical by lower resolution techniques [226, 265]. Even our initial

characterization of the $NtQ_{42}P_{10}$ dimer/trimer based on *R*g and *M* obtained from Guinier analysis of the SANS data suggested this oligomer was mostly spherical [287]. The increased resolution from *ab initio* shape determination has permitted us to better visualize the anisotropy of this oligomer and subsequent early-stage oligomers formed. As described above, the individual *ab initio* structures shown in Figure 5 may not be unique but they certainly demonstrate the most predominant structural features that are present.

Identifying the conformation that is responsible for Htt-exon1 toxicity was possible by comparing the aggregation properties of $NtQ_{42}P_{10}$ and a non-pathogenic Htt-exon1, $NtQ_{22}P_{10}$. Clearly, the role played by the polyQ length is quite substantial as $NtQ_{22}P_{10}$ and $NtQ_{42}P_{10}$ pack and assemble differently. Diminishing the pathologic fragment by 20 glutamines results in conversion of the final fibril internal structure from a hollow β-helix into the more commonly seen amyloid structure type – laminated β-sheets [283]. This decrease in Q-length is also responsible for drastically affecting the assembly pathway. The first events in the oligomerization process of $NtQ_{22}P_{10}$ do not include the formation of a dimer as we recently reported for the pathological length. Instead, the $NtQ_{22}P_{10}$ monomer assembles into a long-lived 8-mer. Indeed, importantly, for the first 15 min. of the $NtQ_{22}P_{10}$ aggregation reaction, no other species, aside from the monomer, could be detected by SANS followed subsequently by the the sole appearance of the 8-mer (n=8 ±1). This suggests that this 8-mer is the critical nucleus required for further aggregation of $NtQ_{22}P_{10}$. A structural characterization of the fibrillization nucleus can be difficult as it is the thermodynamically least-favorable species at low

concentrations. Powers and Powers have shown that it becomes thermodynamically stable and accumulates above supercritical concentration [238]. The nucleus is therefore no longer a thermodynamic nucleus but a structural nucleus. Additionally, Wetzel's team has very recently found that the size of the critical nucleus for naked polyQ peptides is also Q-repeat length dependent, increasing with decreasing Q-repeat length [291]. Our results seem to indicate that we have identified and determined the dimensions of the $NtQ_{22}P_{10}$ critical nucleus by applying time-resolved SANS during the early stages of fibrillization. However, additional experiments are needed to confirm that the formed $NtQ_{22}P_{10}$ 8-mer is the true critical nucleus.

CHAPTER 4.

Conclusions and suggestions for the future work.

4.1 $NtQ_{22}P_{10}$ and $NtQ_{42}P_{10}$: similar peptides, but different mechanisms of aggregation

During the process of our current investigations, very few of the observed experimental outcomes were predictable, based on our experience of working with polyQ peptides [224]. For example, it was anticipated that the kinetic rates of aggregation of normal and pathological-like peptides would show a repeat-length dependence, as observed for model sequences and disease proteins from other polyglutamine extension diseases [96, 292]. However, the effect of the length of polyQ region on the structural dynamics of htt exon-1 aggregation remained unexplored until the onset of our investigation and kept us eager to conduct detailed characterization. A great amount of surprising results were waiting for us upon completion of our study. We found that not only are there distinctive cross-sectional organizations of the mature fibrils but also entirely dissimilar mechanism of aggregation as a result of differences in the length of the polyQ repeat in pathological and normal-like htt exon-1 peptides. The obtained information represents the first unique insight on the role of the polyglutamine length in the mechanism of aggregation.

The initial studies on the aggregation in Huntingtin disease were concentrated on the simple (naked) polyglutamine sequences [224, 293-295]. The

reasoning behind this choice came from the fact that in several polyglutamine-expansion diseases the presence of an increased number of glutamine repeats was consistently associated with a disorder. The challenges in producing long synthetic peptides also contributed to the decision on the model system of study. It was found that $K_2Q_{37}K_2$ peptide demonstrated a nucleated growth polymerization mechanism with a critical nucleus of 1 and it was proposed that all polyQ peptides would show the analogous behavior [284, 294].

Investigations on the nature of toxicity-causing proteins in polyglutamine diseases have shown that the varying length of glutamine repeats correlates with the induction of pathology, depending on the particular disorder. It was also noticed that these proteins affected different regions of the brain and type of tissues, implying that the amino acid sequences flanking the polyQ region might play an important role in the aggregation and therefore toxicity in the disorders. This suggestion raised the issue of whether studying the behavior of naked polyglutamine peptides would have any relevance in HD pathobiology. Studies on the polyQ peptides containing a polyproline sequence at the C-terminus demonstrated that the presence of this region slows aggregation kinetics but does not have any altering effect on the nucleation polymerization mechanism with the monomeric nucleus [225]. Further experiments demonstrated that the N-terminal 17 amino acids of huntingtin placed next to the polyQ sequence radically modify the mechanism of aggregation and increase the rate of aggregates' formation [226]. Thakur et al performed the kinetic studies on the aggregation of $NtQ_{30}P_{10}$ peptide and suggested that it follows the aggregation reaction that is initially non-

nucleated, however, later, amyloid nuclei appears to form stochastically from some early prefibrillar aggregates. Although this statement is potent in giving a rise to multiple interpretations, the difference in the aggregation mechanisms between the simple polyQ and htt exon-1 fragment appears clear. The research study described in this work has provided an insight into the polymerization pathways for the normal and pathological-like htt exon-1 peptides: while the former seems to follow the classical nucleation, the later has a higher propensity to associate through downhill aggregation mechanism. The initial $NtQ_{22}P_{10}$ peptide loss observed by the reverse HPLC assay can be explained by the fragments being incorporated into the forming nucleus and some interspecies association, as observed by our SANS experiments. The positive ThT measurements also suggest that the proposed nucleus might be amyloid-like, further supporting a hypothesis on the nucleation driven mechanism for this type of peptide. On the other hand, all performed experiments on the aggregation of $NtQ_{42}P_{10}$ agree that its behavior is governed by the downhill polymerization. Here we show, for the first time, that the length of polyQ region not only affects the kinetic rates of aggregation but also may be important in the determination of the polymerization mechanism. These findings provide some explanation for the existence of the repeat-length threshold of 37 glutamines for the toxic htt fragment. The further question to be answered, though, is what is the relation between the polyQ length effects on mechanism of aggregation and repeat-length effects on the likelihood of development of Huntington's Disease. Recently Kar et al indicated, that in their HPLC and DLS experiments on the simple polyQ peptides, containing 23 to 26 glutamines, all of

the fragments followed the nucleation-dependent mechanism [291]. However, the surprising finding was that the size of the nucleus was varying with a different length of polyQ: while Q23 peptide aggregation proceeded through tetrameric nucleus, Q24 formed a nucleus of 3 and the size of the nucleus got reduced to n=1 when the length of polyQ reached 26 residues [291]. This observation may not be in contradiction to our own finding. Indeed, it is possible that for the more complex htt exon-1 sequence, changes in the critical nucleus mediated by the length of the polyglutamine repeat results in a change of the entire mechanism of aggregation due to modified inter-peptide interactions from the additional N and C-terminal residues. Clearly many further studies involving concentration dependence and seeding experiments are necessary and planned by our laboratory to provide additional evidence for different polymerization pathways of normal and pathological-like peptides.

4.2 Selected methodology to study htt exon1 aggregation

When one has to deal with an amyloid system, inherently demonstrating a high degree of polymorphism, a particular consideration needs to be paid in using the same parameters and protocols within the entire experimental procedure. The previous structural work on various amyloid peptides and proteins has found that although the formed aggregates show generally similar morphology, the final products of fibrillization can be highly polymorphic: the same amyloidogenic protein or peptide can yield multiple amyloid forms depending on the terms of

induction [35, 280, 296]. Nekooki et al demonstrated that htt exon-1 fragments, containing 42 and 62 glutamines misfold into significantly different conformations, when allowed to aggregate at different temperatures[297]. They also found that varying levels of toxicity associated with the formation of aggregates occurring at 4^0C and 37^0C. For Aβ aggregation, Tycko's team showed that the predominant morphology can be affected by even slight changes in the growth conditions, such as presence or absence of agitation [296]. Since $NtQ_{42}P_{10}$ is prone to aggregate on the minute scale, comparing to the hours-long aggregation for $NtQ_{22}P_{10}$, a reader might suggest running the described time-resolved experiments at significantly lower temperatures for pathological-like peptide. This step would allow slowing down the reaction and potentially could provide more details on the dynamics of the experiment. However, because of the reasons stated above, the same purification and solubilization methods, similar concentrations of 100-150 uM, as well as 20^0C temperature and no agitation conditions were equally utilized when studying aggregation of $NtQ_{22}P_{10}$ and $NtQ_{42}P_{10}$ peptides to ensure experimental consistency in obtaining the presented results.

While working with peptides at the mg/mL range, as required by SANS measurements, we were aware of a potential concern arising from utilization of samples at high concentrations. There is still an open question if a concentration can specify mechanisms of aggregation and structures of the formed intermediates. Although we are comforted with the fact that our data does match the results obtained from AFM and TEM studies performed at lower concentrations[265, 274, 275], a detailed investigation on the structural

consistency within the aggregation pathways occurring at different concentrations still needs to be performed. Our preliminary results (data not presented) have shown that we found the right experimental conditions for running time-resolved SANS at much lower sample concentrations. Performing experiments with the cuvettes ranging from 1 to 5 mm path-lengths would permit the study of structures formed at different concentrations covering an order of magnitude: from 100 down to 10 uM. Identifying the degree of variability of analogous structures formed at different sample concentrations would further boost our confidence in the proposed structural pathways for wild-type and toxic htt exon-1 fragments.

4.3 Future directions

4.3.1 X-ray fiber diffraction studies

The differences in the mechanism of aggregation between the two peptides might provide the major effect on the organization of the internal structure of the mature fibrils, as was observed by SANS. We have demonstrated that this technique represents a powerful tool to study amyloid fibrils structure; and we found that the pathological like-fibrils represent an association of 2-3 hollow filaments, closely resembling Perutz β-helix. At the same time, the fibrils formed by the peptide containing the normal length of polyQ region, follow the single solid cylinder model. In order to provide a further support to our hypotheses, X-ray

fiber diffraction experiments need to be performed. Entirely different X-ray diffraction signals are expected to arise from the different structures predicted by our SANS studies. The classical cross-β diffraction pattern, expected for $NtQ_{22}P_{10}$ fibrils would show a sharp meridional reflection at 4.75 Å and a equatorial signal at 10 Å, representing the length of β-sheet repeat and intra-β-sheet spacing, respectively [298]. The diffraction signal obtained from the hollow $NtQ_{42}P_{10}$ fibrils, though, would differ significantly in the absence of the 10 Å peak, since no β-slab-like conformation is present within the structure of the fibrils.

Also, x-ray fiber diffraction would be advantageous to check the effects of various solution conditions, e.g. reaction temperature and agitation, as mentioned earlier.

4.3.2 Selective deuteration/contrast matching studies

The next step in characterization of the htt amyloid fibrils will include mapping out the contacts between the fragments of htt exon-1 peptide, assigning to the formation of the core (if present) to one of these fragments and providing the details on the structures of these domains within the fibrils. To reach these goals, SANS contrast matching experiments should be performed on both $NtQ_{22}P_{10}$ and NtQ42 selectively deuterated peptides. We have provided the evidence in support of SANS as a valuable technique to study protein aggregation. However SANS utilization is not limited to the general characterization of the fibrils, such as

an overall β-helical or cross-β conformation. SANS, possessing a high sensitivity to the contrast between hydrogen and deuterium isotopes, can also provide information on the distribution of the flanking htt region within the structure of the mature fibrils. Using the recent advances in the production of deuterated FMOC-protected residues it becomes possible to synthesize htt exon-1 peptides, containing the particular regions, such as polyQ or polyPro, being selectively deuterated. Through this approach, the obtained peptides can be treated as multicomponent systems, where conformation and position of each domain can be resolved. During the experimental procedure, the level of deuteration of the solvent can be tuned (by varying the ratio between hydrogen and deuterium) to the level of deuteration of one or more components of the peptide. This allows one "to blank out" the particular portion of the fragment and separate the component of interest, which has a different scattering intensity. This scattering profile yields the position of the selected region as well as its conformational characteristics within the formed fibrils. By performing the series of the experiments involving the htt exon-1 peptide with different domains selectively deuterated, the orientation of the regions relative to each other as well as their contribution to the overall structure of both growing aggregates (in a kinetic experiment) and final fibrils can be assessed. With this particular information in hand, the question on the internal organization of the huntingtin fibrils can be also answered. Not only can the presence of the solid or open core can be detected, but the region of the fragment responsible for the formation of this particular structure can be readily determined.

4.3.3 Transition to a full-length htt exon-1 fragment

The $NtQ_{22}P_{10}$ and $NtQ_{42}P_{10}$ peptides utilized for a current study include the sequences representing the N-terminal exon 1 fragment of huntingtin protein. Containing both 17 amino acids and a Pro-10 region, these constructs reliably reflect the protein context of the toxic fragments, however they do not represent the entire fragment, resulting from the proteolytic cleavage occurring in the cell. Current advances in solid-phase FMOC chemical synthesis allow the production of the longer amino acid stretch constituting the truncated htt peptide, containing an additional polyproline region of 27 residues. Although we demonstrated that the length of polyglutamine repeat has a significant effect on the conformational dynamics and mechanism of htt-exon 1 aggregation; application of our novel structural methods on the full-length fragment will provide the most biologically relevant insight in the process of htt aggregates' formation.

LIST OF REFERENCES

1. Berson, J., et al., *Proprotein convertase cleavage liberates a fibrillogenic fragment of a resident glycoprotein to initiate melanosome biogenesis.* J. Cell Biol, 2003. **161**: p. 521-533.
2. Claessen, D., et al., *A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in Streptomyces coelicolor by forming amyloid-like fibrils* Genes Dev, 2003. **17**: p. 1714-1726.
3. Iconomidou, V., G. Vriend, and S. Hamodrakas, *Amyloids protect the silkmoth oocyte and embryo.* FEBS Lett., 2000. **479**: p. 141-145.
4. Chiti, F. and C. Dobson, *Protein Misfolding, Functional Amyloid and Human Disease.* Annu. Rev. Biochem, 2006. **75**: p. 333-366.
5. Booth, D., et al., *Instability, unfolding and aggregation of human lysozyme variant underlying amyloid fibrillogenesis.* Nature, 1997(385): p. 787-793.
6. Raffen, R., et al., *Physicochemical consequences of amino acid variations that contribute to fibril formation by a immunoglobulin light chains.* Protein Sci, 1999(8): p. 509-517.
7. Sekijima, Y., et al., *The biological and chemical basis for tissue-selective amyloid disease.* Cell, 2005(121): p. 73-85.
8. Stathopulos, P., et al., *Cu/Zn superoxide dismutase mutants associatyed with amyotrophic lateral sclerosis show enhanced formation of aggregates in vitro.* Proc. Natl. Acad. Sci., 2003(100): p. 7021-7026.
9. Lansbury, P. and H. Lashuel, *A century-old debate on protein aggregation and neurodegeneration enters the clinic.* Nature, 2006(443): p. 774-779.
10. Klein, J. and P. Dhurjati, *Protein aggregation kinetics in an Escherichia coli strain overexpressing a Salmonella typhimurium CheY mutant gene.* Appl. Environ. Microbiol., 1995(61): p. 1220-1225.
11. Campioni, S., E. Monsellier, and F. Chiti, *Why proteins misfoid.* Protein misfolding diseases, 2010: p. 3-20.
12. Zhang, H., et al., *The carbonate radical anion-induced covalent aggregation of human copper, zink superoxide dismutase, and alpha-synuclein: intermediacy of tryptophan- and tyrosine-derived oxidation products.* Free Radic. Biol. Med., 2004(36): p. 1355-1365.
13. Kheterpal, I., *Abeta protofibrils possess a stable ore structure resistant to hydrogen exchange.* Biochemistry, 2003(42): p. 14092-14098.
14. Serpell, L., et al., *The protofilament substructure of amyloid fibrils.* J. Mol. Biol., 2000. **300**: p. 1033-1039.
15. Kirschner, D., et al., *Synthetic peptide homologous to beta protein from Alzheimer disease forms amyloid-like fibrils in vitro.* Proc Natl Acad Sci, 1987. **84**: p. 6953-6957.
16. Nilsson, M., *Techniques to study amyloid fibrils formation in vitro.* Methods, 2004. **34**: p. 151-160.
17. Bolder, S., et al., *Thioflavin T and birefringence assays to determine the conversion of proteins into fibrils.* Langmuir, 2007. **23**: p. 4144–4147.
18. LeVine, H.I., *Thioflavine T interaction with synthetic Alzheimer's disease beta-amyloid peptides: detection of amyloid aggregation in solution.* Protein Sci, 1993. **2**: p. 404–410.

19. Hawe, A., M. Sutter, and W. Jiskoot, *Extrinsic fluorescent dyes as tools for protein characterization.* Pharm Res, 2008. **25**: p. 1487–1499.
20. Krebs, M., E. Bromley, and A. Donald, *The binding of thioflavin-T to amyloid fibrils: localisation and implications.* J Struct Biol, 2005. **149**: p. 30-37.
21. Biancalana, M., et al., *Molecular mechanism of thioflavin-T binding to the surface of beta-rich peptide self-assemblies.* J Mol Biol, 2008. **385**: p. 1052–1063.
22. Stsiapura, V., et al., *Computational study of thioflavin T torsional relaxation in the excited state.* J Phys Chem A, 2007. **111**: p. 4829–4835.
23. Inouye, H., P. Fraser, and D. Kirschner, *Structure of beta-crystallite assemblies formed by Alzheimer beta-amyloid protein analogues: analysis by x-ray diffraction.* Biophys J, 1993. **64**: p. 502-519.
24. Kirschner, D., C. Abraham, and D. Selkoe, *X-ray diffraction from intraneuronal paired helical filaments and extraneuronal amyloid fibers in Alzheimer disease indicates cross-beta conformation.* Proc Natl Acad Sci, 1986. **83**: p. 503-507.
25. Terry, C., A. Damas, and P. Oliveira, *Structure of Met30 variant of transthyretin and its amyloidogenic implications.* EMBO J, 1993. **12**: p. 735-741.
26. Jimenez, J., et al., *Cryo-electron microscopy structure of an SH3 amyloid fibrils and model of the molecular packing.* EMBO J., 1999(18): p. 815-821.
27. Smith, D., et al., *A systematic investigation into the effect of protein destabilisation on beta 2-microglobulin amyloid formation.* J Mol Biol, 2003. **330**: p. 943-954.
28. Petkova, A., et al., *A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR.* Proc. Natl. Acad. Sci, 2002(102): p. 315-320.
29. Jaroniec, C., et al., *Molecular conformation of a peptide fragment of transthyretin in an amyloid fibril.* Proc. Natl. Acad. Sci., 2002(99): p. 16748-16753.
30. Ritter, C., et al., *Correlation of structural elements and infectivity of the HET-s prion.* Nature, 2005(435): p. 844-849.
31. Makin, O., et al., *Molecular basis for amyloid fibril formation and stability.* Proc. Natl. Acad. Sci., 2005(102): p. 315-320.
32. Nelson, R., et al., *Structure of the cross-β spine of amyloid-like fibrils.* Nature, 2005(435): p. 773-778.
33. Balbach, J., et al., *Supramolecular structure in full-length Alzheimer's beta-amyloid fibrils: evidence for a parallel beta-sheet organization from solid-state nuclear magnetic resonance.* Biophys. J., 2002(83): p. 1205-1216.
34. Antzutkin, O., et al., *Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of β-sheets in Alzheimer's β-amyloid fibrils.* Proc. Natl. Acad. Sci., 2000(97): p. 13045-12050.
35. Tycko, R., *Molecular structure of amyloid fibrils: insights from solid-state NMR.* Q. Rev. Biohys. , 2006(39): p. 1-55.
36. Goldsbury, C., et al., *Studies on the in Vitro Assembly of Aβ 1–40: Implications for the Search for Aβ Fibril Formation Inhibitors.* J. Struct. Biol., 2000(130): p. 217-231.

37. Torok, M., et al., *Structural and dynamic features of Alzheimer's Aβ peptide in amyloid fibrils studied by site-directed spin labeling.* J. Biol. Chem., 2002(277): p. 40810-40815.
38. Hoshino, M., et al., *Mapping the core of the beta(2)-microglobulin amyloid fibril by H/D exchange.* Nat. Struct. Biol., 2002(9): p. 332-336.
39. Jaroniec, C., et al., *High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy.* Proc. Natl. Acad. Sci., 2004(101): p. 711-716.
40. Jimenez, J., et al., *The protofilament structure of insulin amyloid fibrils.* Proc. Natl. Acad. Sci., 2004(99): p. 9196-9201.
41. Perutz, M., et al., *Amyloid fibers are water-filled nanotubes.* Proc. Natl. Acad. Sci., 2002(99): p. 5591-5595.
42. Morris, A., M. Watzky, and R. Finke, *Protein aggregation kinetics, mechanism, and curve-fitting: A review of the literature.* Biochim Biophys Acta, 2009(1794): p. 375-397.
43. Berg, J., J. Tymoczko, and L. Stryer, *Actin is a polar, self-assembling, dynamic polymer.* Biochemistry, 5th edition, 2002: p. 958-960.
44. Eisenberg, H., *Glutamate dehydrogenase: anatomy of a regulatory enzyme.* Acc. Chem. Res., 1971(4): p. 379-385.
45. Thusius, D., *Mechanism of bovine liver glutamate dehydrogenase self-assembly: II. Simulation of relaxation spectra for an open linear polymerization proceeding via a sequential addition of monomer units.* J. Mol. Biol., 1975(94): p. 367-383.
46. Jullien, M. and D. Thusius, *Mechanism of bovine liver glutamate dehydrogenase self- assembly III. Characterization of the association–dissociation stoichiometry with quasi-elastic light scattering.* J. Mol. Biol., 1976(101): p. 397-416.
47. Roberts, C., *Kinetics of irreversible protein aggregation: analysis of extended Lumry-Eyring models and implication for predicting protein shelf life.* J. Phys. Chem., 2003(107): p. 1194-1207.
48. Wetzel, R., *For protein misassembly, it's the "I" decade.* Cell, 1986(1996): p. 699-702.
49. Sandal, M., et al., *Conformational equilibria in monomeric alpha-synuclein at the single-molecule level.* PLoS Biology, 2008: p. 99-108.
50. Uversky, V., J. Li, and A. Fink, *Evidence for a partially folded intermediate in α-synuclein fibril formation.* J. Biol. Chem., 2001(276): p. 10737-10744.
51. Roberts, C., *Non-native protein aggregation kinetics.* Biotechnol. Bioeng., 2007(98): p. 927–938.
52. Ferrone, F., *Analysis of protein aggregation kinetics.* Methods Enzymol, 1999(309): p. 256-274.
53. Gast, K., et al., *Effects of environmental conditions on aggregation and fibril formation of barstar.* Eur Biophys, 2003(32): p. 710-723.
54. Speed, M., J. King, and D. Wang, *Polymerization mechanism of polypeptide chain aggregation.* Biotechnol Bioeng, 1997(54): p. 333-343.

55. Powers, E. and F. Ferrone, *Kinetic models for protein misfolding and association.* Protein Misfolding Diseases, 2010: p. 73-92.
56. Hurschman, A., et al., *Transthyretin aggregation under partially denaturing conditions is a downhill polymerization.* Biochemistry, 2004(43): p. 7365-7381.
57. Martin, J., *Molecular basis of the neurodegenerative diseases.* Molecular basis of the neurodegenerative diseases., 1999(340): p. 1970–1980
58. Hyman, B. and G. Van Hoesen, *Neuron numbers in Alzheimer's disease: cell-specific pathology.* Neurobiol. Aging 1987(8): p. 555-556.
59. Forno, L., *Neuropathology of Parkinson disease.* J. Neuropathol. Exp. Neurol, 1996(55): p. 259-272.
60. Myers, R., et al., *Decreased neuronal and increased oligodendroglial densities in Huntington's disease caudate nucleus.* J. Neuropathol. Exp. Neurol, 1991(50): p. 729-742.
61. McGeer, P. and E. McGeer, *The inflammatory response system of brain: implications for therapy of Alzheimer and other neurodegenerative diseases. .* Brain Res., 1995(21): p. 195–218
62. Wyss-Coray, T. and L. Mucke, *Inflammation in neurodegenerative disease — a double-edged sword.* Neuron 2002(35): p. 419–432
63. Lindholm, D., H. Wootz, and L. Korhonen, *ER stress and neurodegenerative diseases.* Cell Death Differ, 2006(13): p. 385-392.
64. Martins, V., *Cellular prion protein: on the road for functions.* FEBS Lett, 2002(512): p. 25–28
65. Chiarini, L., et al., *Cellular prion protein transduces neuroprotective signals.* EMBO J, 2002(21): p. 3317–3326.
66. Kuwahara, C., et al., *Prions prevent neuronal cell-line death.* Nature, 1999(400): p. 225-226.
67. Costa, C.D., K. Ancolio, and F. Checier, *Wild-type but not Parkinson's disease-related ala 53 -> Tyr mutant α-synuclein protects neural cells from apoptotic stimuli.* J. Biol. Chem., 1999(275): p. 24065-24069.
68. Loo, D., et al., *Apoptosis is induced by beta-amyloid in cultured central nervous system neurons. .* Proc. Natl Acad. Sci, 1993(90): p. 7951–7955
69. El-Agnaf, O., et al., *Aggregates from mutant and wild-type alpha-synuclein proteins and NAC peptide induce apoptotic cell death in human neuroblastoma cells by formation of beta-sheet and amyloid-like filaments.* FEBS Lett. , 1998(440): p. 71-75.
70. Forloni, G., et al., *Neurotoxicity of a prion protein fragment.* Nature, 1993(362): p. 543–546.
71. McGowan, E., J. Eriksen, and M. Hutton, *A decade of modeling Alzheimer's disease in transgenic mice.* Trends Gemet, 2206(22): p. 218-289.
72. SantaCruz, K., et al., *Tau suppresion in a neurodegenerative mouse model improves memory function.* Science, 2005(309): p. 476-481.
73. Bucciantini, M., et al., *Inherent toxicity of aggregates implies common mechanism for protein misfolding diseases. .* Nature, 2002(416): p. 507-511.

74. Soto, C., *Unfolding the role of protein misfolding in neurodegenerative diseases.* Nature Rev, 2003(4): p. 49-60.
75. Li, K., et al., *Immunocytochemical co-localization of the proteasome in ubiquitinated structures in neurodegenerative diseases and the elderly. .* Exp. Neurol. , 1997(56): p. 125-131.
76. Arispe, N., H. Polland, and E. Rojas, *Giant multilevel cation channels formed by Alzheimer disease amyloid f3-protein [AIBP-(1-40)] in bilayer membranes.* Proc. Natl. Acad. Sc, 1993(90): p. 567-571.
77. Stefani, M., *Generic cell dysfunction in neurodgenerative disorders: role of surfaces in early protein misfolding, aggregation, and aggregate cytotoxicity.* Neuroscientist, 2007(13): p. 519-531.
78. Caughey, B. and P. Landsbury, *Protofibrils, pores, fibrils and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders.* Annu Rev Neurosci, 2003(26): p. 267-298.
79. Kagan, B., R. Azimov, and R. Azimova, *Amyloid peptide channels.* J Membr Biol, 2004(202): p. 1-10.
80. Hirakura, Y. and B. Kagan, *Pore formation by beta-2-microglobulin: a mechanism for the pathogenesis of dialysis-associated amyloidosis.* Amyloid, 2001(8): p. 94-100.
81. Stefani, M. and C. Dobson, *Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution.* J. Mol. Med., 2003(81): p. 678-699.
82. Lashuel, H. and P. Landsbury, *Are amyloid diseases caused by protein aggregates that mimic bacterial pore-forming toxins?* Q Rev Biophys, 2006(39): p. 167-201.
83. Deshpande, A., et al., *Different conformations of amyloid beta induce neurotoxicity by distinct mechanisms in human cortical neurons.* J Neurosci, 2006(26): p. 6011-6018.
84. Demuro, A., et al., *Calcium dysregulation and membrane disruption as a ubiquitous neurotoxic mechanism of soluble amyloid oligomers.* J Biol Chem, 2005(280): p. 17294-17300.
85. DiFiglia, M., et al., *Aggregation of Huntingtin in Neuronal Intranuclear Inclusions and Dystrophic Neurites in Brain.* Science, 1997. **277**: p. 1990-1993.
86. Bates, G., *Huntingtin aggregation and toxicity in Huntington's disease.* Lancet, 2003(361): p. 1642-1644.
87. Salkoe, D., *Cell biology of protein misfolding: the examples of Alzheimer's and Parkinson's diseases.* Nature Cell Biol, 2004(6): p. 1054-1061.
88. Ross, C. and P. MA, *Protein aggregation and neurodegenerative disease.* Nature Med, 2004(10): p. S10-S17.
89. Cohen, F. and J. Kelly, *Therapeutic approaches to protein-misfolding diseases.* Nature, 2003(426): p. 905-909.
90. Terry, R., *Physical basis of cognitive alterations in Alzheimer's disease: synapse loss is the major correlate of cognitive impairment. .* Ann. Neurol, 1991(30): p. 572–580.

91. Tompkins, M. and W. Hill, *Contribution of somal Lewy bodies to neuronal death. .* Brain Res, 1997(775): p. 24-29.
92. Katzman, R., R. Terry, and R. DeTeresa, *Clinical, pathological, and neurochemical changes in dementia: a subgroup with preserved mental status and numerous neocortical plaques.* Ann Neurol, 1988(23): p. 138-144.
93. Gutenkunst, C., et al., *Nuclear and neuropil aggregates in Huntington's disease: relationship to neuropathology.* Ann. Neurol, 1999(19): p. 2522-2534.
94. Lim, K. and J. Tann, *Role of Ubiquitin-proteasome system in Parkinson's disease.* Role of Ubiquitin-proteasome system in Parkinson's disease, 2007(8): p. S13.
95. Majeski, A. and J. Dice, *Mechanisms of chaperone mediated autophagy.* Int. J. Biochem. Cell Biol, 2004(36): p. 2435–2444
96. Orr, A., S. li, and C.-E. Wang, *Trinucleotide repeat disorders.* Ann Rev Neurosci, 2007(30): p. 575-621.
97. McEwan, I., *Molecular mechanisms of androgen receptor- mediated gene regulation: structure-function analysis of the AF- 1 domain.* Endocr Relat Cancer, 2004(11): p. 281-293.
98. Nonhoff, U., M. Ralser, and F. Welzel, *Ataxin-2 interacts with the DEAD/H-Box RNA helicase DDX6 and interferes with P-bodies and stress granules.* Mol Biol Cell, 2007(18): p. 1385–1396.
99. Satterfield, T., S. Jackson, and L. Pallanck, *A drosophila homolog of the polyglutamine disease gene sca2 is a dosage- sensitive regulator of actin filament formation.* Genetics, 2002(162): p. 1687–1702.
100. Wiedemeyer, R., et al., *Ataxin-2 promotes apoptosis of human neuroblastoma cells.* Oncogene 2003(22): p. 401-411.
101. Burnett, B., F. Li, and R. Pittman, *The polyglutamine neuro- degenerative protein ataxin-3 binds polyubiquitylated proteins and has ubiquitin protease activity.* Hum Mol Genet, 2003(28512): p. 3195-3205.
102. Doss-Pepe, E., et al., *Ataxin-3 interactions with Rad23 and valosin-containing protein and its associations with ubiquitin chains and the proteasome are consistent with a role in ubiquitin-mediated proteolysis.* Mol Cell Biol, 2003(23): p. 6469-6483.
103. Frontali, M., *Spinocerebellar ataxia type 6: channelopathy or glutamine repeat disorder?* Brain Res Bull, 2001(56): p. 227-231.
104. Helmlinger, D., S. Hardy, and G. Abou-Sleymane, *Glutamine-expanded Ataxin-7 alters TFTC/STAGA recruitment and chromatin structure leading to photoreceptor dysfunction.* PLoS Biol, 2006(4): p. e67.
105. Maltecca, F., A. Filla, and I. Castaldo, *Intergenerational instability and marked anticipation in SCA-17.* Neurology, 2003(61): p. 1441–1443.
106. Wood, J., J. Yuan, and R. Margolis, *Atrophin-1, the DRPLA gene product, interacts with two families of WW domain-containing proteins.* Mol Cell Neurosci, 1998(11): p. 149-160.
107. Bates, G., P. Harper, and A. Jones, *Huntington's Disease.* Oxford Press, 2002.
108. Walker, F., *Huntington's Disease.* Lancet, 2007(369): p. 218-228.

109. Zuccato, C., M. Valenza, and E. Cattaneo, *Molecular mechanisms and potential therapeutical targets in Huntington's disease.* Physiol Rev, 2010(90): p. 905-981.
110. Reiner, A., et al., *Differential loss of striatal projection neurons in Huntington disease.* Proc Natl Acad Sci, 1988(85): p. 5733–5737.
111. Rosas, H., et al., *Cerebral cortex and the clinical expression of Huntington's disease: complexity and heterogeneity.* Brain, 2008(131): p. 1057–1068.
112. Gusella, J., et al., *A polymorphic DNA marker genetically linked to Huntington's disease.* Nature, 1983(306): p. 234-238.
113. HDCollaborativeResearch, G., *A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes.* Cell, 1993(72): p. 971–983.
114. Myers, R., *Huntington's disease genetics.* NeuroRx, 2004(1): p. 255–62.
115. Snowden, J., et al., *Awareness of involuntary movements in Huntington disease.* Arch Neurol, 1998(55): p. 801–805.
116. Paulsen, J., et al., *Critical periods of suicide risk in Huntington's disease. .* Am J Psychiatry, 2005(162): p. 725-731.
117. Watts, R. and W. Koller, *Movement disorders: neurologic principles and practice.* New York: McGraw-Hill, 1997.
118. Folstein, S., *Huntington's disease: a disorder of families.* Maryland: The Johns Hopkins University Press, 1989.
119. Craufurd, D. and J. Snowden, *Neuropsychological and neuropsychiatric aspects of Huntington's disease.* Bates G, Harper P, Jones L, eds. Huntington's disease, 2002(New York: Oxford University Press): p. 62-94.
120. Baliko, L., B. Csala, and J. Czopf, *Suicide in Hungarian Huntington's disease patients.* Neuroepidemiology, 2004(23): p. 258-260.
121. Maio, L.D., et al., *Suicide risk in Huntington's disease.* J Med Genet, 1993(30): p. 293-295.
122. Mahant, N., et al., *Huntington's disease: clinical correlates of disability and progression.* Neurology 2003(61): p. 1085-1092.
123. Gordon, A., et al., *Coordination of prehensile forces during precision grip in Huntington's disease.* Exp Neurol, 2000(163): p. 136-148.
124. Reilmann, R., et al., *Objective assessment of progression in Huntington's disease: a 3-year follow-up study. .* Neurology, 2001(57): p. 920-924.
125. Hayden, M., *Huntington's chorea.* New York: Springer, 1981.
126. Kremer, B., *Clinical neurology of Huntington's disease.* Bates G, Harper P, Jones L, eds. Huntington's disease, 2002. New York: Oxford University Press: p. 3-27.
127. Rubinsztein, D., *Molecular biology of Huntington's disease (HD) and HD-like disorders.* Pulst S, ed. Genetics of movement disorders., 2003(California: Academic Press): p. 365-377.
128. Rubinsztein, D., J. Leggo, and C. R, *Phenotypic characterization of individuals with 30–40 CAG repeats in the Huntington disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36–39 repeats.* Am J Hum Genet 1996(101): p. 3498–3503.

129. McNeil, M., A. Novelletto, and J. Srinidhi, *Reduced penetrance of the Huntington's disease mutation.* Hum Mol Genet, 1997(6): p. 775-779.
130. Ashizawa, T., et al., *CAG repeat size and clinical presentation in Huntington's disease.* Neurology, 1994(44): p. 1137–1143.
131. Semaka, A., J. Collins, and M. Hayden, *Unstable familial transmis- sions of Huntington disease alleles with 27–35 CAG repeats (inter- mediate alleles).* Am J Med Genet B Neuropsychiatr Genet, 2010(153): p. 314–320.
132. Ranen, N., et al., *Anticipation and instability of IT-15 (CAG)n repeats in parent-offspring pairs with Huntington disease.* Am J Hum Genet, 1995(57): p. 273–279.
133. Pearson, C., *Slipping while sleeping? Trinucleotide repeat expan- sions in germ cells.* Trends Mol Med, 2003(9): p. 490–495.
134. Andrew, S., et al., *The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease.* Nat Genet, 1993(4): p. 398–403.
135. Rubinsztein, D., et al., *Analysis of the huntingtin gene reveals a trinucleotide-length polymorphism in the region of the gene that contains two CCG-rich stretches and a correlation between decreased age of onset of Huntington's disease and CAG repeat number.* Hum Mol Genet, 1993(2): p. 1713–1715.
136. Wexler, N., J. Lorimer, and J. Porter, *Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset.* Proc Natl Acad Sci, 2004(101): p. 3498–3503.
137. Chattapadhyay, B., et al., *Modulation of age at onset of Huntington's disease patients by variations in TP53 and human caspase activated DNase (hCAD) genes.* Neurosci Lett 2005(374): p. 81–86.
138. Rosenblatt, A., R. Brinkman, and K. Liang, *Familial influence on age of onset among siblings with Huntington's disease.* Am J Med Genet, 2001(105): p. 399–403.
139. Kehoe, P., et al., *Age of onset in Huntington disease: sex specific influence of apolipoprotein E genotype and normal CAG repeat length.* J Med Genet, 1999(36): p. 108–11.
140. Andrew, S., et al., *A CCG repeat polymorphism adjacent to the CAG repeat in the Huntington disease gene: implications for diagnostic accuracy and predictive testing. .* Hum Mol Genet, 1994(3): p. 65-67.
141. Farrer, L., et al., *The normal Huntington disease (HD) allele, or a closely linked gene, influences age at onset of HD.* Am J Hum Genet, 1993(53): p. 125–130.
142. Snell, R., et al., *Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease.* Nat Genet, 1993(4): p. 393–397.
143. Vuillaume, I., et al., *Genetic polymorphisms adjacent to the CAG repeat influence clin- ical features at onset in Huntington's disease.* J Neurol Neurosurg Psychiatry, 1998(64): p. 758 –762.
144. Aziz, N., et al., *Normal and mutant HTT interact to affect clinical severity and progression in Huntington disease.* Neurology, 2009(73): p. 1280 –1285.
145. Busch, A., et al., *Mutant huntingtin promotes the fibrillogenesis of wild-type huntingtin: a potential mechanism for loss of huntingtin function in Huntington's disease.* J Biol Chem, 2003(278): p. 41452–41461.

146. Sapp, E., *Huntingtin localization in brains of normal and Huntington's disease patients.* Ann Neurol, 1997(42): p. 604–612.
147. Fusco, F., *Cellular localization of huntingtin in striatal and cortical neurons in rats: lack of cor- relation with neuronal vulnerability in Hunting- ton's disease.* J Neurosci, 1999(19): p. 1189–1202.
148. Vonsattel, J. and M. DiFiglia, *Huntington disease.* J Neuropathol Exp Neurol, 1988(57): p. 369-384.
149. Balleine, B., M. Liljeholm, and S. Ostlund, *The integrative function of the basal ganglia in instrumental conditioning.* Behav Brain Res, 2009(199): p. 43–52.
150. Kravitz, A., *Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry.* Nature, 2010(466).
151. Albin, R., *Abnormalities of striatal projection neurons and N-methyl-D-aspartate receptors in presymptomatic Huntington's disease.* N Engl J Med, 1990(322): p. 1293–1298.
152. Sapp, E., *Evidence for a preferential loss of enkephalin immunoreactivity in the external globus pallidus in low grade Huntington's disease using high resolution image analysis.* Neuroscience, 1995(64): p. 397–404.
153. Petersen, A. and M. Bjorkqvist, *Hypothalamic-endocrine aspects in Huntington's disease.* Eur J Neurosci, 2006(24): p. 961–967.
154. Gabery, S., *Changes in key hypothalamic neuropeptide populations in Huntington disease revealed by neuropathological analyses.* Acta Neuropathol, 2010(120): p. 777–788.
155. Bohanna, I., et al., *Magnetic resonance imaging as an approach towards identifying neuropathological biomarkers for Huntington's disease.* Brain Res Rev, 2008(58): p. 209 –225.
156. Paulsen, J., *Functional imaging in Huntington's disease.* Exp Neurol, 2009(216): p. 272–277.
157. Rosas, H., A. Feigin, and S. Hersch, *Using advances in neuroimaging to detect, understand, and monitor disease progression in*
Huntington's disease. NeuroRx, 2004(1): p. 263–272.
158. Rosas, H., et al., *Complexity and heterogeneity: what drives the ever-changing brain in Huntington's disease?* Ann NY Acad Sci, 2008(1147): p. 196–205.
159. Rosas, H., *Evidence for more widespread cerebral pathology in early HD: an MRI-based morpho- metric analysis.* Neurology, 2003(60): p. 1615–1620.
160. Rosas, H., et al., *Regional cortical thinning in preclinical Huntington dis- ease and its relationship to cognition.* Neurology, 2005(65): p. 745–747.
161. Lodi, R., et al., *Abnormal in vivo skeletal muscle energy metabolism in Huntington's disease and dentatorubropallidoluysian atrophy.* Ann Neurol, 2000(48): p. 72–76.
162. Saft, C., et al., *Mitochondrial impairment in patients and asymptomatic mutation carrier of Huntington's disease.* Mov Disord, 2005(20): p. 674 – 679.
163. Brouillet, E., et al., *3-Nitropropionic acid: a mitochondrial toxin to uncover physiopathological mechanisms underlying striatal degeneration in Huntington's disease.* J Neurochem, 2005(95): p. 1521–1540.

164. Feigin, A., et al., *Thalamic metabolism and symptom onset in preclinical Huntington's disease.* Brain, 2007(130): p. 2858–2867.
165. Eidelberg, D. and D. Surmeier, *Brain networks in Huntington disease.* J Clin Invest, 2011(121): p. 484-492.
166. Beal, M., et al., *Chronic quinolinic acid lesions in rats closely resemble Huntington's disease.* J Neurosci, 1991(11): p. 1649 –1659.
167. Conti, L. and E. Cattaneo, *Neural stem cell system: physiological players or in vitro entities?* Nat Rev Neurosci, 2010(11): p. 176–187.
168. Rigamonti, D., et al., *Loss of huntingtin function complemented by small molecules acting as repressor element 1/neuron restrictive silencer element silencer modulators.* J Biol Chem, 2007(276): p. 14545–14548.
169. Velier, J., et al., *Wild-type and mutant huntingtins function in vesicle trafficking in the secretory and endocytic pathways.* Exp Neurol, 1998(152): p. 34 – 40.
170. Varma, H., et al., *Selective inhibitors of death in mutant huntingtin cells.* Nat Chem Biol, 2007(3): p. 99-100.
171. Davies, S., et al., *Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation.* Cell, 1997(90): p. 537–548.
172. Mangiarini, L., et al., *Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice.* Cell, 1996(87): p. 493–506.
173. Reddy, R., et al., *Behavioral abnormali- ties and selective neuronal loss in HD transgenic mice expressing mutated full-length HD cDNA.* Nat Genet, 1998(20): p. 198–202.
174. Gray, M., et al., *Full-length human mutant huntingtin with a stable polyglutamine repeat can elicit progressive and selective neuropathogen- esis in BACHD mice.* J Neurosci, 2008(6182–6195).
175. Lin, C., et al., *Neurological abnormalities in a knock-in mouse model of Huntington's disease.* Hum Mol Genet, 2001(10): p. 137–144.
176. Shelbourne, P., et al., *A Huntington's disease CAG expansion at the murine Hdh locus is unstable and associated with behavioural abnormalities in mice.* Hum Mol Genet, 1999(8): p. 763–774.
177. Ishiguro, H., et al., *Age-dependent and tissue-specific CAG repeat instability occurs in mouse knock-in for a mutant Huntington's disease gene.* J Neurosci Res, 2001(65): p. 289–297.
178. Menalled, L., et al., *Time course of early motor and neuropathological anomalies in a knock-in mouse model of Huntington's disease with 140 CAG repeats.* J Comp Neurol, 2003(465): p. 11-26.
179. Wheeler, V., et al., *Length-dependent gametic CAG repeat instability in the Huntington's disease knock-in mouse.* Hum Mol Genet, 1999(8): p. 115–122.
180. Almeida, L.D., et al., *Lentiviral-mediated delivery of mutant huntingtin in the striatum of rats induces a selective neuropathology modulated by polyglutamine repeat size, huntingtin expression levels, and protein length.* J Neurosci, 2002(22): p. 3473–3483.

181. Yang, S., et al., *Towards a transgenic model of Huntington's disease in a non-human primate.* Nature, 2008(453): p. 921–924.
182. Krobitsch, S. and S. Lindquist, *Aggregation of huntingtin in yeast varies with the length of the polyglutamine expansion and the expression of chaperone proteins.* Proc Natl Acad Sci, 2000(97): p. 1589 –1594.
183. Meriin, A., et al., *Huntington toxicity in yeast model depends on polyglutamine aggregation mediated by a prion-like protein Rnq1.* J Cell Biol, 2002(157): p. 997–1004.
184. Faber, P., et al., *Polyglutamine- mediated dysfunction and apoptotic death of a Caenorhabditis elegans sensory neuron.* Proc Natl Acad Sci 1999(96): p. 179 – 184.
185. Parker, J., et al., *Expanded polyglutamines in Caenorhabditis elegans cause axonal abnormalities and severe dysfunction of PLM mechanosensory neurons without cell death.* Proc Natl Acad Sci, 2001(98): p. 13318–13323.
186. Ebert, A., et al., *Induced pluripotent stem cells from a spinal muscular atrophy patient.* Nature, 2009(457): p. 277-280.
187. Andrade, M. and P. Bork, *HEAT repeats in the Huntington's disease protein.* Nat Genet, 1995. **11**: p. 115-116.
188. Goldberg, Y., et al., *Cleavage of huntingtin by apopain, a proapoptotic cysteine protease, is modulated by the polyglutamine tract.* Nat Genet, 1996. **13**: p. 442–449.
189. Graham, R., et al., *Cleavage at the caspase-6 site is required for neuronal dysfunction and degeneration due to mutant huntingtin.* Cell, 2006. **125**: p. 1179 –1191.
190. Hermel, E., et al., *Specific caspase interactions and amplification are involved in selective neuronal vulner- ability in Huntington's disease.* Cell Death Differ, 2004. **11**(424 – 438).
191. Wellington, C., et al., *Caspase cleavage of mutant huntingtin precedes neurodegeneration in Huntington's disease.* J Neurosci, 2002. **22**: p. 7862–7872.
192. Bizat, N., et al., *Calpain is a major cell death effector in selective striatal degeneration induced in vivo by 3-nitropropionate: implications for Huntington's disease.* J Neurosci, 2003. **23**: p. 5020 –5030.
193. Gafni, J. and L. Ellerby, *Calpain activation in Huntington's disease.* J Neurosci, 2002. **22**: p. 4842– 4849.
194. Goffredo, D., et al., *Calcium-dependent cleavage of endogenous wild-type huntingtin in primary cortical neurons.* J Biol Chem, 2002. **277**: p. 39594–39598.
195. Lunkes, A., et al., *Proteases acting on mutant huntingtin generate cleaved products that differentially build up cytoplasmic and nuclear inclusions.* Mol Cell, 2002. **10**: p. 259–269.
196. Kim, Y., et al., *Caspase 3-cleaved NH2-terminal fragments of wild-type and mutant huntingtin are present in normal and Huntington's disease brains, associate with membranes, undergo calpain-dependent proteolysis.* Proc Natl Acad Sci, 2001. **98**: p. 12784-12789.

197. Dyer, R. and C. McMurray, *Mutant protein in Huntington disease is resistant to proteolysis in affected brain.* Nat Genet, 2001. **29**: p. 270–278.
198. Kalchman, M., et al., *Huntingtin is ubiquitinated and interacts with a specific ubiquitin- conjugating enzyme.* J Biol Chem, 1996. **271**: p. 19385–19394.
199. Humbert, S., et al., *The IGF-1/Akt pathway is neuroprotective in Huntington's disease and involves Huntingtin phosphorylation by Akt.* Dev Cell, 2002(2): p. 831–837.
200. Rangone, H., et al., *The serum- and glucocorticoid-induced kinase SGK inhibits mutant huntingtin-induced toxicity by phosphorylating serine 421 of huntingtin.* Eur J Neurosci, 2004. **19**: p. 273–279.
201. Warby, S., et al., *Huntingtin phosphorylation on serine 421 is significantly reduced in the striatum and by polyglutamine expansion in vivo.* Hum Mol Genet, 2005. **14**: p. 1569–1577.
202. Luo, S., et al., *Cdk5 phosphorylation of huntingtin reduces its cleavage by caspases: implications for mutant huntingtin toxicity.* J Cell Biol, 2005. **169**: p. 647–656.
203. Schilling, B., et al., *Huntingtin phosphorylation sites mapped by mass spectrometry. Modulation of cleavage and toxicity.* J Biol Chem, 2006. **281**: p. 23686–23697.
204. Khoshnan, A., et al., *Activation of the IkappaB kinase complex and nuclear factor-kappaB contributes to mutant huntingtin neurotoxicity.* Neurosci, 2004. **24**: p. 7999 – 8008.
205. Thompson, L., et al., *IKK phosphorylates Huntingtin and targets it for degradation by the proteasome and lysosome.* J Cell Biol, 2009. **187**: p. 1083–1099.
206. Geiss-Friedlander, R. and F. Melchior, *Concepts in sumoylation: a decade on.* Nat Rev Mol Cell Biol, 2007. **8**: p. 947–956.
207. Steffan, J., et al., *SUMO modification of Huntingtin and Huntington's disease patholog.* Science, 2004. **304**: p. 100–104.
208. Yanai, A., et al., *Palmitoylation of huntingtin by HIP14 is essential for its trafficking and function.* Nat Neurosci, 2006. **9**: p. 824–831.
209. Duyao, M., et al., *Inactivation of the mouse Huntington's disease gene homolog Hdh.* Science, 1995. **269**: p. 407– 410.
210. Zeitlin, S., et al., *ncreased apoptosis and early embryonic lethality in mice nullizygous for the Huntington's disease gene homologue.* Nat Genet, 1995. **11**: p. 155-163.
211. Reiner, A., et al., *Wild-type huntingtin plays a role in brain development and neuronal survival.* Mol Neurobiol, 2003. **28**: p. 259 –276.
212. Rigamonti, D., et al., *Wild-type huntingtin protects from apoptosis upstream of caspase-3.* J Neurosci, 2000. **20**: p. 3705–3713.
213. Rigamonti, D., et al., *Huntingtin's neuroprotective activity occurs via inhibition of procaspase-9 processing.* J Biol Chem, 2001. **276**: p. 14545–14548.
214. Ho, L., et al., *Wild type Huntingtin reduces the cellular toxicity of mutant Huntingtin in mammalian cell models of Huntington's disease.* J Med Genet, 2001. **38**: p. 450–452.

215. Zhang, Y., et al., *Huntingtin inhibits caspase-3 activation.* EMBO J, 2006. **25**: p. 5896–5906.
216. Gunawardena, S., et al., *Disruption of axonal transport by loss of huntingtin or expression of pathogenic polyQ proteins in Drosophila.* Neuron, 2003. **40**: p. 25-40.
217. Trushina, E., et al., *Mutant huntingtin impairs axonal trafficking in mammalian neurons in vivo and in vitro.* Mol Cell Biol, 2004. **24**: p. 8195– 8209.
218. Gunawardena, S. and L. Goldstein, *Polyglutamine diseases and transport problems: deadly traffic jams on neuronal highways.* Arch Neurol, 2005. **62**: p. 46–51.
219. Smith, R., P. Brundin, and J. Li, *Synaptic dysfunction in Huntington's disease: a new perspective.* Cell Mol Life Sci, 2005. **62**: p. 1901–1912.
220. Temussi, P., L. Masino, and A. Pastore, *From Alzheimer to Huntington: why is a structural understanding so difficult?* The EMBO J, 2003(22): p. 355-361.
221. Foguel, D. and J. Silva, *New insights into the mechanisms of protein misfolding and aggregation in amyloidogenic diseases derived from pressure studies.* Biochemistry, 2004(43): p. 11361–11370.
222. Ollesch, J., et al., *Prion protein α-to-β transition monitored by time-resolved Fourier transform infrared spectroscopy.* Appl. Spectrosc, 2007(61): p. 1025-1031.
223. Perálvarez-Marín, A., A. Barth, and A. Gräslund, *Time-resolved infrared spectroscopy of pH-induced aggregation of the Alzheimer Aβ1–28 peptide.* J. Mol. Biol., 2008(379): p. 589–596.
224. Burkoth, T., et al., *Struture of the β-amyloid(10-35) fibril.* J. Am. Chem. Soc. , 2000(122): p. 7883-7889.
225. Yong, W., et al., *Structure determination of micelle-like intermediates in amyloid β-protein fibril assembly by using small angle neutron scattering.* Proc. Natl. Acad. Sci. , 2002(99): p. 150-154.
226. Hamill, A., S. Wang, and C. Lee, *Solution structure of an amyloid-forming protein during photoinitiated hexamer-dodecamer transitions revealed through small angle neutron scattering.* Biochemistry, 2007(46): p. 7694-7705.
227. Vestergaard, B., et al., *A helical structural nucleus is the primary elongating unit of insulin amyloid fibrils.* PLoS Biol, 2007(5): p. e134.
228. Naiki, H. and F. Gejyo, *Kinetic analysis of amyloid fibril formation.* Methods Enzymol 1999(309): p. 305-318.
229. Waugh, D., et al., *Studies of the nucleation and growth reactions of selected types of insulin fibrils.* J Am Chem Soc, 1953(75): p. 2592–2600.
230. Powers, E. and D. Powers, *The kinetics of nucleated polymerizations at high concentrations: Amyloid fibril formation near and above the "super-critical concentration.'.* Biophys J, 2006(91): p. 122–132.
231. Waugh, D., *A mechanism for the formation of fibrils from protein molecules.* J Cell Comp Physiol, 1957(49): p. 145–164.
232. Oosawa, F. and S. Asakura, *Thermodynamics of the polymerization of protein.* London: Academic Pres, 1975.

233. Jacques, D. and J. Trewhella, *Small-angle scattering for structural biology—expanding the frontier while avoiding the pitfalls.* Prot Sci, 2010. **19**: p. 642-657.
234. Neylon, C., *Small angle neutron and X-ray scat- tering in structural biology: recent examples from the literature.* Eur Biophys J Biophys Lett, 2008(37): p. 531-541.
235. Svergun, D. and M. Koch, *Small-angle scattering studies of biological macromolecules in solution.* Rep Prog Phys, 2003(66): p. 1735–1782.
236. Guinier, A., *The diffusion of X-rays under the extremely weak angles applied to the study of fine particles and colloidal suspension.* Comptes Rendus Hebdomadaires Des Seances De L Acad Des Sci, 1938(206): p. 1374-1376.
237. Guinier, A. and G. Fournet, *Small-angle scattering of X-rays (structure of matter series).* New York: Wiley, 1955.
238. Hjelm, R., *The small-angle approximation of X-ray and neutron scatter from rigid rods of non-uniform cross-section and finite length.* J Appl Crystallogr, 1985(18): p. 452-460.
239. Chen, S. and D. Bendedouch, *Structure and interactions of proteins in solution studied by small-angle neutron-scattering.* Methods Enzymol, 1986(130): p. 79-116.
240. Stanley, C., et al., *Protein Structure and Hydration Probed by SANS and Osmotic Stress.* Biophys J, 2008(94): p. 2777-2789.
241. VanOudenhove, J., et al., *Analysis of PKR structure by small-angle scattering.* J Mol Biol, 2009(387): p. 910-920.
242. Svergun, D., et al., *Protein hydration in solution: experimental observation by x-ray and neutron scattering.* Proc. Natl. Acad. Sci, 1998(95): p. 2267-2272.
243. Whitten, A. and J. Trewhella, *Small-angle scattering and neutron contrast variation for studying bio-molecular complexes.* Methods Mol. Biol., 2009. **544**: p. 307-323.
244. Kuwamoto, S., S. Akiyama, and T. Fujisawa, *Radiation damage to a protein solution, detected by synchrotron X-ray small-angle scattering: dose-related considerations and suppression by cryoprotectants.* J Synchrotron Radiat, 2004(11): p. 462-468.
245. Svergun, D., *Restoring three-dimensional structure of biopolymers from solution scattering.* J Appl. Crystallogr., 1999. **30**: p. 792-797.
246. Konarev, P., et al., *ATSAS 2.1, a program package for small-angle scattering data analysis.* J Appl Crystallogr, 2006(39): p. 277–286.
247. Walther, D., F. Cohen, and S. Doniach, *Reconstruction of low-resolution three-dimensional small-angle X-ray solution scattering data for biomolecules.* J. Appl. Chrystallogr., 2000. **33**: p. 350-363.
248. Pedersen, J., D. Posselt, and K. Mortensen, *Analytical treatment of the resolution function for small-angle scattering.* J. Appl. Chrystallogr., 1990(23): p. 321-333.
249. Kirkpatrick, S., C. Gelatt, and M. Vecci, *Optimization by simulated annealing.* Science, 1983(220): p. 671-680.

250. Volkov, V. and D. Svergun, *Uniqueness of ab initio shape determination in small angle scattering.* J. Appl. Chrystallogr., 2003(36): p. 860-864.
251. Svergun, D., *Small-angle X-ray and neutron scattering as a tool for structural systems biology.* Biol. Chem, 2010. **391**: p. 737-743.
252. Whitten, A., et al., *Cardiac myosin-binding protein C decorates F-actin: implications for cardiac function.* Proc. Natl. Acad. Sci., 2008. **105**: p. 18360–18365.
253. Kuemmerle, S., et al., *Huntingtin aggregates may not predict neuronal death in Huntington's disease.* Ann. Neurol, 1999. **46**(842-849).
254. Saudou, F., et al., *Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions.* Cell, 1998. **95**: p. 55-66.
255. Slow, E., et al., *Absence of behavioral abnormalities and neurodegeneration in vivo despite widespread neuronal huntingtin inclusions.* Proc Natl Acad Sci, 2005(102): p. 11402–11407.
256. Legleiter, J., et al., *Mutant huntingtin fragments form oligomers in a polyglutamine length-dependent manner in vitro and in vivo.* J. Biol. Chem., 2010. **285**: p. 14777-14790.
257. Wacker, J., et al., *Hsp70 and Hsp40 attenuate formation of spherical and annular polyglutamine oligomers by partitioning monomer.* Nat. Struct. Mol. Biol, 2004. **11**: p. 1215-1222.
258. Olshina, M., et al., *Tracking mutant huntingtin aggregation kinetics in cells reveals three major populations that include an invariant oligomer pool.* J. Biol. Chem., 2010. **285**: p. 21807-21816.
259. Nayak, A., et al., *A universal pathway for amyloid nucleus and precursor formation for insulin.* Proteins, 2009. **74**: p. 556-565.
260. Berthelier, V. and R. Wetzel, *An assay for characterizing in vitro the kinetics of polyglutamine aggregation.* Methods Mol. Biol., 2003. **217**: p. 295-303.
261. Rasband, W., *ImageJ.* U. S. National Institutes of Health, Bethesda, MD, USA, http://rsb.info.nih.gov/ij/. 1997-2004.
262. Svergun, D., C. Barberato, and M. Koch, *CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates.* J. Appl. Crystallogr, 1995. **28**: p. 768-773.
263. Lee, C., et al., *A three-stage kinetic model of amyloid fibrillation.* Biophys. J, 2007. **92**: p. 3448-3458.
264. Thakur, A., et al., *Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism.* Nat. Struct. Mol. Biol, 2009. **16**: p. 380-389.
265. Chalikian, T. and B. KJ, *On volume changes accompanying conformational transitions of biopolymers.* Biopolymers, 1996. **39**: p. 619-626.
266. Fitzkee, N. and G. Rose, *Reassessing random-coil statistics in unfolded proteins.* Proc. Natl. Acad. Sci., 2004. **101**: p. 12497-12502.
267. Ossato, G., et al., *A two-step path to inclusion formation of huntingtin peptides revealed by number and brightness analysis.* Biophys. J, 2010. **98**: p. 3078-3085.

268. Poirier, M., et al., *Huntingtin spheroids and protofibrils as precursors in polyglutamine fibrilization.* J. Biol. Chem., 2002. **277**: p. 41032-41037.
269. Chen, S., et al., *Polyglutamine aggregation behavior in vitro supports a recruitment mechanism of cytotoxicity.* J. Mol. Biol., 2001. **311**: p. 173-182.
270. Scherzinger, E., et al., *Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology.* Proc. Natl. Acad. Sci., 1999. **96**: p. 4604-4609.
271. Pedersen, J., *Analysis of small-angle scattering data from colloids and polymer solutions: modeling and least-squares fitting.* Adv. Colloid Interface Sci., 1997. **70**: p. 171-210.
272. Sharma, D., et al., *Polyglutamine homopolymers having 8-45 residues form slablike β-crystallite assemblies.* Proteins, 2005. **61**: p. 398-411.
273. Sikorski, P. and E. Atkins, *New model for crystalline polyglutamine assemblies and their connection with amyloid fibrils.* Biomacromolecules, 2005. **6**: p. 425-432.
274. Paravastu, A., et al., *Molecular structural basis for polymorphism in Alzheimer's β-amyloid fibrils.* Proc. Natl. Acad. Sci., 2008. **105**: p. 18349-18354.
275. Kim, M., et al., *Secondary structure of Huntingtin amino-terminal region.* Structure, 2009. **17**: p. 1205-1212.
276. Perutz, M., et al., *Aggregation of proteins with expanded glutamine and alanine repeats of the glutamine-rich and asparagine-rich domains of Sup35 and of the amyloid β-peptide of amyloid plaques.* Proc. Natl. Acad. Sci., 2002. **99**: p. 5596-5600.
277. Sawaya, M., et al., *Atomic structures of amyloid cross-β spines reveal varied steric zippers.* Nature, 2007. **447**: p. 5596-5600.
278. Bhattacharyya, A., A. Thakur, and R. Wetzel, *Polyglutamine aggregation nucleation: thermodynamics of a highly unfavorable protein folding reaction.* Proc. Natl. Acad. Sci. USA, 2005. **102**: p. 15400–15405.
279. Aiken, C., et al., *Phosphorylation of Threonine 3 Implications for huntingtin aggregation and neurotoxicity.* J Biol Chem, 2009. **284**: p. 29427-29436.
280. Atwal, R., et al., *Kinase inhibitors modulate huntingtin cell localization and toxicity.* Nat Chem Biol, 2011. **7**: p. 453-460.
281. Stanley, C., T. Perevozchikova, and V. Berthelier, *Structural Formation of Huntingtin Exon 1 Aggregates Probed by Small-Angle Neutron Scattering.* Biophys J, 2011. **100**: p. 2505-2513.
282. Rasband, W.S., *ImageJ,* U. S. National Institutes of Health, Bethesda, MD, USA. 1997-2004.
283. Svergun, D., *Mathematical methods in small-angle scattering data analysis.* J. Appl. Cryst., 1991. **24**: p. 485-492.
284. Chen, S., F. Ferrone, and R. Wetzel, *Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation.* Proc Natl Acad Sci U S A, 2002. **99**: p. 11884–11889.
285. Southwell, A., J. Ko, and P. Patterson, *Intrabody Gene Therapy Ameliorates Motor, Cognitive, and Neuropathological Symptoms in Multiple Mouse Models of Huntington's Disease.* J. Neurosci., 2009. **29**: p. 13589 –13602.

286. Ross, C., *Polyglutamine pathogenesis: Emergence of unifying mechanisms for Huntington's disease and related disorders.* Neuron, 2002. **35**: p. 819-822.
287. Chen, S., et al., *Amyloid-like features of polyglutamine aggregates and their assembly kinetics.* Biochemistry, 2002. **41**(7391–7399).
288. Khare, S., et al., *Molecular orirign of polyglutamine aggregation in neurodegenerative diseases.* PLoS Comp Bio, 2005. **1**: p. e30.
289. Bhattacharyya, A., et al., *Oligoproline effects on polyglutamine conformation and aggregation.* J Mol Bio, 2005. **355**: p. 524-535.
290. Kar, K., et al., *Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent.* Nature Stru, Mol Bio, 2011. **18**: p. 328-337.
291. Petkova, A., et al., *Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils.* Science, 2005. **307**: p. 262-265.
292. Nekooki-Machidaa, Y., et al., *Distinct conformations of in vitro and in vivo amyloids of huntingtin-exon1 show different cytotoxicity.* Proc. Natl. Acad. Sci., 2009. **106**: p. 9679-9684.
293. Jahn, T., et al., *The common architecture of cross-beta amyloid.* J. Mol. Biol., 2010. **395**: p. 717-727.

APPENDIX

Probing the Internal Organization of the Mature Htt Fibrils.

A described in the chapter 3, the htt fibrils formed by the normal and pathological htt exon-1 fragments demonstrated a significantly different morphology. While fibrils formed by $NtQ_{22}P_{10}$ peptide show a dense cross-beta structure (when followed from the initial stages of fibrillzation to the mature fibrils), aggregates formed by $NtQ_{42}P_{10}$ peptide resemble the beta-helical conformation of cross-section as one proposed by Perutz.

In order to directly probe the internal structure of fibrils, measured by number of internal waters, and confirm the presence of the hollow pore within pathological Huntingtin fibrils, we performed contrast variation SANS experiments.

Probing the amount of water molecules related to proteins, and especially of alteration in hydration associated with conformational changes and binding reactions, are essential in extending of our knowledge about the role of water. The classification of hydration water usually depends on the specifics of the technique studying water structure [2, 3]. Here, we concentrate on protein-associated water that excludes small molecules by a preferential hydration of external protein surface or by a steric exclusion from internal cavities [4-7].

Protein hydration as characterized by solute exclusion can be studied using a number of methods, such as light, x-ray, and neutron scattering, as well as sedimentation equilibrium [8, 9], using the distinction in properties of water and solute/water mixtures. Small-angle neutron scattering (SANS) is exceptionally suitable for investigating waters, correlated with protein and excluding small solutes[10]. The following three regions of scattering contrast are made with added solute: the protein, the protein-associated water, and the bulk water/ solute solution. The layer of excess water adjacent to the protein that keeps solutes out has a distinctive scattering length density from both the protein and the bulk solution containing solute. The scattering intensity of the protein then consists partially of the protein itself and its preferential hydration layer. In experimental terms, SANS determines the dry protein volume, or that inaccessible to the solvent, when in H_2O, D_2O, and mixtures of the two [11]. The protein-associated water volume determined in solute/D_2O mixtures is then the volume inaccessible to the solute but accessible by water.

To describe htt-exon 1 aggregate structure and hydration, we conducted static SANS experiments on htt fibrils in the presence of hydrogenated osmolytes. These osmolyte contrast variation studies are comparable to conventional H_2O/D_2O contrast variation but with the supplementary advantage of elucidating water present around and within the structures. We investigate here the exclusion of TEG from the fibrils formed by $NtQ_{42}P_{10}$ peptide. Changes in I(0) and the apparent Rg, in the presence of the increasing osmolyte concentration can be utilized for gaining the information on the number of water molecules associated

with proteins. I(0) is sensitive to the number of water while Rg depends both on the number of waters and their location.

Generally, the scattering intensity of the protein and water treated as composite particle can be described as follows:

$$I_s(0) = \frac{c_p N_A}{M_p}(\Delta\rho_p V_p + \Delta\rho_w V_w)^2 \qquad (3)$$

where c_p is the protein concentration, N_A is Avogadro's number, M_p is the protein weight, $\Delta\rho_w$ (=ρ_w - ρ_o) is the contrast between protein-associated water (D_2O) and bulk solution, and V_p and V_w are the molecular volumes of the protein and protein-associated water, respectively. For fitting, we using the intensity ratio as a function of the volume fraction of osmolyte, f_v

$$\frac{I_s(0)}{I_0(0)} = \left(1 + f_v \frac{\rho_w - \rho_s}{\rho_p - \rho_w}\left(\frac{V_p + V_w}{V_p}\right)\right)^2 \qquad (4)$$

Likewise, a composite R_g can be defined as:

$$R_g^2 = \frac{\Delta\rho_p V_p R_{g,p}^2 + \Delta\rho_w V_w R_{g,w}^2}{\Delta\rho_p V_p + \Delta\rho_w V_w} \qquad (5)$$

with $R_{g,p}$ and $R_{g,w}$ representing the individual radii of gyration of the protein and the associated water, respectively (obtained relative to their common center of scattering). Both Equations 4 and 5 were used to fit a series of data with varying TEG f_v assuming V_w is a constant across the entire solute concentration range [12, 13].

Similarly, we can derive the composite expressions for the 2D analogs: the composite cross-sectional zero-angle scattering intensity, $I_c(0)$

$$I_c(0) = \frac{c_p M_L}{N_A} \pi \left(\Delta\rho_p \bar{v}_p + \Delta\rho_w \bar{v}_w \right)^2 \qquad (6)$$

which will result in an analogous expression as Equation 6 to describe the zero-angle intensity ratio for data fitting purposes. The water volume, V_w, is obtained from

$$V_w = \frac{\bar{v}_w M_w}{N_A} \qquad (7)$$

where M_w is the molecular weight of water (for D_2O, M_w = 20 g/mol). The composite cross-sectional radius of gyration, R_c, is derived from the general equation

$$R_c^2 = \frac{\int_0^R r^2 dA}{\int_0^R dA} \qquad (8)$$

where $dA = 2\pi r$, to yield the core protein $R_{c,p} = R_p / \sqrt{2}$ with R_p being the protein cylindrical radius. Similarly, the cylindrical shell R_c, $R_{c,shell}$, is found from integrating over the limits R_p to R_w and defining the effective water shell thickness, $t = R_w - R_p$ to yield

$$R_{c,shell}^2 = \frac{1}{2} \left[\frac{\left(R_p + t \right)^4 - R_p^4}{\left(R_p + t \right)^2 - R_p^2} \right] \qquad (9)$$

and to obtain the composite $R_{c,comp}$

$$R_{c,comp}^2 = \frac{\Delta\rho_p R_p^2 R_{c,p}^2 + \Delta\rho_w \left[\left(R_p + t \right)^2 - R_p^2 \right] R_{c,shell}^2}{\Delta\rho_p R_p^2 + \Delta\rho_w \left[\left(R_p + t \right)^2 - R_p^2 \right]} \qquad (10)$$

From a determined water accessible surface area or volume of water, the number of preferentially hydrating water molecules, N_w, is calculated by utilizing 9 $Å^2$ surface area and 30 $Å^3$ volume per water molecule, respectively [14].

Osmolyte contrast variation SANS experiments on fibrils from pathological htt-exon1 peptides were tested in D_2O PBS buffer at 20 ^oC. We mixed these fibrils that were initially formed in the absence of any osmolyte into TEG solutions of concentrations ranging from 3 to 14% including a background run without protein for each condition. Equations (6), (9), and (10) were used for analysis. It was assumed that from $I(0)$, R_g, and the shape of the SANS curve, $I(Q)$, the level and localization of hydration within and around the fibril cross-section could be elucidated.

Upon completion of the described experiments the number of waters associated with $NtQ_{42}P_{10}$ was found. The obtained curve of $I_{c,s}(0)/I_{c,0}(0)$ vs. the varying volume fraction of TEG yield a slope, indicative of 78 ± 15 waters (Figure 2A). The localization of the water molecules was significantly hampered as the error bars on graph Rc vs. fv were prohibitory large.

As it was mentioned above, the changes in the R_c as a function of f_v are supposed to provide information on the localization of the waters. It is assumed that in the case of solid cylinder all the calculated waters would be predicted to be on the outside of the fibrils. This would be expressed in the more pronounced slope of R_c vs. f_v, since that radius of cross-section would appear smaller as the concentration of an osmolyte increases. This phenomena can be explained by the fact that solute-excluded hydration layer around the protein has a contrast relative

to the bulk solution, $\Delta\rho_w$ ($=\rho_w - \rho_o$), that is opposite in sign from the protein, $\Delta\rho_p$ ($=\rho_p - \rho_o$) [10].

Alternatively, if the mature fibrils represent a hollow cylinder model or a set of filamentous hollow cylinders, the substantial number of waters would be contained on the inside, and the observed radius of cross-section would not be strongly affected by the increasing percent value of osmolyte.

For $NtQ_{42}P_{10}$, the slope of R_c vs. f_v is shown on the figure 3. The localization of the water molecules is significantly hampered as the error bars were prohibitory large. The error in obtained R_c increases with elevating osmolyte concentration, since the adding up a solute diminish the contrast between protein and solution and also enlarge the incoherent background scattering that appears mostly due to the solute hydrogens [10].

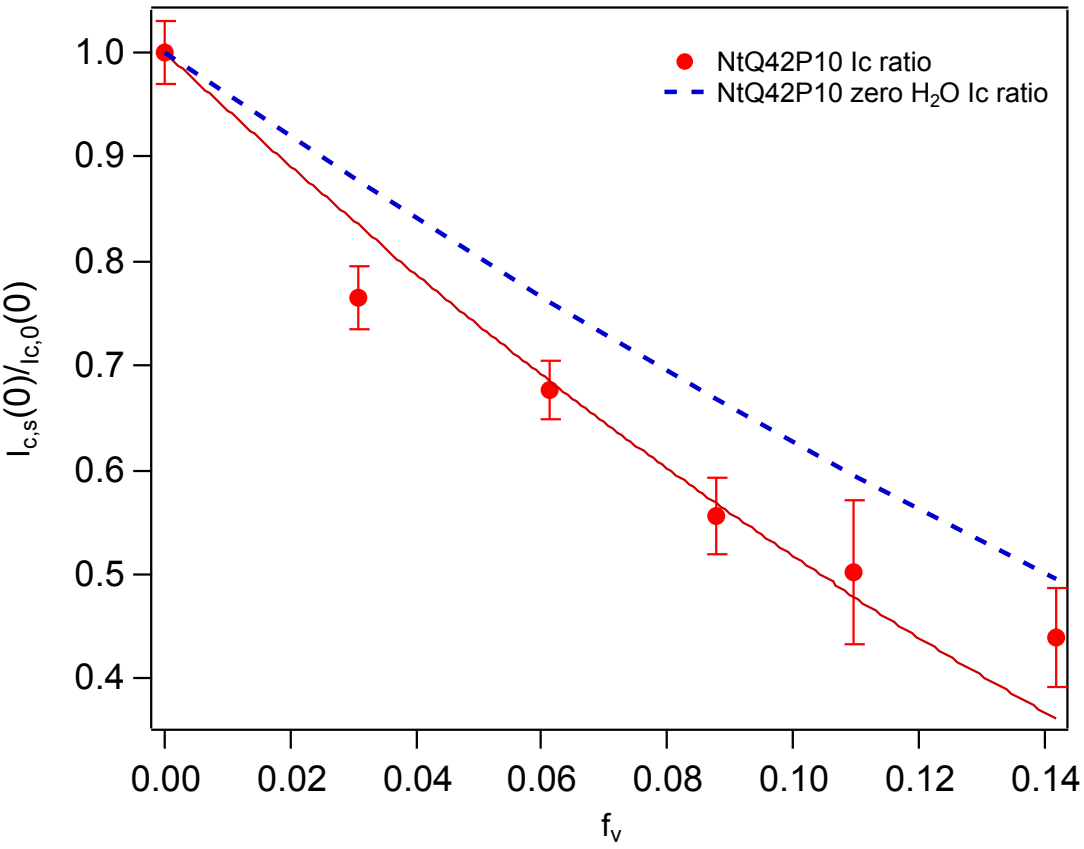


Figure A1. The $NtQ_{42}P_{10}$ zero-angle neutron scattering intensity ratio, $I_{c,s}(0)/I_{c,0}(0)$; as a function of TEG volume fraction, fv.

Solid line is a fit obtained using Eq. 6.The obtained number of waters is 78 ± 15. Dashed lines are given for a case of zero water present within the system to guide the eye.

The obtained (from $I_{c,s}(0)/I_{c,0}(0)$ vs. fv) number of waters was used to put constraints on the internal structure of the fibrils through calculated geometrically volume accessible for waters. Three models have been assumed: a semi-solid cylinder with peptide arranged in a loose way; a single hollow cylinder; and a fibril consisting of 2-3 hollow filaments.

1). In the case of single, semi-solid cylinder model, the experimental values of R_c (27.5 Å) and M_L (1475.3 Da/Å) were used to calculate the area inside the fibrils, potential for accommodating water molecules:

$$Area = \pi R_c^2 - \frac{M_L}{N_A}\vartheta \qquad (11)$$

with a resulting area of 2960 Å, yielding 329 waters, that would be required to fill fibrils with the given model of cross-section and length 4.75 Å. This number significantly exceeds the obtained from contrast variation value, excluding this particular case from the potential pool of models.

2). The structural parameters for the next model, representing a single hollow cylinder were calculated using the previously mentioned M_L and the expression:

$$R_{c,hcyl} = \sqrt{\frac{\left(R_i^2 + R_0^2\right)}{2}} \qquad (12)$$

where R_i and R_0 represent internal and outer radius of fibrils. The obtained values for these parameters were found to be 32.3 Å and 21.6 Å, respectively. From these numbers, the predicted number of waters present within the volume inside the fibrils, is 163.7, which is again larger than experimental value.

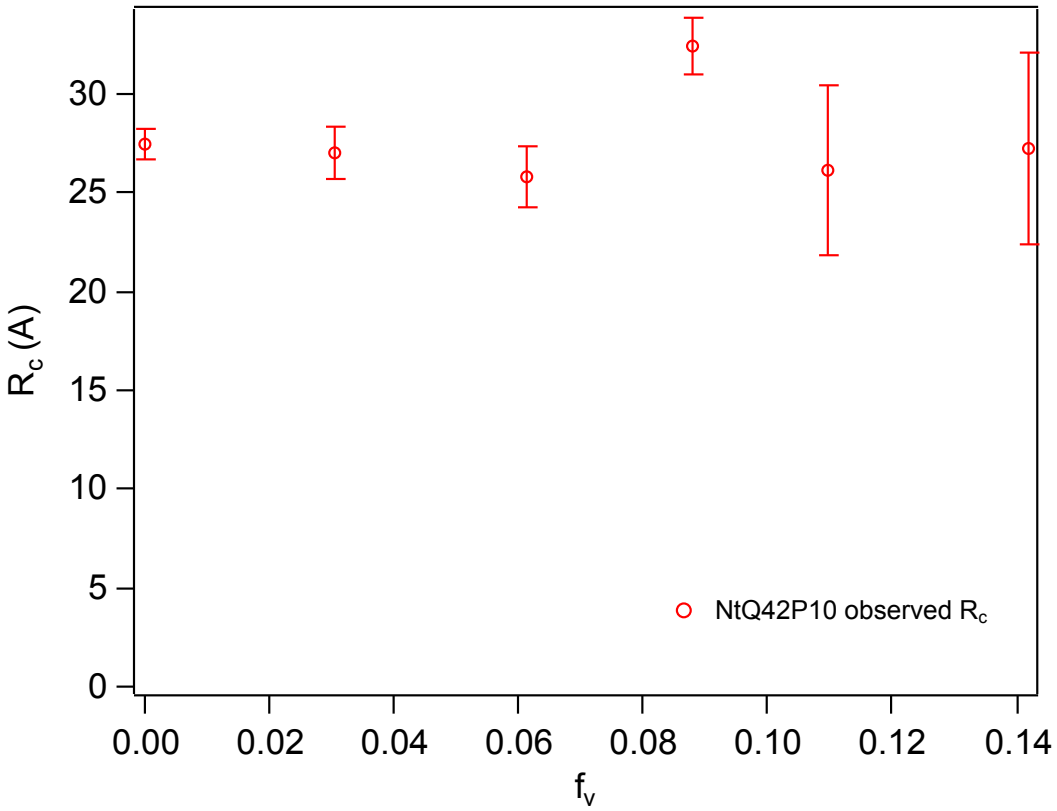


Figure A2. Measured change in $NtQ_{42}P_{10}$ R_c versus solute f_v.

The large value of errors observed at the higher concentration of TEG does not allow the reliable determination of the number of water present, as well as their location.

One of the alternative possibilities is that in the case of the fibril with internal radius of 21.6 Å, the TEG molecules could enter inside the cylinder. Even though the chance for such event is sufficiently low since the entrance of osmolyte to the constricted pore is unfavorable, sterically it is still feasible (Rg of TEG is 5.5 Å, which yields radius equal to 7.1 Å and diameter of 14.2 Å). In this case, the area available for water molecules is significantly reduced and we assume a single (3 Å) layer of water adjacent to the internal wall of the fibrils. According to this assumption, the estimated number of internal of waters in the presence of TEG inside constitutes 40 molecules, with the remaining amount possibly hydrating the outside of the fibrils.

3). The third model represents a mixture (50:50) of fibrils consisting of 2 and 3 hollow filaments with the internal radius equal to 5.9 and external 15.9 Angstroms. According to the internal area, representing the average between 2 and 3 filaments' areas, the number of waters residing on the inside is 31, while the rest is hydrating the outer surface of the fibrils. The predicted, optimal hydration layer of the external part of fibrils would contain averagely 74.7 water molecules per 4.75 Å, however it is possible that the layer of water repelling TEG is not uniform.

Performed mathematical analysis on hydration profiles of the proposed models revealed that the semi-solid cylinder approximation could be excluded from the potential cross-sectional organization of the mature fibrils. At the same time, there are some levels of possibility that the mature fibrils can be either a

single hollow cylinder (less probable) or a set of 2 or 3 hollow filaments constituting mature fibrils.

The computational coarse-grained modeling can be potentially applied to build the pdb structures of $NtQ_{42}P_{10}$ mature fibrils with the numbers and locations of water molecules as predicted in models 2 and 3. These profiles can be utilized to obtain theoretical SANS scattering curves, which subsequently can be compared with the experimental scattering data for the identification of the most possible cross-sectional organization of $NtQ_{42}P_{10}$ fibrils.

List of References.

1. Perutz, M., et al., *Amyloid fibers are water-filled nanotubes.* Proc. Natl. Acad. Sci., 2002(99): p. 5591-5595.

2. Svergun, D., et al., *Protein hydration in solution: experimental observation by x-ray and neutron scattering.* Proc. Natl. Acad. Sci, 1998(95): p. 2267-2272.

3. Chen, S., et al., *Observation of fragile-to-strong dynamic crossover in protein hydration water.* Proc. Natl. Acad. Sci, 2006(103): p. 9012-9016.

4. Arakawa, T. and T. SN, *The stabilization of proteins by osmolytes.* Biophys. J., 1985(47): p. 411-414.

5. Courtenay, E., et al., *Vapor pressure osmometry studies of osmolyte-protein interac- tions: implications for the action of osmoprotectants in vivo and for the interpretation of "osmotic stress" experiments in vitro.* Biochemistry, 2000(39): p. 4455-4471.

6. Colombo, M., D. Rau, and V. Parsegian, *Protein solvation in allosteric regulation.: a water effect on hemoglobin.* Science, 1992(256): p. 655-659.

7. Reid, C. and R. RP, *Probing protein hydration and conformational states in solution.* Biophys. J, 1997(72): p. 1022-1030.

8. Eisenberg, H., *Biological Macromolecules and Polyelectrolytes in Solutions.* Oxford University Press, New York, 1976.

9. Eisenberg, H., *Protein and nucleic acid hydration and cosolvent interactions: establishment of reliable baseline values at high cosolvent concentrations.* Biophys. Chem, 1994(53): p. 57-68.

10. Stanley, C., et al., *Protein Structure and Hydration Probed by SANS and Osmotic Stress.* Biophys J, 2008(94): p. 2777-2789.

11. Jacrot, B., *The study of biological structures by neutron scattering from solution.* Rep. Prog. Phys., 1976(39): p. 911-953.

12. Harries, D., D. Rau, and V. Parsegian, *Solutes probe hydration in specific association of cyclodextrin and adamantane.* J. Am. Chem. Soc, 2005(127): p. 2184–2190.

13. Stanley, C. and D. Rau, *Preferential hydration of DNA: the magnitude and distance dependence of alcohol and polyol interactions.* Biophys. J, 2006(91): p. 912-920.

14. Parsegian, V., R. Rand, and D. Rau, *Osmotic stress, crowding, preferential hydration, and binding: a comparison of perspectives.* Proc. Natl. Acad. Sci., 2000(97): p. 3987–3992.

VITA

Tatiana Perevozchikova was born in Pershotravensk, Ukraine in March 1981. In December 2007, Tatiana received her Bachelor of Science in Microbiology with a concentration in Computational Bacterial Genomics from University of Tennessee, Knoxville. After graduation, she began her graduate studies in the Genome Science and Technology Program at the University of Tennessee, where she obtained her Doctorate of Philosophy Degree in Life Sciences.