AGGREGATE POLYMORPHISM IN PROTEIN DEPOSITION DISEASES: INVESTIGATIONS BY MAGIC ANGLE SPINNING SOLID STATE NMR AND TRANSMISSION ELECTRON MICROSCOPY

by

Jennifer Christine Boatz

B.S., Indiana University Bloomington, 2013

Submitted to the Graduate Faculty of

The School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Jennifer Christine Boatz

It was defended on

October 23, 2018

and approved by

Dr. Angela M. Gronenborn, UPMC Rosalind Franklin Professor and Department Chair, Department of Structural Biology

Dr. Sunil Saxena, Professor and Department Chair, Department of Chemistry

Dr. Tatyana Polenova, Professor, Department of Chemistry and Biochemistry, University of Delaware

Dissertation Director: Dr. Patrick C. A. van der Wel, Associate Professor, Department of Structural Biology

AGGREGATE POLYMORPHISM IN PROTEIN DEPOSITION DISEASES: INVESTIGATIONS BY MAGIC ANGLE SPINNING SOLID STATE NMR AND TRANSMISSION ELECTRON MICROSCOPY

Jennifer Christine Boatz, Ph.D.

University of Pittsburgh, 2018

The deposition of normally soluble protein can occur in any organ in the human body and is associated with tissue dysfunction, cell death, and the progression of disease. Protein aggregation is concomitant with blindness as an outcome of cataract, life-threatening organ failure as a consequence of amyloidosis, and pronounced degeneration of the brain.

The mutation responsible for Huntington’s disease causes an expansion of the polyglutamine domain of huntingtin exon 1 that directly promotes misfolding and refolding of huntingtin and huntingtin N-terminal fragments into amyloid-like fibrils in the basal striatum and cortex of the brain. Several fibril polymorphs have been identified, however the relationship between neurotoxicity and amyloid polymorphism is poorly understood. The P23T mutant of gamma-D-crystallin is associated with cataract formation in the eyes of very young children. Crystallins have been shown to form amyloid-like, native-like, as well as amorphous looking aggregates *in vitro*, accordingly it is unclear which class of aggregates P23T gamma-D-crystallin is most likely to form in cataract. Apolipoprotein A-I is a known anti-atherosclerotic factor and oxidation at methionine residues enhances its function. However, this oxidation also induces aggregation in vascular amyloidosis, which is interlinked with atherosclerosis progression. It is unclear whether apolipoprotein A-I aggregates misfold into amyloid-like fibrils as is usually the case in amyloidosis.

Magic angle spinning solid state NMR (MAS ssNMR) is ideally suited to provide atomic resolution information on the structure and dynamics of insoluble, non-crystalline protein aggregates. Transmission electron microscopy (TEM) allows for the visualization of morphological features of aggregates that cannot be observed by optical microscopy and can be used to identify polymorphs and aid in distinguishing between different classes of aggregates. In this dissertation, I use both MAS ssNMR and TEM in addition to other biophysical and structural techniques to investigate the differences in structure and dynamics between polymorphs of huntingtin exon 1, P23T gamma-D-crystallin, and apolipoprotein A-I. Enabled by my experiments, I narrow down the potential molecular mechanisms involved in these three distinct types of protein deposition diseases. I show that depending on the milieu, proteins have the potential for varied amyloidogenic and non-amyloidogenic self-assembly.

TABLE OF CONTENTS

TABLE OF CONTENTS VI

LIST OF TABLES XIII

LIST OF FIGURES XIV

PREFACE XVIII

DEDICATION XX

LIST OF PUBLICATIONS XXI

LIST OF ABBREVIATIONS XXIII

1.0 INTRODUCTION 1

- 1.1 PROTEIN DEPOSITION IN HUMAN DISEASE 1
- 1.2 BASICS OF PROTEIN STRUCTURE 2
- 1.3 CLASSES OF PROTEIN AGGREGATES 3
  - 1.3.1 Amyloid and amyloid-like fibrils 4
  - 1.3.2 Aggregates with a conserved native-like structure 7
  - 1.3.3 Amorphous aggregates 8
  - 1.3.4 Functional aggregates 9
- 1.4 PROTEIN DEPOSITION DISEASES OF INTEREST 9
  - 1.4.1 Huntington's disease 10
  - 1.4.2 Cataract 12

1.4.2.1 Native structures of α, β, and γ crystallins 13
1.4.2.2 Congenital cataract: P23T hγD crystallin 14
1.4.3 Vascular amyloidosis and atherosclerosis 15
1.5 METHODOLOGIES 15
1.5.1 Magic angle spinning solid state NMR (MAS ssNMR) 16
1.5.1.1 Sample preparation 16
1.5.1.2 Building blocks of MAS ssNMR experiments 18
1.5.1.3 Two dimensional homonuclear experiments 21
1.5.1.4 Heteronuclear experiments 23
1.5.1.5 Secondary structure determination 25
1.5.1.6 Folded structure and beyond 25
1.5.2 Transmission electron microscopy (TEM) 26
1.5.2.1 Negative stain transmission electron microscopy 26
1.5.3 Fourier transform infrared spectroscopy (FTIR) 31
1.5.4 X-ray diffraction 33
1.5.5 Protein production and purification 35
1.5.6 Kinetics of aggregation and amyloid formation 36
1.5.6.1 Sodium dodecyl sulfate polyacrylamide gel electrophoresis 36
1.5.6.2 Electrospray ionization time of flight mass spectrometry 37
1.5.6.3 High performance liquid chromatography (HPLC) based sedimentation assays 39
1.5.6.4 Thioflavin T (ThT) fluorescence 40
1.5.6.5 Amyloid seeding assays 41

1.5.7 Toxicity studies 44

2.0 STRUCTURAL INSIGHTS ON EXPANDED HUNTINGTIN EXON 1 46

2.1 SUMMARY: HUNTINGTIN EXON 1 FIBRILS FEATURE AN INTERDIGITATED β-HAIRPIN-BASED POLYGLUTAMINE CORE 46

2.1.1 Proposed stochastic formation of β-hairpins 48

2.2 SUMMARY: FIBRIL POLYMORPHISM AFFECTS IMMOBILIZED NON-AMYLOID FLANKING DOMAINS OF HUNTINGTIN EXON1 RATHER THAN ITS POLYGLUTAMINE CORE 48

2.2.1 Proposed model of htt exon1 fibrils 50

3.0 ROLES OF CONCENTRATION AND THE N-TERMINAL FLANKING DOMAIN IN MUTANT HUNTINGTIN EXON 1 FIBER POLYMORPHISM 52

3.1 INTRODUCTION 54

3.2 METHODS 57

3.2.1 Protein expression and purification 57

3.2.2 Fibril formation 58

3.2.3 Transmission electron microscopy 59

3.2.4 X-ray powder diffraction 60

3.2.5 Thioflavin T fluorescence and seeding assay 60

3.2.6 Cleavage kinetics assays by SDS-PAGE and ESI-TOF MS 61

3.2.7 Potential trypsin cleavage products of MBP-Q44-httEx1 62

3.2.8 Mass Spectrometry 62

3.2.9 Magic Angle Spinning solid state NMR spectroscopy 63

3.3 RESULTS 64

3.3.1 Concentration of MBP-Q44-httEx1 affects width of mature fibrils 64

3.3.2 Kinetics of Q44-httEx1 monomer release by Factor Xa protease 66

3.3.3 The rate of Q44-httEx1 release from MBP-Q44-httEx1 affects sample heterogeneity 67

3.3.4 Heterogeneity between and within Q44-httEx1 fibrils 71

3.3.5 Trypsin protease releases ΔN15-Q44-httEx1 from MBP-Q44-httEx1 with high activity 72

3.3.6 Aggregates of ΔN15-Q44-httEx1 fibrils 75

3.3.7 Comparison between Q44-httEx1 and ΔN15-Q44-httEx1 fibrils using MAS ssNMR 76

3.3.8 Dynamics of the PRD of ΔN15-Q44-httEx1 fibrils.................................. 80

3.4 DISCUSSION 82

3.4.1 Threshold concentration of Q44-httEx1 monomer necessary to induce the production of thick Q44-httEx1 fibrils 82

3.4.2 Comparison of ΔN15-Q44-httEx1 and Q44-httEx1 fibrils...................... 82

3.4.3 The effect of monomer polymorphism on fibril polymorphism 84

3.4.4 Fibril polymorphism may be explained by secondary nucleation events on the fibril surface 85

3.4.5 Conclusion 88

3.5 ACKNOWLEDGEMENTS 88

3.6 AUTHOR CONTRIBUTIONS.. 89

4.0 CATARACT-ASSOCIATED P23T γ-D-CRYSTALLIN RETAINS A NATIVE-LIKE FOLD IN AMORPHOUS-LOOKING AGGREGATES FORMED AT PHYSIOLOGICAL PH 90
4.1 INTRODUCTION 92
4.2 METHODS 95
4.2.1 Expression and purification of P23T hγD 95
4.2.2 Preparation of P23T hγD aggregates 97
4.2.3 Transmission electron microscopy 98
4.2.4 X-ray powder diffraction 99
4.2.5 Thioflavin T fluorescence measurements 99
4.2.6 MAS ssNMR spectroscopy 100
4.2.7 Synthetic MAS ssNMR spectra 102
4.3 RESULTS 103
4.3.1 Polymorphic aggregation of P23T hγD 103
4.3.2 The amorphous-looking aggregates have high internal order 104
4.3.3 Solid-state NMR of the acid-induced fibrils 107
4.3.4 Detection of dynamic regions by ssNMR 108
4.3.5 Secondary structure content of the aggregates 111
4.3.6 Spectral modeling of misfolded and native-like aggregates 113
4.3.7 Acid-induced P23T hγD fibrils lack native-like structure 114
4.3.8 The amorphous aggregates have a native-like structure 115
4.4 DISCUSSION 119
4.5 ACKNOWLEDGMENTS 125

4.6 AUTHOR CONTRIBUTIONS 126

5.0 MAGIC ANGLE SPINNING SOLID STATE NMR STUDIES OF METHIONINE OXIDIZED APOLIPOPROTEIN A-I AGGREGATES 127

5.1 INTRODUCTION 129

5.2 METHODS 131

5.2.1 Preparation of apoA-I samples 131

5.2.2 Thioflavin T fluorescence assays 132

5.2.3 Fourier transform infrared (FTIR) spectroscopy 132

5.2.4 Transmission electron microscopy 133

5.2.5 MAS solid state NMR spectroscopy 133

5.3 RESULTS 134

5.3.1 Amyloid formation occurs in non-oxidized apolipoprotein A-I upon seeding with pre-formed amyloid fibrils 134

5.3.2 Non-oxidized apolipoprotein A-I adopts a partial β-sheet structure upon seeding with pre-formed oxidized amyloid seeds 137

5.3.3 Seeded non-oxidized apolipoprotein A-I aggregates retain a partial native structure 138

5.4 DISCUSSION 143

5.5 ACKNOWLEDGEMENTS 145

5.6 AUTHOR CONTRIBUTIONS 145

6.0 DISSERTATION SUMMARY AND THE BIG PICTURE 147

APPENDIX A 149

A.1. SUPPORTING INFORMATION FOR CHAPTER 3 149

A.2. TRYPSIN CLEAVAGE PRODUCTS OF MBP-Q44-HTTEX1 AND FREE MBP 157

APPENDIX B 165

B.1. SUPPORTING INFORMATION FOR CHAPTER 4 165

APPENDIX C 171

C.1. SUPPORTING INFORMATION FOR CHAPTER 5 171

BIBLIOGRAPHY 174

LIST OF TABLES

Table A.1. Observed trypsin cleavage products of MBP-Q44-httEx1. 162

Table A.2. Observed trypsin cleavage products of MBP-Q44-httEx1 by ESI-TOF MS prior to aggregation of the ΔN15-Q44-httEx1 monomer.. 163

Table A.3. Observed trypsin cleavage products of MBP-Q44-httEx1 by ESI-TOF MS following aggregation of ΔN15-Q44-httEx1 into mature fibrils. 163

Table A.4. Experimental conditions for MAS ssNMR experiments. 164

Table B.1. Detailed experimental conditions of the MAS ssNMR experiments on P23T hγD crystallin... 168

Table B.2. Chemical shifts of residues assigned in amorphous P23T hγD aggregates at physiological pH determined by MAS ssNMR 169

LIST OF FIGURES

Figure 1.1. Secondary and tertiary structure of proteins 2

Figure 1.2. Kinetics of amyloid formation 7

Figure 1.3. Model structure of huntingtin exon 1 11

Figure 1.4. Native structures of eye lens crystallins 13

Figure 1.5. Native states of wild type and P23T hγD crystallin 14

Figure 1.6. Schematic of MAS ssNMR sample preparation and experimentation 17

Figure 1.7. Process for packing ssNMR 3.2 mm rotors 18

Figure 1.8. 1D MAS ssNMR spectra of httEx1 fibrils 20

Figure 1.9. MAS ssNMR results on an amyloid forming acyl-peptide chimera assembly 22

Figure 1.10. Heteronuclear NMR 24

Figure 1.11. Negative stain TEM images of biological samples. 27

Figure 1.12. Huntingtin exon1 fibril formation monitored by TEM. 28

Figure 1.13. TEM data comparing polyQ-expanded htt exon1 fibrils formed at 22 and 37 °C ... 29

Figure 1.14. Fourier transform filtering of negatively stained httEx1 fibrils 30

Figure 1.15. Bundling of Q44-httEx1 fibrils. 31

Figure 1.16. FTIR on httEx1 fibril polymorphs 32

Figure 1.17. PolyQ amyloid structure 34

Figure 1.18. Huntingtin exon 1 sequence and domain structure 36

Figure 1.19. Kinetics of httEx1 release from fusion construct monitored by SDS-PAGE........... 37

Figure 1.20. MBP release from MBP-Q44-httEx1 fusion protein monitored by ESI-TOF MS... 38

Figure 1.21. ThT fluorescence of httEx1 monomer and mature fibrils 41

Figure 1.22. PolyQ protein recruitment through seeding 42

Figure 1.23. Seeding assay results on htt^{NT} $Q_{23}P_{10}K_2$ peptide... 43

Figure 1.24. Treated and untreated human dopaminergic neurons.. 45

Figure 2.1. Schematic proposed model of htt exon1 fibrils... 51

Figure 3.1. Huntingtin exon 1 and MBP based fusion protein 55

Figure 3.2. Dependence of Q44-httEx1 fibril width on the starting concentrations of MBP-Q44-httEx1 prior to cleavage by FXA at 37 °C... 65

Figure 3.3. Classification of fibril polymorphs by average width in relation to the initial concentration of MBP-Q44-httEx1 and the Fusion Protein to Protease ratio.............................. 69

Figure 3.4. Fibril heterogeneity as observed by TEM 70

Figure 3.5. TEM images of Q44-httEx1 fibrils formed at 28.6 μM MBP-Q44-httEx1 and Fusion Protein to Protease ratio of 62.5:1... 71

Figure 3.6. Trypsin protease cleavage of MBP-Q44-httEx1. 73

Figure 3.7. TEM of [U-^{13}C-^{15}N] ΔN15-Q44-httEx1 fibrils 76

Figure 3.8. Comparison of ΔN15-Q44-httEx1 and Q44-httEx1 by Magic Angle Spinning solid state NMR 78

Figure 3.9. Hypothetical models of Q44-httEx1 fibril growth 86

Figure 4.1. Structure of P23T hγD... 93

Figure 4.2. P23T hγD aggregation 98

Figure 4.3. Characterization of aggregates formed at 37° C under different pH conditions 105

Figure 4.4. MAS ssNMR spectra of uniformly ^{13}C,^{15}N-labeled P23T hγD aggregates formed at pH 7 106

Figure 4.5. MAS ssNMR spectra of U-^{13}C,^{15}N-labeled P23T hγD aggregates formed at pH 3 . 108

Figure 4.6. MAS ssNMR spectra of mobile residues within aggregated P23T hγD 110

Figure 4.7. NMR chemical shift analysis of residue-specific secondary structure 112

Figure 4.8. Synthetic ^{13}C-^{13}C ssNMR spectra predicted for aggregates containing either typical amyloid-like misfolded proteins or natively folded P23T hγD 114

Figure 4.9. Experimental ^{13}C-^{13}C ssNMR spectrum of amorphous P23T hγD aggregates annotated with the liquid-state NMR resonances of P23T hγD in solution 116

Figure 4.10. Mapping of NMR-detected structural changes and similarities 118

Figure 4.11. Proposed potential aggregation mechanisms 122

Figure 5.1. Seeding of non-oxidized wt-ApoA-I with molar 10% and 1% of pre-formed H_2O_2-wt-ApoA-I aggregates 135

Figure 5.2. TEM analysis of H_2O_2-ApoA-I seeds and seeded non-oxidized apoA-I aggregates 136

Figure 5.3. TEM analysis of non-oxidized apoA-I in the absence of H_2O_2-ApoA-I seeds 137

Figure 5.4. FTIR spectra in the amide I region (1602-1708 cm^{-1}) of non-oxidized wt-ApoA-I after seeding with pre-formed H_2O_2-wt-ApoA-I aggregates 138

Figure 5.5. MAS ssNMR spectral analysis of pelleted [U-^{13}C,^{15}N] wt-ApoA-I aggregates 139

Figure 5.6. Comparison of experimental MAS NMR 2D ^{13}C-^{13}C spectrum of H_2O_2 wt-^{13}C,^{15}N-ApoA-I aggregates to a simulated spectrum based on the X-ray determined α-helical native structure of C-terminal truncated Δ(185-243)-ApoA-I 142

Figure 5.7. Schematic of the consequences of the destabilization of self-associated apolipoprotein A-I 144

Figure A.1. Q44-httEx1 fibril dimensions. 149

Figure A.2. Cleavage kinetics of MBP-httEx1 fusion proteins by FXA protease. 150

Figure A.3. ThT Fluorescence following cleavage with trypsin and FXA. 151

Figure A.4. Distribution of fibril widths visible by TEM for Q44-httEx1 fibrils formed at 37 °C, varying concentrations of MBP-Q44-httEx1, and varying molar ratios of MBP-Q44-httEx1 to FXA protease. 152

Figure A.5. Branching and bundling of Q44-httEx1 fibrils. 153

Figure A.6. PDSD spectrum of ΔN15-Q44-httEx1. 154

Figure A.7. Polarization transfer curves. 155

Figure A.8. Hypothetical flanking domain distribution and implications for secondary nucleation 156

Figure A.1. Cleavage products of Q44-httEx1 by trypsin protease. 158

Figure A.2. Trypsin cleavage of free MBP. 160

Figure B.1. MAS ssNMR chemical shift assignments of amorphous P23T hγD aggregates formed at physiological pH 165

Figure B.2. MAS ssNMR chemical shift differences between amorphous P23T hγD aggregates and the simulated amyloid-like model 167

Figure C.1. Full 1D MAS ssNMR spectra of [U-^{13}C,^{15}N] apoA-I 171

Figure C.2. MAS ssNMR spectral comparison between a replicate H_2O_2 wt-^{13}C,^{15}N-ApoA-I sample and the seeded intact wt-^{13}C,^{15}N-ApoA-I aggregates 172

Figure C.3. Amyloid-propensity prediction results. 173

PREFACE

I am incredibly thankful to Drs. Patrick C. A. van der Wel and Angela M. Gronenborn, not only for their mentorship throughout my graduate career but also for the exceptional kindness and understanding they showed me during a difficult transition phase of my life. Patrick was always eager to offer his advice both towards my research and professional development, and he often encouraged me to pursue extracurricular activities focused on science outreach and student government. I owe heartfelt gratitude to Angela, who provided me several amazing opportunities and taught me the importance of transdisciplinary science. I also thank my dissertation committee, research collaborators, and several members of the Van der Wel and Gronenborn labs, especially Drs. Matthew Whitley and Cody Hoop. All these individuals taught me essential skills that were vital towards my research and the completion of my PhD requirements.

I am thankful to the Biomedical Graduate Student Association, which provided me an outlet to develop my interpersonal, organizational, and leadership skills. The BSGA made me feel included and helped me realize I can have an impact on policy. It is still unbelievable to me that my involvement in the BGSA directly resulted in the installation of monthly BGSA travel awards, a comprehensive list of mental health resources on official university webpages, and an overhaul in how the Graduate Studies Office mass distributes information to students. None of that would have been possible without the encouragement of Lauren Zielinski, Drs. John Horn and Joyce D’Antonio, and previous BGSA presidents Drs. Greg Logan, Aliyah Weinstein, and Abby Stahl.

I am thankful to ARCS Pittsburgh Chapter and especially Ann Fromm and Bill Payne for their support and wonderful company. I am grateful to several people at Indiana University who encouraged me to pursue my PhD. In no particular order, I thank: Drs. George Rebec, Joel Ybe, Sarah Fontaine, Ana Estrada-Sánchez, Collin Hobbs, Martha Oakley, Todd Stone, Yan Yu, and the late Dr. Karlijn Keijzer. I also thank my dear friends Lisa Clark, Yoojin Lee, Karen Randolph, Erin Cox, and Patrick Pottorff, my RCIA sponsor Mary Bosack, and everybody at the Upper Room. I would like to thank my family for their fundamental support, especially my mother, late father, brother David and grandma Joan. Finally, I owe heartfelt thanks to my husband Dr. Abhishek Mandal, who brings balance to my life and brings out the best in me. Ever since my interview weekend at the University of Pittsburgh, Abhishek has been there to offer his hospitality, insight, compassion, humor, and encouragement.

DEDICATION

I dedicate this dissertation to my grandmother and “favorite big kid” Shirley, whose fight with an atypical parkinsonian condition inspired me to pursue research on protein aggregation in human disease.

LIST OF PUBLICATIONS

Over the course of my Ph.D. research, I contributed to the following published works:

Hoop, C. L., Lin H. K., Kar K., Magyarfalvi, G., Lamley, J. M., **Boatz, J. C.**, Mandal, A., Lewandowski, J. R., Wetzel, R., van der Wel, P. C. A. *Huntington exon 1 fibrils features an interdigitated β-hairpin-based polyglutamine core*, PNAS 113(6):1546-1551. Copyright (2016) Hoop, C. L., et. al. [1]

Merg, A. D., **Boatz, J. C.**, Mandal A., Zhao, G., Mokashi-Punekar, S., Liu, C., Wang, X., Zhang, P., van der Wel, P. C. A., Rosi, N. L. *Peptide-Directed Assembly of Single-Helical Gold Nanoparticle Superstructures Exhibiting Intense Chiroptical Activity*, J. Am. Chem. Soc. 138(41): 13655-13663. Copyright (2016) American Chemical Society. [2]

Lin, H. K., **Boatz, J. C.**, Krabbendam, I. E., Kodali, R., Hou, Z., Wetzel, R., Dolga, A. M., Poirier, M. A., van der Wel, P. C. A. *Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core,* Nature Communications 8:15462. Copyright (2017) Lin, H. K., et. al. [3]

Mandal, A., **Boatz, J. C.**, Wheeler, T., van der Wel, P. C. A. *On the use of ultracentrifugal devices for routine sample preparation in biomolecular magic-angle-spinning NMR*, J. Biomol. NMR 67(3):165-178. Copyright (2017) Springer Science + Business Media Dordrecht. [4]

Boatz, J. C., Whitley, M. J., Li, M., Gronenborn, A. M., van der Wel, P. C. A. *Cataract-associated P23T γD-crystallin retains a native-like fold in amorphous-looking aggregates formed at physiological pH*, Nature Communications 8:15137. Copyright (2017) Boatz, J. C., et al. [5]

Witkowski, A., Chan, G. K. L, **Boatz, J. C.**, Li, N. J., Wong, J. C., van der Wel, P. C. A., Cavigiolio, G. *Methionine oxidized apolipoprotein A-I at the crossroads of HDL biogenesis and amyloid formation*, The FASEB Journal 32(6):3149-3165. Copyright (2018) Witkowski, A., et. al. [6]

Smith, A., Märker, K., Piretra, T., **Boatz, J. C.**, Matlahov, I., Kodali, R., Hediger, S., van der Wel, P. C. A., De Paëpe, G. *Structural fingerprinting of protein aggregates by DNP-enhanced solid-state NMR at natural isotopic abundance*, JACS 140(44):14576-14580. Epub ahead of print 2018 Oct 26. Copyright (2018) American Chemical Society. [7]

LIST OF ABBREVIATIONS

AD	Alzheimer's disease
ApoA-I	Apolipoprotein A-I
CANCO	^{13}Cα to ^{15}N magnetization transfer followed by transfer to ^{13}CO
CCD	Charged-coupled device
CNS	Central nervous system
CONCA	^{13}CO to ^{15}N magnetization transfer followed by transfer to ^{13}Cα
CP	Cross polarization
CSI	Chemical shift indexing
CTD	C-terminal domain
DARR	Dipolar assisted rotational resonance
DTT	Dithiothreitol
ESI-TOF MS	Electrospray ionization time-of-flight mass spectrometry
FPLC	Fast protein liquid chromatography
FTIR	Infrared and Fourier transform infrared spectroscopy
HbS	Hemoglobin S
HD	Huntington's disease
HDL	High density lipoprotein
HγD	Human gamma-D crystallin

HPLC	High performance liquid chromatography
Htt	Huntingtin protein
HttEx1	Huntingtin exon 1
FXA	Factor Xa
IDP	Intrinsically disordered protein
INEPT	Insensitive nuclei enhanced by polarization transfer
IPTG	Isopropyl beta-D-1-thiogalactopyranoside
IR	Infrared
LB	Lysogeny broth, also known as Luria-Bertani broth
MAS	Magic angle spinning
MBP	Maltose binding protein
MES	2-(N-morpholino)ethanesulfonic acid
MPO	Myeloperoxidase
MS	Mass spectrometry
MW	Molecular weight
NA	Natural abundance
NCA	^{15}N to ^{13}Cα magnetization transfer
NCACX	^{15}N to ^{13}Cα magnetization transfer followed by transfer to CX carbons
NCO	^{15}N to ^{13}CO magnetization transfer
NCOCX	^{15}N to ^{13}CO magnetization transfer followed by transfer to CX carbons
NMR	Nuclear magnetic resonance

NTD	N-terminal domain
OD	Optical density
PBS	Phosphate buffered saline
PD	Parkinson’s disease
PDD	Protein deposition disease
PDSD	Proton driven spin diffusion
PolyQ	Polyglutamine domain
PPII	Polyproline II, a helical structure in protein
PRD	Proline rich domain
PSP	Progressive supranuclear palsy
PTM	Post-translational modification
RC	Random coil
RF	Radiofrequency
RMSD	Root mean squared deviation
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
ssNMR	Solid state nuclear magnetic resonance
TEM	Transmission electron microscopy
TOBSY	Through bond correlation spectroscopy
TPPM	Two-pulse phase-modulated decoupling
UV	Ultraviolet
UV-VIS	Ultraviolet-visible
VA	Vascular amyloidosis
WT	Wild type

1.0 INTRODUCTION

1.1 PROTEIN DEPOSITION IN HUMAN DISEASE

Protein deposition diseases (PDDs) encompass several diseases including cataract, cystic fibrosis, amyloidosis-associated cancers and kidney diseases, and several neurodegenerative diseases [8]. The term amyloidoses is reserved for a class of PDDs that affects organs other than the central nervous system (CNS) [8]. For example, islet amyloidosis affects the pancreas and is found in type II diabetes [8-12]. Hereditary apolipoprotein A-I (apoA-I) amyloidosis can be localized (e.g. to the kidney, liver, etc.) or systemic, however its nonhereditary form is associated with localized plaque buildup in arteries defined as atherosclerosis [13-16]. Neurodegenerative PDDs are associated with deterioration of the CNS and include Alzheimer's (AD), Parkinson's (PD), and Huntington's disease (HD), several dementias, and atypical parkinsonian conditions such as progressive supranuclear palsy (PSP) [8, 17-19]. Some neurodegenerative PDDs are contagious, e.g. kuru, which is transmitted through ritualistic cannibalism, and bovine spongiform encephalopathy (commonly called mad cow disease or variant Creutzfeldt-Jakob disease in humans) which is transmitted through contaminated beef [20-22].

Protein aggregation in tissues is a hallmark of PDDs [8]. In this dissertation, I define *aggregation* as the chaperone-free self-assembly of normally soluble protein into an insoluble, or solid, state [8]. Protein deposition is often associated with misfolding of the native protein structure and refolding into a disease state. This is pathogenic in nature, leading to tissue dysfunction, disease and death. Understanding the mechanisms by which proteins aggregate can aid in structure-based drug design for treatment or prevention of PDDs.

1.2 BASICS OF PROTEIN STRUCTURE

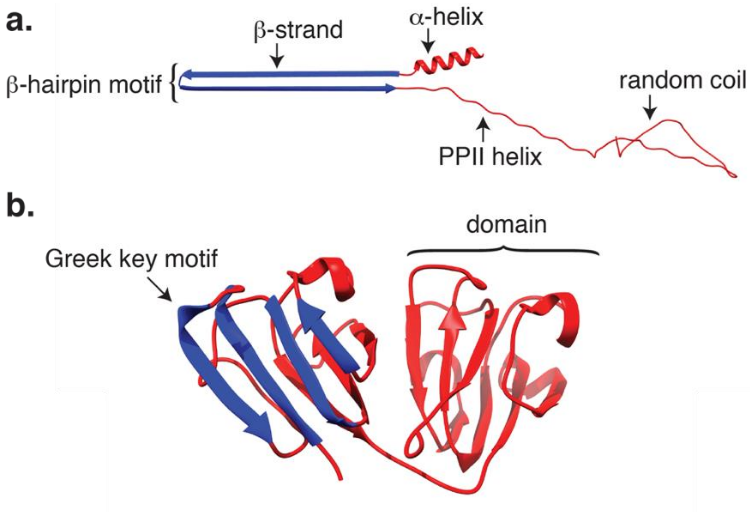


Figure 1.1. Secondary and tertiary structure of proteins

(a) Model structure of amyloid httEx1 that includes a β-hairpin motif (highlighted blue). Secondary structure components are indicated. The model was built in UCSF Chimera [23]. (b) Solution NMR structure of P23T γD crystallin, generated in UCSF Chimera [23, 24]. The protein contains two domains, and each domain contains two Greek key motifs (highlighted blue).

Proteins are translated from RNA as unstructured polypeptide chains that can fold into many different three-dimensional molecular structures. Most proteins are well folded in their functional, or *native state*, however some proteins are natively unfolded: the latter are defined as intrinsically disordered proteins (IDPs) [25]. Protein folding is driven by molecular interactions and often facilitated by molecular chaperones that assist the folding process and prevent aggregation of misfolded protein [26]. Folded proteins are composed of secondary structures such as α-helices and β-sheets that are stabilized through hydrogen bonds and can interact with one another to form supersecondary structures called motifs (e.g. β-hairpins, Figure 1.1, a) and domains (Figure 1.1, b). A domain can remain stable independent from the rest of the protein, however a motif cannot.

1.3 CLASSES OF PROTEIN AGGREGATES

In this section, I present an overview of aggregate morphologies and molecular structures for proteins that are associated with human disease. *Morphology* refers to the appearance of the aggregates on the order of nanometers and micrometers. The morphology of a protein aggregate may be fibrillar or amorphous. Fibrils are threadlike polymeric assemblies that experience periodic growth on one axis. Amorphous aggregates contain molecules that are not arranged into a periodic repeating pattern. Some individuals also categorized crystals as another form of protein aggregates. Crystals contain molecules arranged into a periodic repeating pattern that can grow by the addition of molecules in three dimensions. Fibrils and amorphous aggregates are common in PDDs; crystal formation is rare and has only been reported for immunoglobulins in monoclonal gammopathies [27-30].

An aggregate's morphology is governed by its molecular structure, but the molecular structure cannot be discerned from the morphology alone. The aggregates in PDDs are typically amyloid-like, native-like, or disordered in their molecule structure. Some PDDs are associated with the formation of multiple classes of aggregate structures, and some varieties of proteins (e.g. immunoglobulins) can aggregate into different structure classes in different diseases [8, 9, 31, 32]. It is not uncommon for proteins to aggregate into multiple structure classes within the same disease [8, 9, 31-35].

1.3.1 Amyloid and amyloid-like fibrils

To date, amyloid-forming peptides and proteins have been identified in 70 diseases broadly categorized as either amyloidoses or neurodegenerative amyloid-related PDDs [8]. Neurodegenerative amyloid-related PDDs are associated with amyloid deposition within the central nervous system (CNS), while the term amyloidoses is usually reserved for systemic or localized buildup of amyloid fibrils in tissues other than the CNS [8]. Alzheimer's (AD) and Parkinson's (PD) disease are considered predominantly sporadic neurodegenerative amyloid-related PDDs, although hereditary forms have been documented [8, 17, 18]. In contrast, Huntington's disease (HD) is distinctly hereditary, resulting from a specific mutation of the huntingtin gene (see section 1.4.1) [19].

Many PDDs are associated with the formation of amyloid and amyloid-like fibrils [8]. Amyloid fibrils are characterized by an arrangement of β-sheets parallel to the axis of the fibrils, with the strands arranged perpendicular to the axis [36]. The β-sheets are often arranged in a parallel fashion, but antiparallel arrangements have also been observed [1, 37]. Fibrils are typically microns in length, less than 30 nm in width, and composed of twisting or laterally associating filaments (although fibrils made up of only one filament have been documented) [38]. Many amyloid forming proteins are IDPs in their native state [39]. The amyloid hypothesis proposes that the accumulation of amyloid fibrils or their precursors is directly associated with progression of AD, however this hypothesis is recently under debate [8, 40].

Amyloid fibrils are frequently polymorphic [37, 38]. *Polymorphism* is defined as a state where a protein can form two or more molecular structures, which are often easily distinguished from one another by their morphologies. For example, individual polymorphs of A$\beta_{1\text{-}40}$, an amyloid found in AD, vary in the number of and arrangement of β-sheets and side-chain interactions. This molecular level polymorphism leads to the formation of fibrils that can be distinguished at the atomic and molecular level by magic angle spinning solid state NMR (MAS ssNMR) and at the nanoscopic and microscopic level by transmission electron microscopy (TEM) [41-46]. Amyloid polymorphs can vary not only in their structure but also in their toxicity, and there is increasing evidence that amyloid polymorphism plays a key role in the neuropathology of neurodegenerative diseases [47]. For example, A$\beta_{1\text{-}40}$ fibrils generated *in vitro* from extracts derived from AD brains revealed several polymorphs of A$\beta_{1\text{-}40}$ that differ between patients but not between different areas of individual brains [44].

The formation of amyloid fibrils from soluble protein regularly displays sigmoidal kinetics consisting of three phases: lag, elongation, and plateau (Figure 1.2, a) [8, 48, 49]. The *lag phase* represents the time during which the protein misfolds from its native state and nuclei are formed through a thermodynamically unfavorable process. *Nuclei* are defined as the smallest stable misfolded species from which fibril growth can be directly initiated, a process called *primary nucleation*. Nuclei can form from either monomeric (Figure 1.2, b) [8, 48, 50-52] or multimeric species (Figure 1.2, c) [8, 52-54]. Once nuclei or fibrils exist, they grow rapidly through the addition of monomers to the fiber ends [8]. Fibril growth dominates during the *elongation phase*, which is concluded with a *plateau phase* that represents equilibrium, when most proteins are typically in the fibrillar state [8].

Mature fibrils can act as nuclei to promote the formation of new fibrils through *secondary nucleation* (Figure 1.2, d), which may in some cases appear as *branching* (Figure 1.2, e), multiply through *fragmentation* (Figure 1.2, f). Moreover, fibrils can pack together isotropically to form a gel-like state (Figure 1.2, g), or associate laterally to form bundles with either blunt or frayed ends (Figure 1.2, h) [55, 56]. Both *in vitro* and *in vivo*, pre-existing fibrils can act as '*seeds*' to dramatically shorten the lag phase; accordingly, individual polymorphs can propagate their structure through seeding [44, 47, 57]. Seeding is thought to be responsible for the observation that a single polymorph is observed in samples of A$\beta_{1\text{-}40}$ propagated from varying regions of the same AD brain [44].

Several mathematical models exist to describe amyloid formation kinetics [58]. These models must account for nuclei formation and fibril elongation but may also account for seeding, misfolding, oligomerization and dissociation, conformational rearrangements, secondary nucleation, fragmentation, coagulation, or bundling [8, 48-54, 58-60]. Experimentally, the kinetics of amyloid formation *in vitro* can be monitored with dyes such as thioflavin T (ThT), as discussed in section 1.5.6.4 [61]. Alternatively, kinetics of aggregation can be monitored with methods where the soluble protein and the aggregates are separated through sedimentation or filtering, followed by quantification of the soluble protein. One method available for monitoring aggregation is a high performance liquid chromatography (HPLC) based sedimentation assay, as discussed in section 1.5.6.3 [54, 62, 63].

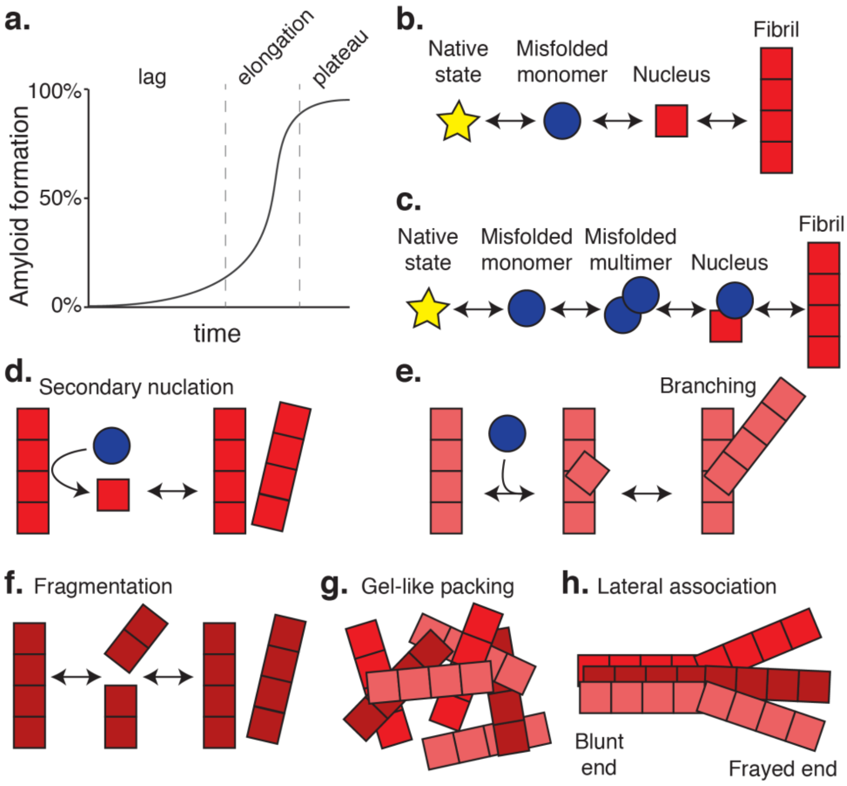


Figure 1.2. Kinetics of amyloid formation

(a) Amyloid formation displays sigmoidal kinetics with three phases: lag, elongation, and plateau. (b) Formation of monomeric nuclei from a misfolded monomer, followed by nucleated polymerization of a fibril. (c) Formation of misfolded multimer followed by nuclei formation through conformational conversion. (d) Secondary nucleation of a misfolded monomer followed by fibril growth. (e) Branching from preexisting fibril. (f) Fragmentation of a fibril followed by fibril growth. (g) Fibril packing into an isotropic gel-like phase, and (h) into bundles through lateral association.

1.3.2 Aggregates with a conserved native-like structure

In many PDDs, the native and functional molecular structure is lost during aggregation. However, there are documented cases where the native structure is mostly retained in the aggregates, e.g. as seen in the nonamyloid fibrils of hemoglobin S (HbS) that are found in sickle cell anemia [33]. The monomers that make up HbS fibrils aggregate through an end-to-end stacking mechanism, reminiscent of the formation of actin and tubulin fibrils [33]. Native-like aggregates can also form through domain swapping. In domain-swapping, domains or motifs swap between molecules, aided by loops and turns that act as hinges [34, 35]. This mechanism has been suggested for aggregation of crystallins, which has implications in cataract (see section 1.4.2) [64-67]. However, recent solution NMR and small angle X-ray scattering results from the Gronenborn lab contest this mechanism [68]. Alternatively, small changes in charged or hydrophobic surface residues can lead to protein aggregation with a preserved native structure via a ‘condensation’ or ‘precipitation’ mechanism that does not rely on substantial misfolding or hinging [69-71]. Natively folded protein *precipitates* are neither fibrillar nor crystalline in nature. Interestingly, eye lens proteins have been observed to phase separate and precipitate in cold-induced temporary cataracts in bovine, rat, and fish lenses and also in living mice [72-75].

1.3.3 Amorphous aggregates

In several PDDs, aggregates that appear amorphous by TEM are often observed in addition to amyloid fibrils [8, 52]. It is usually assumed that amorphous-looking aggregates are poorly structured and disordered. Their appearance is considered nonpathogenic as the result of aggregation that is off-pathway to the formation of amyloid [8, 52, 76]. On the other hand, disordered oligomers that are on-pathway to amyloid fibril formation can have cytotoxic properties [77]. Yet again, off-pathway amorphous aggregates can actually have protective properties in many amyloid PDDs [78, 79]. Therefore, it is desirable to identify and implement therapeutics that redirect aggregation from amyloid fibrils to off-pathway amorphous aggregates [79].

1.3.4 Functional aggregates

It is important to note that not all peptide and protein aggregates are harmful and associated with human disease. For example, pre-melanosomal protein fibrils are important for regulating melanin condensation within melanosomes [80]. It is becoming increasingly recognized that aggregates can be desirable and have functional roles [81-84].

1.4 PROTEIN DEPOSITION DISEASES OF INTEREST

The diseases of interest described in this dissertation illustrate several general aspects of PDDs. Huntington's disease (section 1.4.1) illustrates one of several neurodegenerative diseases where the formation of amyloid fibrils in the brain is directly correlated with disease progression [8, 17-19]. The mechanism(s) of protein aggregation implicated in cataract (section 1.4.2) is under intense debate, and could involve the formation of amyloid fibrils, native-like aggregates, or amorphous aggregates [64-68, 72-75, 85-88]. The role of apolipoprotein A-I aggregation in atherosclerosis (section 1.4.3) is unclear; oxidation at methionine is associated with an increased function and presumably preserves the native structure [89]. However, oxidation also promotes aggregation and aggregates of apolipoprotein A-I have been identified in vascular amyloidosis [89].

1.4.1 Huntington’s disease

Huntington’s disease (HD) is a neurodegenerative disease that causes movement abnormalities. Disease progression is correlated with the aggregation and accumulation of amyloid-like fibrils made up of huntingtin (htt) and htt N-terminal fragments in the cortex and basal striatum of the brain [19]. Huntingtin is a large (~350 kDa) and promiscuous protein that acts as a scaffold for several binding partners that play important roles in cell signaling, membrane trafficking, and transcription regulation [90]. Wild type htt has anti-apoptotic properties, however apoptosis is accelerated in HD brains, resulting in a significant atrophy of cerebral tissue [19, 90].

The autosomal dominant mutation responsible for HD results in an elongation of a polyglutamine (polyQ) domain within the first exon (httEx1) of htt [19]. As shown in Figure 1.3, the polyQ domain of httEx1 makes up an antiparallel β-sheet structure within the core of mature fibrils and is flanked by a highly conserved 17 residue α-helical N-terminal domain (htt^{NT}) and a mixed PPII helix/random coil proline-rich domain (PRD) [1, 91-94]. Expansion beyond 35 glutamine residues destabilizes the polyQ domain and causes htt to misfold, self-assemble, and aggregate; further polyQ expansion is directly correlated with increased aggregation kinetics [95-97]. The mean age of onset occurs at 35 years of age, however this is strongly dependent on the length of the expanded polyQ domain; each addition of a glutamine residue within the polyQ domain is correlated with an earlier age of onset of HD symptoms and extreme cases of juvenile HD have been documented [98].

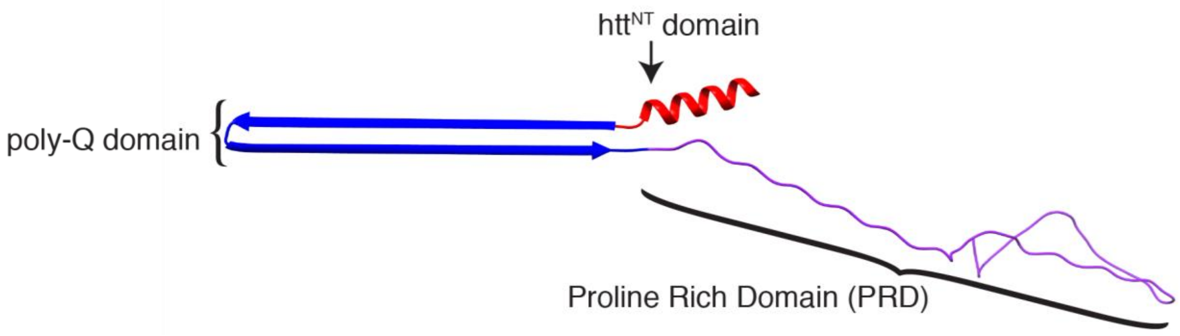


Figure 1.3. Model structure of huntingtin exon 1

Model structure of huntingtin exon 1, built in UCSF Chimera [23]. The htt^{NT} *domain (red) forms an* α*-helix. The polyQ domain (blue) forms an antiparallel* β*-hairpin. The proline rich domain (PRD, purple) has PPII helices and random coil components. This model was built using UCSF Chimera software [23].*

Expanded httEx1 is sufficient to induce HD-like pathology and fibril formation in mouse models [99, 100]. Studies on neuronal cell cultures and mice brain extracts reveal that httEx1 aggregates are polymorphic and exhibit a wide range of neurotoxicity [101, 102]. Additional studies on neuronal cells reveal the formation of prominent inclusion bodies whose contents vary between cell types and can contain amorphous and fibrillar aggregates [103-105]. Model polyQ peptides and httEx1 aggregates are also polymorphic *in vitro* [47, 106]. For example, httEx1 fibrils formed *in vitro* by Dr. Nekooki-Machida and coworkers at 4 and 37 °C respectively were found to exhibit varying levels of cytotoxicity, indicating that changes in the structure and dynamics of the polyQ core and flanking domains have implications in HD cytotoxicity [47]. Although prior studies have attributed the polymorphism found in httEx1 aggregates to structural differences in the polyQ domain, these results are low-resolution and not definitive [47]. Given that fiber formation is related to cytotoxicity and disease progression, it is essential to gain insight into the structure and dynamics of httEx1 fibrils and the molecular mechanism underlying fibril formation and toxicity.

1.4.2 Cataract

Cloudiness or coloring of the eye lens, medically defined as cataract, affects an estimated 22 million (17.1%) Americans over age 40 and another 88 million (68.3%) people over age 80 globally [107-109]. The number of affected individuals is expected to increase by another 39 million by 2030 [107-109]. Surgical replacement of the lens is presently the only treatment for cataracts. Although fairly straightforward, this procedure is not globally available and in the US costs over $3.4 billion/year through Medicare; Thus, cheaper, more readily available, and less intrusive treatments are desirable [109]. A prerequisite for the development of new treatments is a thorough understanding of the mechanism of cataract formation through basic science and structural biology.

The α and βγ crystallins make up 90% of lens proteins and are major constituents of cataracts [110, 111]. The α crystallins are multimeric complexes that act as chaperones, while the βγ crystallins are well-structured proteins that are necessary to maintain the high refractive index of the lens [112, 113]. Crystallins are exceptionally stable and soluble proteins that can withstand many environmental insults and undergo no protein turnover following infancy [70, 114]. The long lifetime of crystallins requires unusually high solubility: bovine crystallins can reach an estimated concentration of 520 mg/mL in the native eye lens nucleus [114, 115]. Crystallins likely have key structural features which allow them to be many times more soluble and stable than most proteins.

It is generally accepted that damaged crystallins become insoluble, as opposed to degraded, in the lens during cataract formation [116, 117]. However, the mechanism by which crystallins associate and form insoluble aggregates is debated [64-68, 85-88]. Recently proposed mechanisms that have been advanced involve amyloid-like fibril formation or the formation of highly-ordered domain-swapped aggregates [64-68, 85, 86]. Neither model is exclusive and additional models exist including the above-mentioned condensation based self-assembly observed in cold cataract [72-75, 87, 88].

1.4.2.1 Native structures of α, β, and γ crystallins

The α crystallins are small proteins that form very large multimers (Figure 1.4, a) [118, 119]. The β and γ crystallins have very similar structures to each other, however the β crystallins (Figure 1.4, b) form homodimers while γ crystallins (Figure 1.4, c) remain monomers in the native state [120, 121]. Each molecule of β and γ crystallins contains two structural domains connected by a short linker, and each domain contains two highly conserved Greek key motifs (see Figure 1.1, b). The N-terminal and C-terminal domains are very similar in sequence and structure [120, 121].

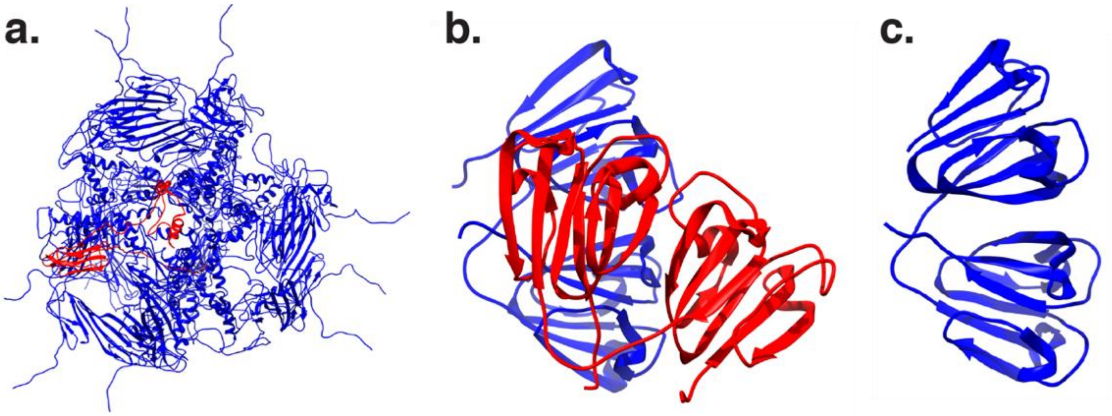


Figure 1.4. Native structures of eye lens crystallins

(a) Model of αB-crystallin as a 24mer [119]. A single monomer is highlighted red. (b) Crystal structure of truncated βB1-crystallin dimer [120]. (c) Crystal structure of WT γD-crystallin [121].

1.4.2.2 Congenital cataract: P23T hγD crystallin

The most abundant crystallin in the human eye lens is hγD crystallin. Accordingly, many mutations of hγD are associated with congenital cataract [122]. The autosomal dominant P23T mutation of hγD affects very young children, and the cataracts that form can be blue to white in color, coralliform (shaped similar to branching coral), aculeiform (needle-shaped, resembling crystals), or silica-like, and can be located in either the lamellar layers or nucleus of the lens [123-129]. P23T hγD is insoluble above 5 mg/mL at 37 °C, considerably less soluble than WT hγD. However, the native structure of P23T hγD is very similar that of WT hγD, so it is not clear why this mutation reduces the solubility of hγD drastically (Figure 1.5, a,b) [24, 130]. The most obvious structural difference involves the rotamers of His 22, which is sequentially next to the mutation site (Figure 1.5, c) [24, 130].

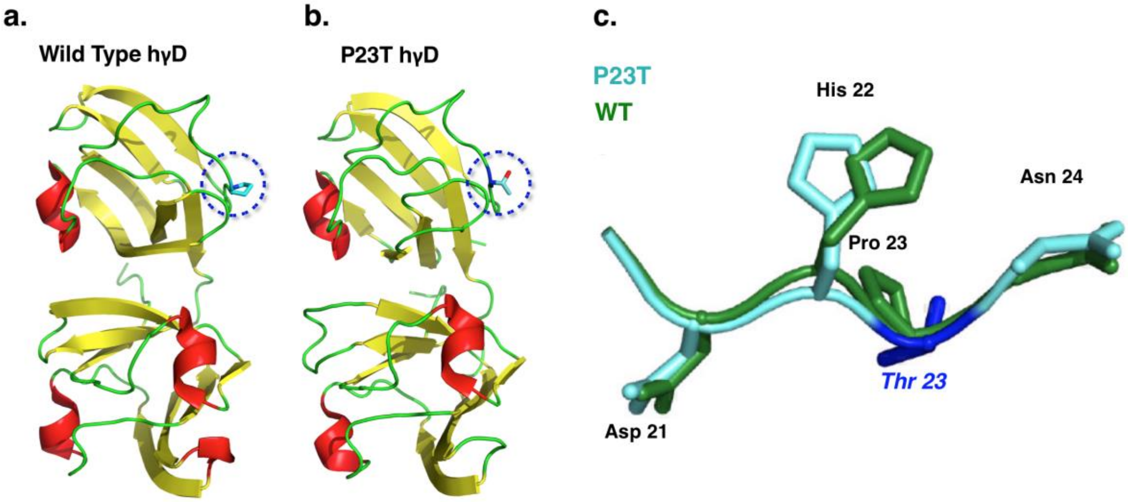


Figure 1.5. Native states of wild type and P23T hγD crystallin

(a) Crystal structure of WT hγD and (b) solution NMR structure of P23T hγD [24, 130]. The mutation site is circled. (c) The His 22 residue experiences small rotamer changes between WT and P23T hγD. This figure was generated using PyMol software [131].

1.4.3 Vascular amyloidosis and atherosclerosis

The deposition of amyloid fibrils in vascular tissue is defined as vascular amyloidosis (VA) [89]. VA is common in aortic tissue and was found to be present in more than 97% of samples from patients over 50 years in age [132]. Atherosclerosis and VA are also highly correlated [89]. Several amyloidogenic proteins accumulate within atherosclerotic arteries, including transthyretin, immunoglobin γ, medin, and in the highest abundance apolipoprotein A-I (apoA-I) [132-142]. The presence of apoA-I in atherosclerotic plaques is in part surprising because apoA-I, the main constituent of high-density lipoproteins (HDL), is known to have anti-atherosclerotic properties [16, 143]. Amyloidogenic apoA-I identified in atherosclerotic lesions was found to be oxidized in at least one, and possibly three, methionine sites [144]. The frequency of apoA-I deposition in VA and atherosclerosis combined with a poor understanding of the process of apoA-I amyloid formation warrants structural and mechanistic studies on methionine oxidized apoA-I. This topic is discussed further in Chapter 5.0.

1.5 METHODOLOGIES

Protein aggregates are insoluble and lack crystallographic order; they are therefore untenable for structural studies by traditional methods such as solution NMR and X-ray crystallography. In this section, I present an overview of some of the methodologies used in this dissertation, including methodologies I used to gain insight into aggregate structures. I provide several examples, primarily from results I gathered on httEx1 from a manuscript on which I was a co-author early in my Ph.D. research [3].

1.5.1 Magic angle spinning solid state NMR (MAS ssNMR)

Magic angle spinning (MAS) solid state nuclear magnetic resonance (ssNMR) is ideally suited to study the structure and dynamics and of insoluble, non-crystalline aggregates at atomic resolution [145]. Its power has been convincingly demonstrated previously for many amyloid-like fibrils and aggregated proteins [1, 37, 44, 92, 146, 147]. In this section, I provide an overview of a few MAS ssNMR concepts used later in this dissertation.

1.5.1.1 Sample preparation

Aggregated protein samples are most commonly analyzed by ^{13}C and ^{15}N MAS ssNMR. These isotopes exist infrequently (1.1% and 0.37%, respectively) in proteins at isotopic natural abundance (NA), therefore protein must be labeled with ^{13}C and ^{15}N isotopes in order to be detected in more than one dimension by MAS ssNMR. I overexpressed MBP-Q44-httEx1 in *Escherichia coli* for several experiments described in this and in the next chapters. For most non-NMR experiments, the *E. coli* are grown in rich medium called lysogeny broth (LB) (also known as Luria-Bertani broth [148]) that contains nutrients at NA. In order to uniformly substitute the ^{12}C and ^{14}N isotopes found at NA, ^{13}C and ^{15}N isotopes are added to modified minimal M9 medium during the protein expression process [149]. The M9 medium itself contains no carbon or nitrogen sources but does contain several additives necessary for *E. coli* growth.

After expression, purification, and aggregation (as described in section 1.5.5), micrograms of protein are packed into small MAS ssNMR rotors (Figure 1.6) using a home-built ultracentrifugal packing tool over several hours (Figure 1.7) and the rotor is then sealed with a finned cap [4]. I found this packing tool extremely advantageous over traditional packing methods using a spatula in part because it ensures dense and near complete sample packing and retains sample hydration [4]. An NMR spectrometer equipped with an MAS ssNMR probe is necessary to perform MAS ssNMR experiments. I primarily used MAS probes capable of detecting ^{1}H, ^{13}C, and ^{15}N at their resonance frequencies, measured as chemical shifts in parts per million (ppm).

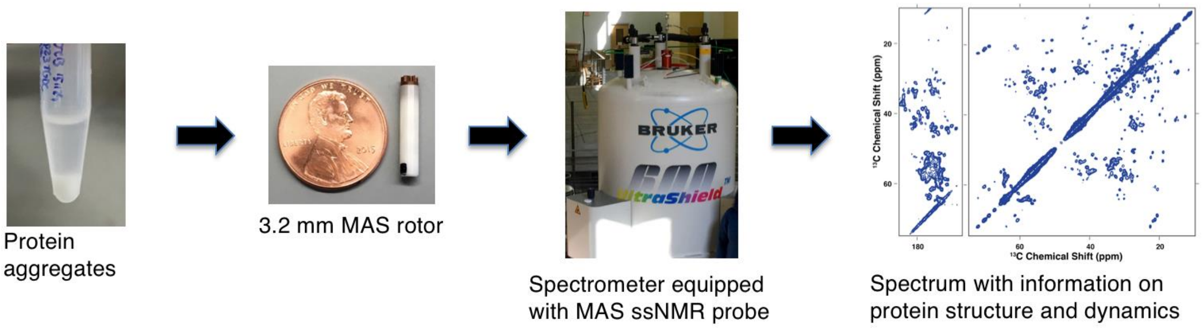


Figure 1.6. Schematic of MAS ssNMR sample preparation and experimentation

Protein aggregates (far left) are packed into small MAS rotors prior to experimentation in an NMR spectrometer equipped with an MAS ssNMR probe. The obtained MAS ssNMR spectrum (far right) contains information on both protein structure and dynamics. I obtained the spectrum shown here from uniformly labeled ^{13}*C,* ^{15}*N P23T hγD.*

Inside the probe the rotor is tilted to precisely 54.7356°, which is referred to as the magic angle. A stream of air spins the rotor fast (e.g. 13 kHz or 13,000 times per second) during acquisition of the MAS ssNMR spectra. The spinning at the magic angle averages out the chemical shift anisotropy that would otherwise be observed in NMR studies of a static solid sample, resulting in isotopically averaged chemical shifts and sharp peaks comparable to what is obtainable by solution NMR [150]. The ^{13}C-^{13}C spectrum I obtained on P23T hγD shown in Figure 1.6 contains peaks that are sensitive to both protein structure and protein dynamics. This spectrum contains many well resolved peaks, indicative of a homogenous sample (see Chapter 4.0).

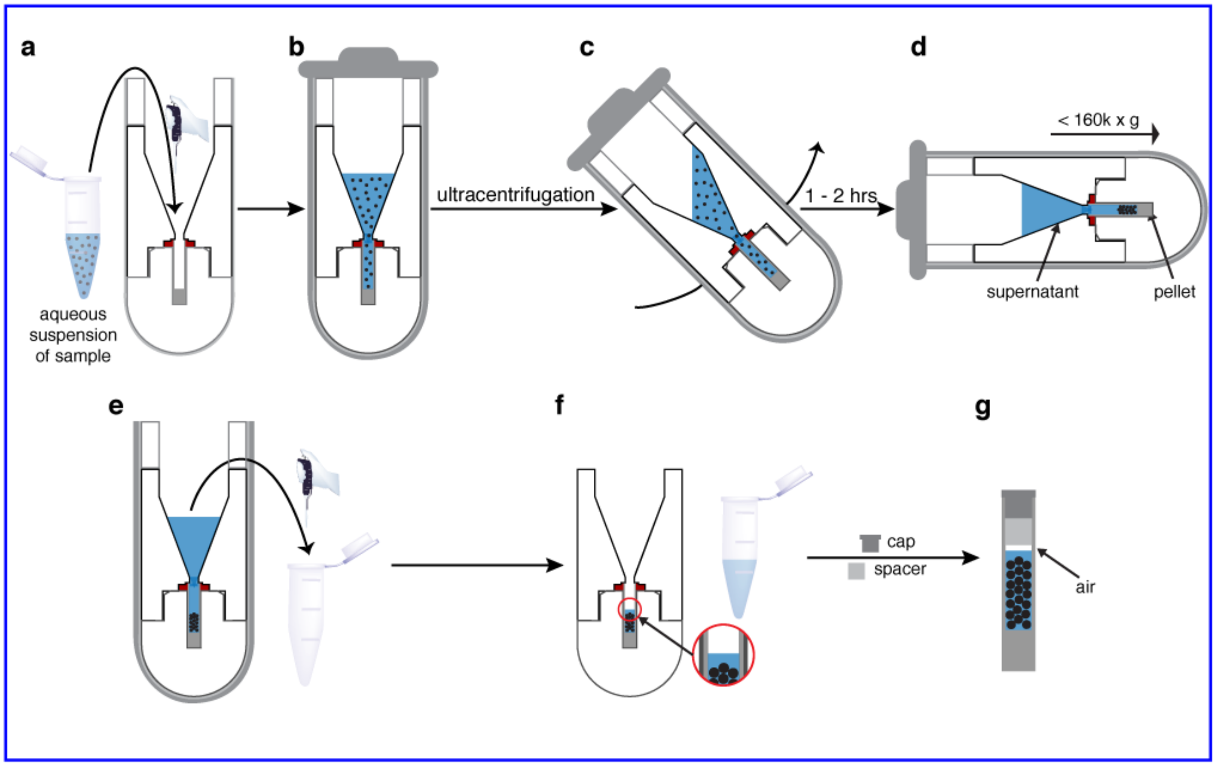


Figure 1.7. Process for packing ssNMR 3.2 mm rotors

(a) The sample suspension is placed in the funnel, up to a 1 mL volume. (b) The packed tool is inserted in its swinging bucket and sealed. (c, d) During ultracentrifugation the bucket achieves a horizontal orientation that ensures even pelleting of the sample within the rotor. (e) After completion of the ultracentrifugation, excess supernatant is removed. (f) A small amount of supernatant is left to maintain excess hydration. (g) After disassembly of the packing tool the rotor is capped. Note that it is important to leave a small gap between the liquid and the spacer or cap. Reprinted with permission from Mandal, A., ***Boatz, J. C.****, Wheeler, T., van der Wel, P. C. A. On the use of ultracentrifugal devices for routine sample preparation in biomolecular magic-angle-spinning NMR, J. Biomol. NMR, 67(3):165-178. Copyright (2017) Springer Science + Business Media Dordrecht.*

1.5.1.2 Building blocks of MAS ssNMR experiments

Many of the ^{13}C and ^{15}N MAS ssNMR experiments I performed over the course of my research were based on either direct polarization (DP), cross polarization (CP) or insensitive nuclei enhanced by polarization transfer (INEPT) experiments. In DP-based experiments, ^{13}C or ^{15}N nuclei are directly polarized or excited by a radiofrequency (RF) pulse. The signal to noise of the peaks in a DP-based spectrum is related to the gyromagnetic ratio of the polarized nuclei [151]. The gyromagnetic ratio of a nucleus, e.g. ^{13}C, is directly proportional to its magnetic strength, or more specifically, to the relationship between its magnetic dipolar moment and angular momentum [152]. The gyromagnetic ratio of ^{1}H is higher than the gyromagnetic ratios of either ^{13}C or ^{15}N, therefore signal to noise can be enhanced by transferring polarization from ^{1}H to ^{13}C or ^{15}N. In the CP-based experiments, RF pulses are applied on two channels: ^{1}H and either ^{13}C or ^{15}N. This allows ^{1}H polarization to be transferred to either ^{13}C or ^{15}N nuclei through heteronuclear dipole-dipolar interactions [153].

Dipolar couplings occur between nuclei that are near each other in space. The strength of dipolar coupling constants depends on the product of the gyromagnetic ratio of each nuclei, and is inversely proportionate to the third power of the distance between the nuclei [154, 155]. Molecular motion can dramatically reduce dipolar couplings, consequently CP-based experiments are used to selectively measure the most rigid residues within protein aggregates. CP is often advantageous over DP; the low sensitivity of ^{13}C and ^{15}N nuclei is enhanced by multiple factors when ^{13}C and ^{15}N polarization is transferred from the highly sensitive ^{1}H nuclei by CP. However, since DP experiments do not depend on dipolar couplings, spectra obtained via DP can capture signals from both rigid and mobile residues.

Like CP-based experiments, in INEPT experiments polarization is transferred from highly sensitive ^{1}H nuclei to nuclei with lower sensitivity [156]. However, polarization is transferred through J-based scalar couplings instead of through dipolar couplings. J-couplings are dependent on the covalent bonds between the donor ^{1}H and acceptor nuclei, while dipolar couplings are not. Contrary to CP-based experiments, INEPT experiments can capture signals from flexible residues that experience similar mobility to soluble proteins [156]. At the same time INEPT experiments are very sensitive to T_2 (also called spin-spin and transverse) relaxation. Rigid residues relax quickly (microseconds), therefore in typical INEPT experiments their polarization is lost prior to acquisition. However, highly flexible residues relax slowly (milliseconds) and are amenable to detection by INEPT-based experiments.

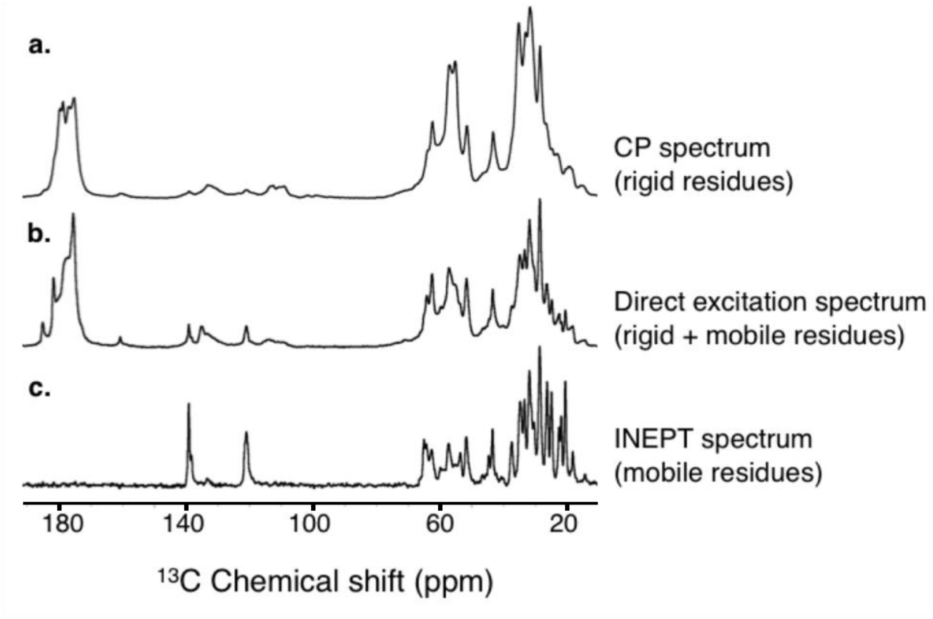


Figure 1.8. 1D MAS ssNMR spectra of httEx1 fibrils

(a) Cross polarization, (b) direct excitation/DP, and (c) INEPT spectra of uniformly labeled ^{13}C, ^{15}N *Q44-httEx1 fibrils. The data were obtained at a set temperature of 277 K and 10 kHz MAS, independently of the results presented in Chapter 1.5 [3].*

An example of 1D ^{13}C DP, CP, and INEPT spectra are shown in Figure 1.8. I obtained these spectra on Q44-httEx1 fibrils at 10 kHz MAS and a set temperature of 277 K. The CP spectrum (Figure 1.8, a) is dominated by glutamine residues found within the rigid amyloid core, while the INEPT (Figure 1.8, c) spectrum selectively reports on mobile residues primarily found in the flanking domains. Direct excitation (DP) of the ^{13}C nuclei results in a spectrum that reports on both rigid and mobile residues (Figure 1.8, b), allowing for a more nuanced comparison between the CP and INEPT spectra. These spectra closely match the spectra of httEx1 fibrils recently reported by others and the Van der Wel lab [3, 157].

1.5.1.3 Two dimensional homonuclear experiments

Insight into protein structures can be gained through looking at two dimensional CP-based spectra, e.g. ^{13}C-^{13}C dipolar assisted rotational resonance (DARR) and proton driven spin diffusion (PDSD) experiments, which selectively report on immobilized residues [158]. In DARR experiments, polarization is transferred through a CP step to ^{13}C, and then is transferred again to nearby ^{13}C nuclei through a process called DARR mixing. Low power ^{1}H irradiation that matches the MAS rate is applied during the mixing period, resulting in rotational resonance recoupling that enhances the transfer between ^{13}C sites [159].

At short DARR mixing times (e.g. 8 ms), correlations between covalently bound ^{13}C nuclei with a bond distance of approximately 1.5 Å are primarily observed, e.g. between Cα and Cβ nuclei within the same residue. Increasing the DARR mixing time (e.g. to 100 ms) allows for additional ^{13}C-^{13}C correlations to be observed that correspond to distances that do not exceed 7 Å [160]. Because dipolar coupling does not depend on chemical bonds, intramolecular correlations (e.g. Cα and Cδ nuclei within the same residue), *and* intermolecular correlations (e.g. two Cα nuclei in closely interacting but different molecules) can be observed [160]. At even longer mixing times, we use PDSD experiments in which ^{1}H irradiation is turned off during the ^{13}C-^{13}C mixing step in order to prevent damage to both the sample and the MAS ssNMR probe.

As already noted, CP-based DARR and PDSD report only on the most rigid residues in aggregated protein samples. In order to instead identify ^{13}C-^{13}C correlations within flexible residues, a 2D INEPT-based experiment with total through bond correlation spectroscopy (TOBSY) mixing based on scalar couplings can be applied [161].

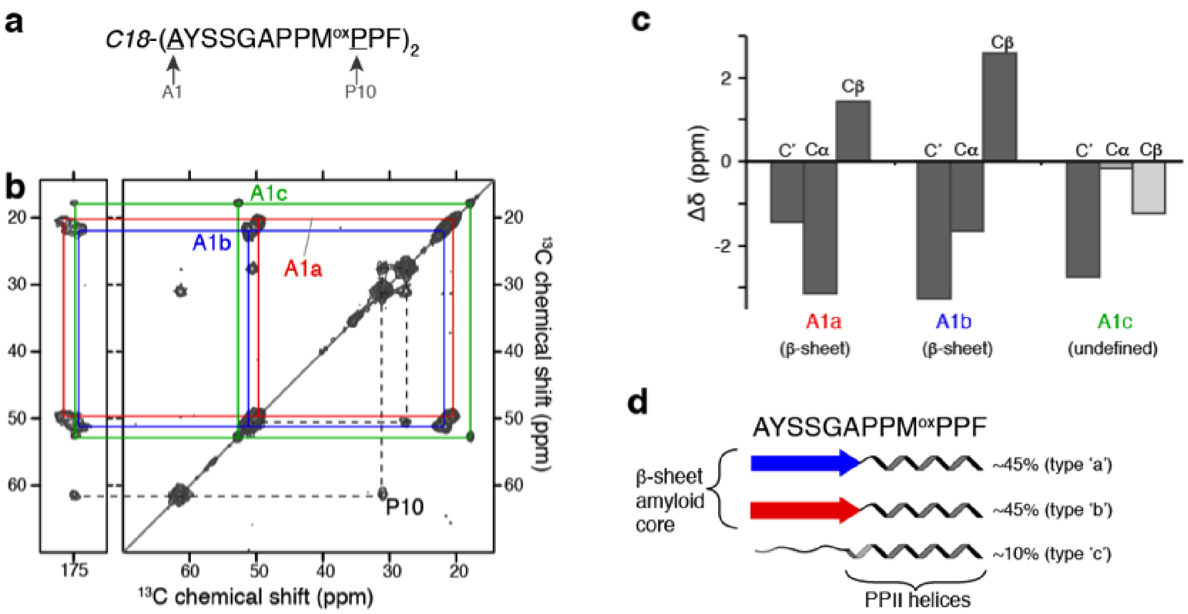


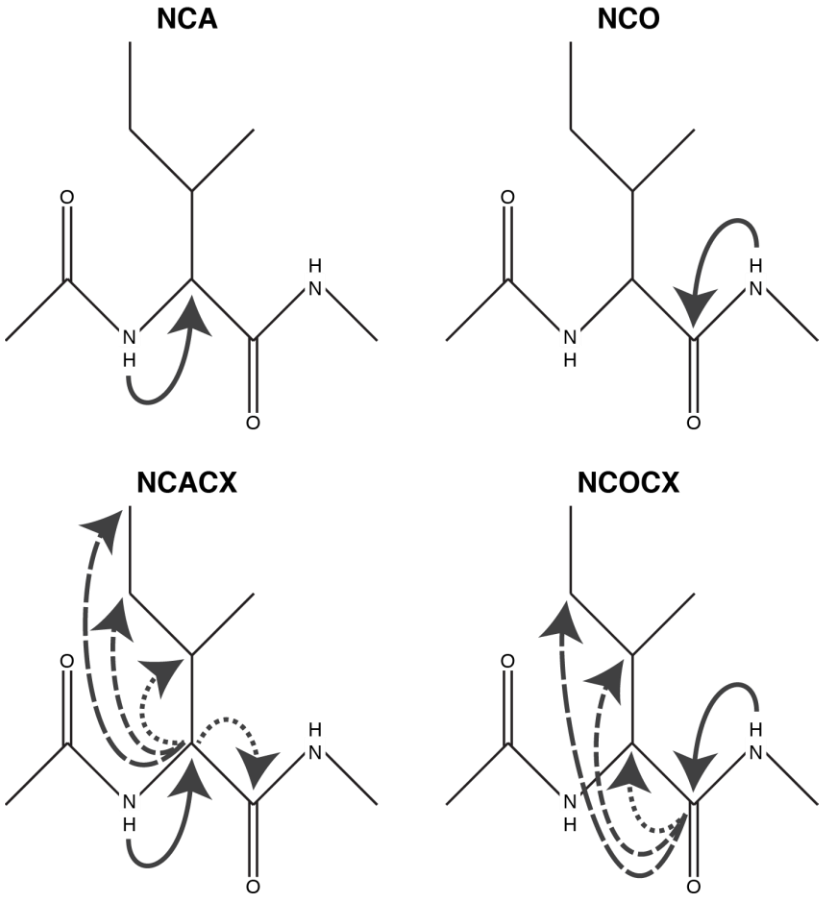
Figure 1.9. MAS ssNMR results on an amyloid forming acyl-peptide chimera assembly

a) Position of residue-specific $^{13}C,^{15}N$*-labeling (arrows). (b) 2D* ^{13}C*-*^{13}C *MAS ssNMR spectrum of labeled* $[C_{18}\text{-}(PEP_{Au})_2]^{M\text{-}Ox}$ *assemblies. Dashed and colored lines connect sets of peaks from labeled P10 (black dashed) and A1 residues (solid lines). Three A1 conformations are marked with red (A1a), blue (A1b), and green (A1c) lines. (c) Secondary structure analysis of A1 ssNMR signals, showing A1a and A1b to be part of the β-sheet core. (d) Secondary structure distribution in the three peptide conformers seen by ssNMR, along with their relative ssNMR peak intensities (right). Modified and reprinted with permission from Merg, A. D.,* ***Boatz, J. C.****, Mandal A., Zhao, G., Mokashi-Punekar, S., Liu, C., Wang, X., Zhang, P., van der Wel, P. C. A., Rosi, N. L. Peptide-Directed Assembly of Single-Helical Gold Nanoparticle Superstructures Exhibiting Intense Chiroptical Activity, J. Am. Chem. Soc. 138(41): 13655-13663. Copyright (2016) American Chemical Society.*

An example DARR spectrum I ran on an amyloid-forming synthetic acyl-peptide chimera (sequence in Figure 1.9, a) is presented in Figure 1.9, b [2]. Isotopic ^{13}C labels were introduced at two residues: Ala1 and Pro10. Interestingly, three peaks are observed that represent correlations between Cα and Cβ in Ala1. The same phenomenon is observed for Cα/CO and Cβ/CO correlations. This indicates that Ala1 exists in three distinct structure types (‘a’, ‘b’, and ‘c’) within the amyloid assembly.

1.5.1.4 Heteronuclear experiments

Two- and three-dimensional heteronuclear MAS ssNMR experiments are essential for assigning ^{13}C and ^{15}N resonances to specific residues within a protein’s sequence. Many heteronuclear MAS ssNMR experiments exist, including NCA, NCO, NCACX, NCOCX, CONCA, and CANCO [162]. In NCA experiments, a ^{1}H-^{15}N CP step is followed by a ^{15}N-^{13}C CP step that selectively transfers polarization from backbone ^{15}N nuclei to the Cα nuclei within the same residue (Figure 1.10, top left) [162]. In NCO experiments, the polarization is instead transferred to the CO carbons in the preceding n-1 residue (Figure 1.10, top right). A DARR mixing period can be added to NCA and NCO experiments to allow mixing from the Cα and CO nuclei to nearby carbons: these experiments are referred to as NCACX (Figure 1.10, bottom left) and NCOCX (Figure 1.10, bottom right), respectively. Alternatively, multiple CP steps can be combined to transfer polarization from the Cα to the backbone ^{15}N to the CO nuclei (CANCO), or to transfer polarization in the reverse direction (CONCA). The process of assigning ssNMR peaks to individual nuclei from such experiments is described in more detail by Hong [163] and Castellani, et. al. [164].



NCA

NCO

NCACX

NCOCX

Figure 1.10. Heteronuclear NMR

During NCA experiments magnetization is transferred from the ^{15}N *amide in the peptide backbone to the* $^{13}C\alpha$ *position within the same residue (top left, solid black arrow), while in NCO experiments (top right) the magnetization is transferred instead to the backbone carbonyl* ^{13}C *in the n-1 position. An addition of a DARR mixing step to the NCA or NCO experiment causes magnetization to transfer from the* $C\alpha$ *or CO to nearby* ^{13}C*; these experiments are referred to as NCACX (bottom left) and NCOCX (bottom right) experiments, respectively. Short DARR mixing times will mostly result in* ^{13}C*-*^{13}C *magnetization transfers over short distances through space, e.g. the equivalent of a 1-bond distance (dotted arrow). Increasing the DARR mixing time allows magnetization to transfer over a longer distance through space, as depicted by the dashed arrows.*

1.5.1.5 Secondary structure determination

A chemical shift indexing (CSI) method can be applied to determine the secondary structure for assigned amino acids [2, 165]. This method subtracts average random coil chemical shifts from the observed chemical shifts for each atom [166]. Atoms that likely exist in a β-sheet secondary structure will display a negative secondary chemical shift difference ($\Delta\delta$); conversely, atoms that exist in an α-helix will have a positive $\Delta\delta$. For example, in in Figure 1.9, c, the CSI results indicate that 'a' and 'b' type Ala1 adopt a β-sheet structure, likely within the amyloid core, while the 'c' type Ala1 adopts an undefined structure (Figure 1.9, d).

A similar phenomenon is observed in polyQ peptides and httEx1 fibrils; two sets of glutamine signals ('a' and 'b') are observed, indicative of two distinct glutamine structures present within the rigid polyQ core in equal amounts [3]. CSI indicates that both 'a' and 'b' type glutamine have a β-sheet secondary structure. A smaller set of signals represents 'c' type glutamine residues that exist outside of the rigid polyQ core.

1.5.1.6 Folded structure and beyond

Although I do not discuss these approaches in Chapters 3 - 5, it is important to recognize that several strategies are available that can lead to three-dimensional structure determination by MAS ssNMR. These strategies include experiments and simulations that can be used to determine several structural restraints including torsion angles and long-range intramolecular and intermolecular contacts [167-173]. MAS ssNMR been successfully applied to solve the folded structure for several proteins and peptides including membrane proteins, protein crystals, and amyloid fibrils [45, 145, 173-184].

In additional to structural information, MAS ssNMR can also be used to probe molecular dynamics. Some of the methods available that yield insight into the dynamics of protein aggregates were discussed above. Many more strategies are available to probe dynamics, including simulations and experiments that utilize anisotropic interactions or spin relaxation to gain insight into dynamics at various timescales and motional amplitudes [185].

1.5.2 Transmission electron microscopy (TEM)

Transmission electron microscopy (TEM) allows for the visualization of protein assemblies that cannot be observed by optical microscopy. A few microliters of sample are sufficient to deposit sample onto a TEM grid, which can then be imaged to obtain thousands of TEM micrographs. Features as small as a few nm can be visualized within a single electron micrograph [186].

1.5.2.1 Negative stain transmission electron microscopy

Negative staining is a method of increasing contrast in electron micrographs. In negatively stained grids, protein samples do not take up stain and therefore appears lighter against a dark background. Electron dense salts are often used as staining agents due to their ability to provide decent contrast and protect the sample from radiation damage [186]. However, a focused beam can rapidly degrade the sample and nonetheless, contrast is low in focused images [186]. Images are usually obtained slightly under focus in order to maintain sample integrity and contrast.

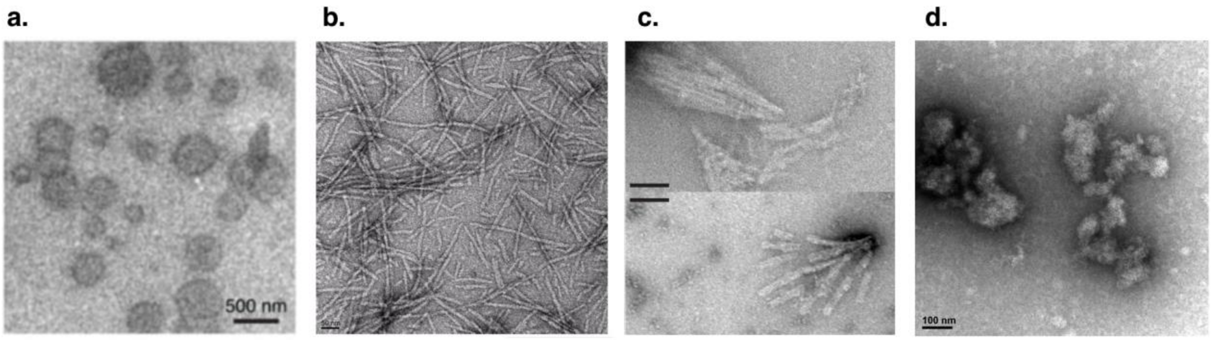


Figure 1.11. Negative stain TEM images of biological samples.

(a) TOCL/DOPC/cholesterol vesicles bound to cytochrome c. Scale bar is 500 nm (b) $K_2Q_{31}K_2$ *fibrils. Scale bar is 50 nm. (c) Oxidized apolipoprotein A-I fibrils. Scale bars are 100 nm. (d) Aggregates of V75D hγD. Scale bar is 100 nm. Panel (a) reprinted with permission from Mandal, A.,* ***Boatz, J. C.****, Wheeler, T., van der Wel, P. C. A. On the use of ultracentrifugal devices for routine sample preparation in biomolecular magic-angle-spinning NMR, J. Biomol. NMR, 67(3):165-178. Copyright (2017) Springer Science + Business Media Dordrecht. Panel (c) reprinted with permission from Witkowski, A., Chan, G. K. L,* ***Boatz, J. C.****, Li, N. J., Wong, J. C., van der Wel, P. C. A., Cavigiolio, G. Methionine oxidized apolipoprotein A-I at the crossroads of HDL biogenesis and amyloid formation, The FASEB Journal 32(6):3149-3165. Copyright (2018) Witkowski, A., et. al.*

I used negative stain TEM to describe a range of samples in the Van der Wel lab (Figure 1.11), mostly to study the macromolecular morphology of both amorphous-looking and fibrillar aggregates [1, 3-6]. I also used negative stain TEM to distinguish amyloid polymorphs from one another and to visualize the aggregation of amyloid fibrils over time. For example, I prepared unlabeled Q44-httEx1 samples and analyzed macromolecular features of temperature-dependent polymorphs by TEM (Figure 1.12, a-d) [1].

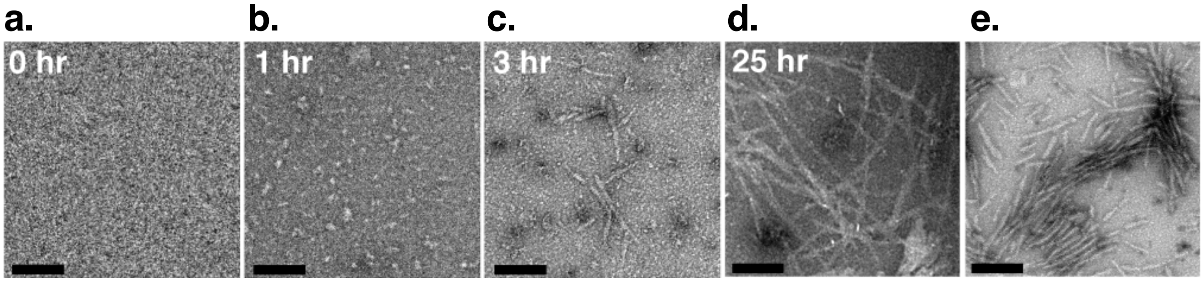


Figure 1.12. Huntingtin exon1 fibril formation monitored by TEM.

(a–e) Negatively stained TEM as a function of time after Factor Xa release of unlabelled exon1. (a) Uncleaved htt exon1 MBP fusion protein. (b) Oligomers observed 1 h after cleavage. (c) By 3 h, fibrils have begun to form. (e) After 25 h, fibrils have grown and oligomers are no longer visible on the grid. (e) Mature [U-^{13}C,^{15}N]-labelled fibrils prepared for ssNMR. (Scale bars: 200 nm.) Modified and reprinted with permission from Hoop, C. L., Lin H. K., Kar K., Magyarfalvi, G., Lamley, J. M., ***Boatz, J. C.****, Mandal, A., Lewandowski, J. R., Wetzel, R., van der Wel, P. C. A. Huntington exon 1 fibrils features an interdigitated β-hairpin-based polyglutamine core, PNAS 113(6):1546-1551. Copyright (2016) Hoop, C. L., et. al.*

I also used negative stain TEM to distinguish the difference in fibril widths between different to fibril polymorphs formed at 22 and 37 °C, respectively. I primarily analyzed TEM micrographs with ImageJ software, which contains several tools to aid in fibril width analysis [187]. For example, a histogram of fibril widths can be obtained by carefully measuring several fibrils with the straight-line Free-Hand tool. Analysis of TEM micrographs revealed that narrow fibrils can be formed at 37 °C and have an average width of 6-7 nm (Figure 1.13, a,b,e) [3]. At 22 °C, thicker fibrils form, which have an average width of 15-16 nm (Figure 1.13, c,d,f) [3]. These results are reproducible in independently prepared samples, and were the focus of one of my co-authored publications [3].

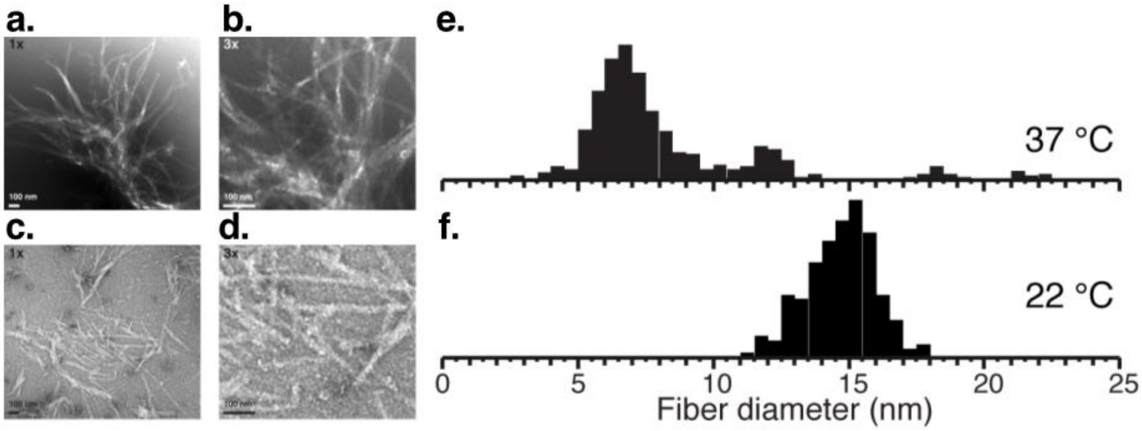


Figure 1.13. TEM data comparing polyQ-expanded htt exon1 fibrils formed at 22 and 37 °C

(a,b) Negative stain TEM of htt exon1 fibrils formed at 37°C, at two levels of magnification. (c-d) Negative stain TEM of htt exon1 fibrils formed at 22°C, at two levels of magnification. All scale bars are 100 nm. (e,f) Fibril width derived from negative-stain TEM on the mature fibrils formed at 37 °C (597 measurements over 99 fibrils) and 22 °C (219 measurements over 73 fibrils). Modified and reprinted from Lin, H. K., ***Boatz, J. C.****, Krabbendam, I. E., Kodali, R., Hou, Z., Wetzel, R., Dolga, A. M., Poirier, M. A., van der Wel, P. C. A. Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core, Nature Communications 8:15462. Copyright (2017) Lin, H. K., Boatz, J. C., et. al. under Creative Commons Attribution 4.0 International License.*

Alternatively, individual fibrils can be manually ‘picked’ and vertically aligned (Figure 1.14, a). In these cases, Fourier transform filters may be used to increase visibility of the fibril edges. I manually picked, aligned, and applied Fourier transform filters to several individual fibrils. High pass filters (Figure 1.14, b) keep the high frequencies from the Fourier transform of the macrographs while removing the low frequencies; the reverse is true for low pass filters. High pass filters are useful for observing the fine details of TEM micrographs. However, signal is decreased while the noise remains the same; therefore, the contrast between the grid and the sample worsens. Low pass filters are useful for increasing the contrast of micrographs, however in doing so the fine details are lost (Figure 1.14, c). A band pass filter can be applied to micrographs in order to increase image contrast while retaining most of the fine details (Figure 1.14, d,e). Band pass filters are also useful to inspect dense bundles of fibrils, which can appear as large black masses in micrographs with decreased visibility of fibrils frayed far from the bundle center (Figure 1.15).

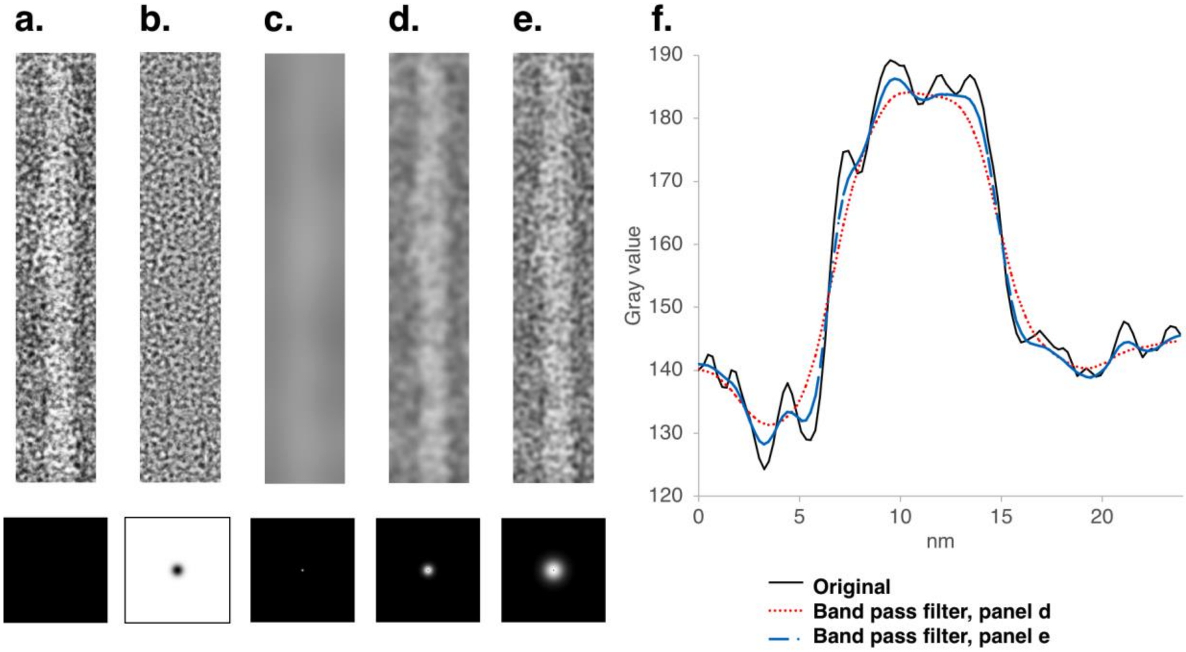


Figure 1.14. Fourier transform filtering of negatively stained httEx1 fibrils

(a) Manually 'picked' and vertically aligned fibril without applying Fourier transform filters, (b) after applying a high-pass filter, (c) low-pass filter, and (d,e) band pass filters. The Fourier transform filters applied are shown below the TEM image, depicted in reciprocal space. The white areas indicate frequencies which have been removed. (f) Profile plots from panels a, d, and e, overlaid. Filtering was done in ImageJ [187].

After picking individual fibrils, I obtained profiles of gray values along the fiber axis (Figure 1.14, g). High gray values are associated with areas of negative stain. The width of individual fibrils can then be obtained by measuring the half-height of a plot of gray values (Figure 1.14, g). Band pass filters can also be used to improve profile plots of gray values in order to better estimate the half-height when measuring fibril width. Profiles are also valuable for distinguishing individual filaments within fibrils. The stain penetrates the area between individual filaments in a manner that is not always easily detected by eye. By plotting the average gray value over the fiber axis, filaments interacting side-by-side can be distinguished with increased certainty.

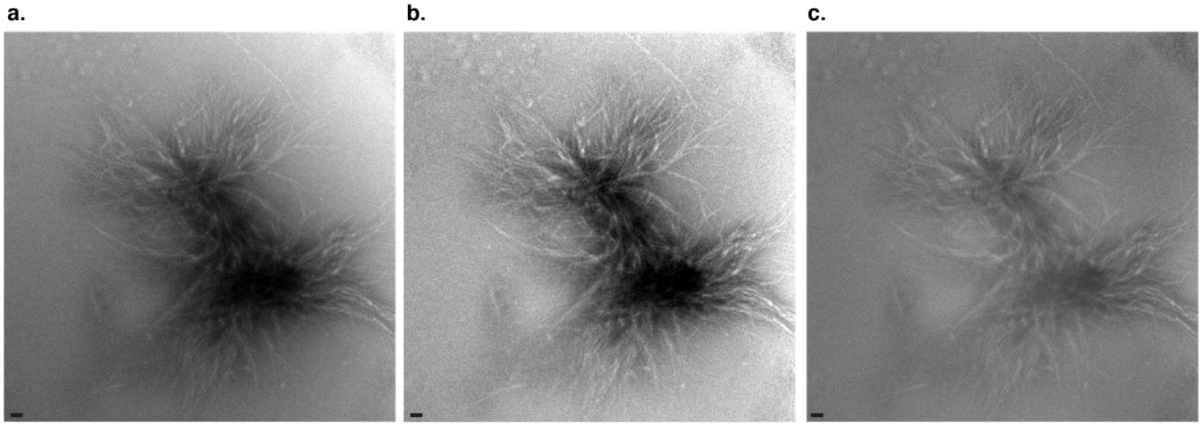


Figure 1.15. Bundling of Q44-httEx1 fibrils.

(a) Bundle of Q44-httEx1 fibrils. (b) Image was subjected to a band-pass filter to increase contrast and highlight the fibrils flayed from the bundle and the density of the bundled fibrils. (c) Image was subjected to a band-pass filter to deemphasize the density of the bundled fibrils and better visualize individual fibrils. All scale bars are 100 nm. Filtering was done in ImageJ [187].

1.5.3 Fourier transform infrared spectroscopy (FTIR)

Fourier transform infrared spectroscopy (FTIR) is a useful technique for determining the secondary structure elements present within protein structures, and it is advantageous as a complementary technique to MAS ssNMR. The amide I and amide II bands in an FTIR spectrum respectively report on the vibrational stretching of the C=O bond and the bending of the N-H bond [188]. These bonds are involved in hydrogen bonding that stabilizes α-helices, β-sheets, and other secondary structure elements, therefore the position of the amide I and II bands are sensitive to secondary structure [188]. FTIR requires much less sample than MAS ssNMR, does not require isotopic labeling, and can be used to rapidly analyze both soluble and the aggregated proteins.

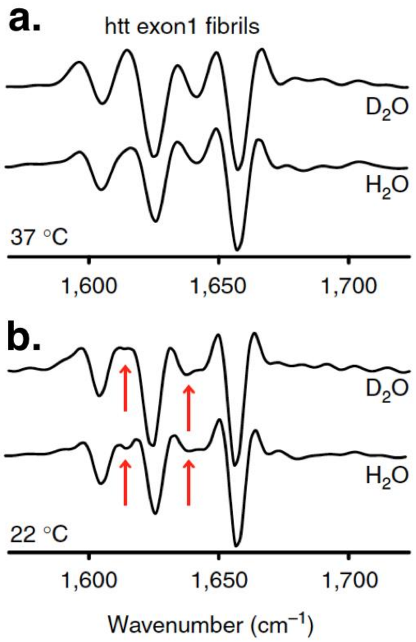


Figure 1.16. FTIR on httEx1 fibril polymorphs

(a) Second-derivative FTIR of htt exon1 fibrils formed at 37 °C and (b) 22 °C, for fibrils dispersed in either H_2O or D_2O. The colored arrows mark the most notable differences between the fibril types. (c) Reference data on fibrillar $K_2Q_{31}K_2$, HNTF ($htt^{NT}Q_{30}P_{10}K_2$) fibrils, and aggregated α-helical htt^{NT} in PBS buffer. (d) Resonance frequencies of different secondary structure elements. Reprinted from Lin, H. K., ***Boatz, J. C.****, Krabbendam, I. E., Kodali, R., Hou, Z., Wetzel, R., Dolga, A. M., Poirier, M. A., van der Wel, P. C. A. Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core, Nature Communications 8:15462. Copyright (2017) Lin, H. K., Boatz, J. C., et. al. under Creative Commons Attribution 4.0 International License. Panel (f) was adapted with permission from Sivanandam, V. N., et al, Copyright (2011) American Chemical Society.*

Dr. Ravindra Kodali obtained FTIR results on samples that I prepared, which were used to observe subtle structural differences between temperature-dependent Q44-httEx1 polymorphs [3]. FTIR measurements show that both fibril types display the same dominant secondary structure features present in model polyQ peptide aggregates (Figure 1.16), however there are minor but distinct differences in signals associated with the PPII helices in the PRD (red arrows, Figure 1.16), which can be understood by the magic angle spinning solid state NMR results discussed in section 2.2 [3]. The FTIR results are reproducible in independently prepared samples [3].

1.5.4 X-ray diffraction

X-ray diffraction of amyloid fibrils generate specific patterns that indicate a high order of repetitive β-sheet structure. Specifically, the β-sheets are arranged parallel to the axis of the fibrils [189]. Two reflections are observed in amyloid, which correspond to 4.7-4.9 and 8-12 Å repetitive spacings [189]. The former reflection reveals the distance between β-strands that are interacting via hydrogen bonds, while the latter reveals the distance between β-sheets within the core. Un-oriented fibrils produce diffraction patterns referred to as a “powder” pattern; because the fibrils within the sample are not aligned with respect to one another, the reflections associated with intra- and inter-β-sheets experience a complete radial averaging [189]. Aligned fibril samples diffract X-rays into a cross-β fiber diffraction pattern where half-moon shaped reflections appear perpendicular to each other along the meridian and equator of the pattern [189]. Of note, X-ray diffraction of fibrils does not permit the reconstruction of the atomic structure, however it can be used to provide valuable structural constraints than can supplement higher resolution structural techniques.

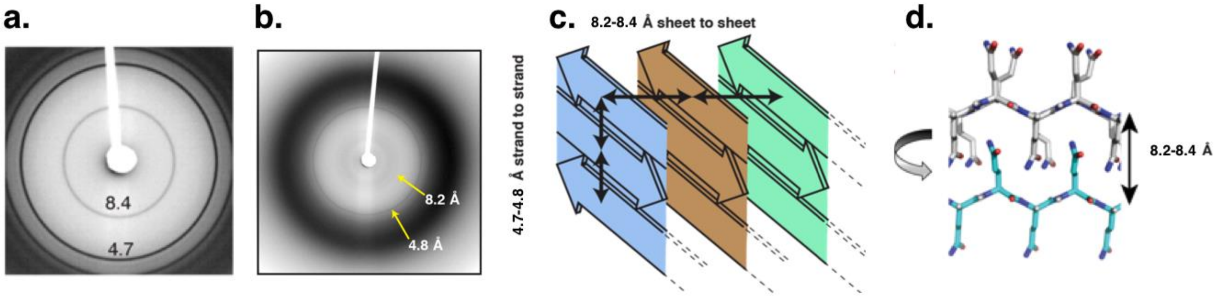


Figure 1.17. PolyQ amyloid structure

(a) X-ray powder diffraction on hydrated $K_2Q_{31}K_2$ *fibrils shows the cross-β dimensions of polyQ amyloid. (b) X-ray powder diffraction on hydrated Q44-httEx1 fibrils. (c) Cartoon depicting spacings between β-strands and β-sheets in polyQ peptides and httEx1. (d) The glutamine side chains interdigitate in polyQ peptides and httEx1. Panels a and d modified and reprinted with permission from Hoop, C. L., Lin H. K., Kar K., Magyarfalvi, G., Lamley, J. M.,* ***Boatz, J. C.****, Mandal, A., Lewandowski, J. R., Wetzel, R., van der Wel, P. C. A. Huntington exon 1 fibrils features an interdigitated β-hairpin-based polyglutamine core, PNAS 113(6):1546-1551. Copyright (2016) Hoop, C. L., et. al.*

The X-ray powder diffraction data I collected from $K_2Q_{31}K_2$ peptide fibrils reveal that the backbone-to-backbone distances between β-strands and β-sheets are 4.7 and 8.4 Å respectively (Figure 1.17, a). The diffraction pattern of mature Q44-httEx1 aggregates is less well resolved, however the observed distances of ~ 4.8 Å and 8.2 Å (Figure 1.17, b) are not significantly different from the distances observed for the $K_2Q_{31}K_2$ peptide fibrils and are consistent with prior findings [190]. These results point to the idea that the amyloid core of polyQ peptides and polyQ containing proteins is very similar (Figure 1.17, c). Of particular note is the short 8.2 - 8.4 inter-sheet distance, which indicates that the β-sheet backbones are very close together in space and the glutamine side chains interdigitate in a steric zipper fashion (Figure 1.17, d).

1.5.5 Protein production and purification

Structural and biophysical studies require the ability to reconstitute the aggregation process *in vitro*. This requires the production of protein. HttEx1 is expressed as a fusion construct to maltose-binding protein (MBP), designated MBP-httEx1, with a Factor Xa specific proteolytic cleavage site located on a short linker between the MBP and the httEx1 (Figure 1.18, a) [3]. HttEx1 has several binding partners and can be modified post-translationally, as indicated in Figure 1.18, b. Cleavage of MBP-Q44-httEx1 by Factor Xa (FXA) results in the release of monomeric Q44-httEx1, which rapidly forms oligomeric species and later fibrils that elongate over time. A C-terminal his-tag allows for purification of the MBP-httEx1 fusion protein by affinity over a nickel-column, and size-exclusion with a gel column can subsequently remove lysozyme if it is introduced during cell lysis. The yield in minimal media is sufficient to produce milligrams of uniformly isotopically labeled protein (^{13}C, ^{15}N) that is necessary for MAS ssNMR structural studies. A construct with a disease-relevant 44-residue-long polyQ domain was used for several experiments.

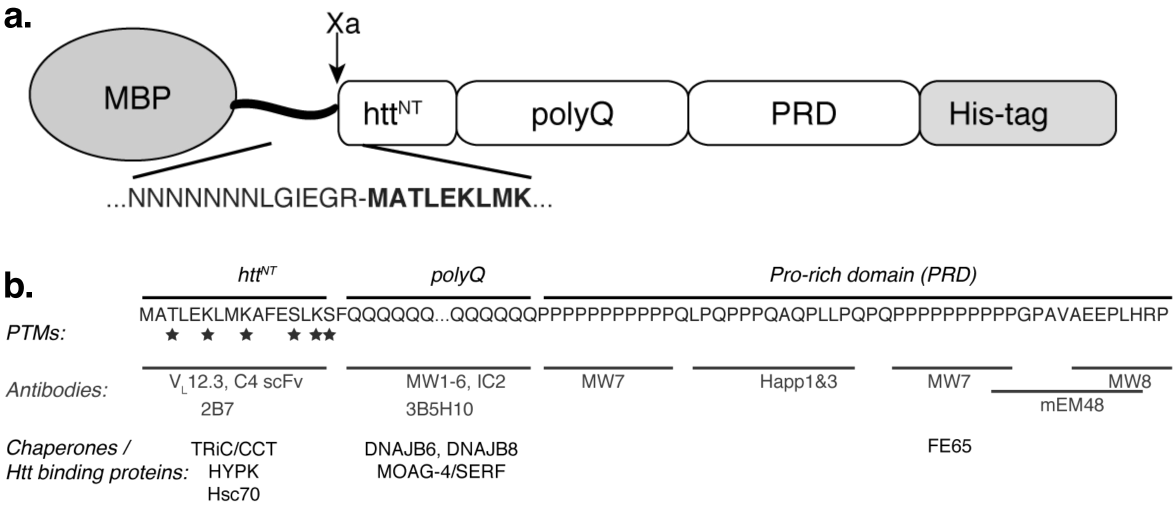


Figure 1.18. Huntingtin exon 1 sequence and domain structure

(a) Design of the MBP fusion protein, with the sequence of the Factor Xa cleavage site in the linker shown below. (b) The domain structure and sequence of htt exon1 is shown at the top. The locations of post-translational modifications (PTMs), binding sites of various antibodies, and other htt-binding proteins are also indicated [93, 191-200]. Modified and reprinted from Lin, H. K., ***Boatz, J. C.****, Krabbendam, I. E., Kodali, R., Hou, Z., Wetzel, R., Dolga, A. M., Poirier, M. A., van der Wel, P. C. A. Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core, Nature Communications 8:15462. Copyright (2017) Lin, H. K., Boatz, J. C., et. al. under Creative Commons Attribution 4.0 International License.*

1.5.6 Kinetics of aggregation and amyloid formation

1.5.6.1 Sodium dodecyl sulfate polyacrylamide gel electrophoresis

I used SDS-PAGE during sample preparation and in order to estimate the rate of httEx1 release from the MBP-httEx1 fusion protein (Figure 1.19). The MW of MBP-Q44-httEx1 is 57 kDa at isotopic natural abundance (NA). The release of MBP (43 kDa) and the disappearance of the fusion protein can both be monitored over time after proteolysis by FXA. By combining SDS-PAGE cleavage assays and TEM (see section 1.5.2), I found that the morphology of mature Q44-

httEx1 fibrils depends partially on the rate of release of the Q44-httEx1 monomer (see Chapter 3.0).

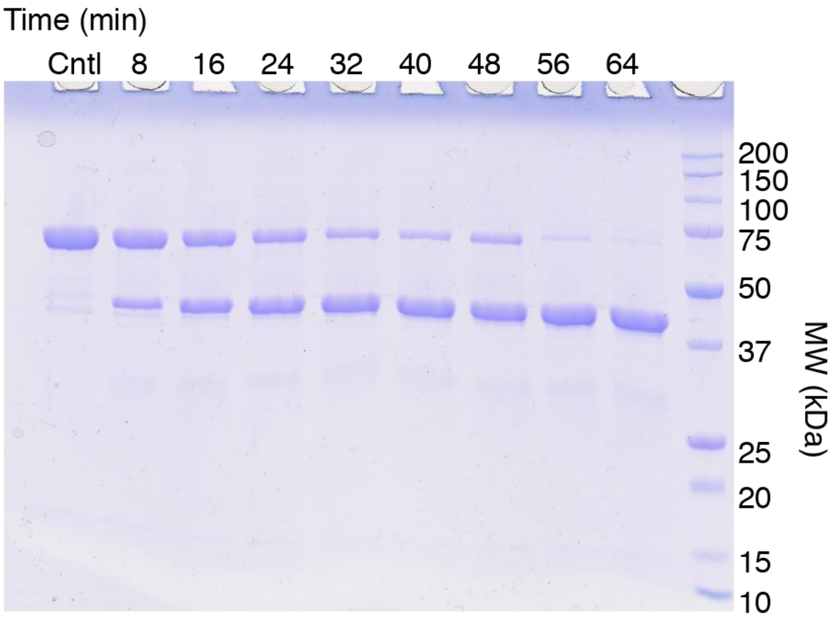


Figure 1.19. Kinetics of httEx1 release from fusion construct monitored by SDS-PAGE

The MBP fusion protein (Cntl lane) is exposed to FXA, and release of the MBP (new band in additional lanes) and disappearance of the fusion protein is monitored over time. The release of MBP correlates 1:1 to the release of httEx1.

1.5.6.2 Electrospray ionization time of flight mass spectrometry

Mass spectrometry (MS) is an analytical technique used to detect protein(s) within sample(s) and identify their MWs. Protein samples are ionized into a vapor, and the ions are then separated through an electromagnetic field based on their mass-to-charge ratio (m/z), which is plotted vs intensity as a mass spectrum. Depending on the ionization method used, MS can detect as little as attomole concentrations of protein with MWs of up to 300,000 Da [201]. When combined with other analytical techniques, MS can be a powerful tool for protein quantification and is therefore ideal for assessing protein purity.

The specific form of MS employed in these studies was electrospray ionization time of flight mass spectrometry (ESI-TOF MS). There are several advantages of ESI-TOF over other MS techniques. The most notable advantage is its impressive sensitivity; ESI-TOF can accurately determine the MW of intact protein within ± a single Da [201]. This is essential to quantify the percentage of ^{13}C and ^{15}N atoms (vs ^{12}C and ^{14}N) within an isotopically labeled sample.

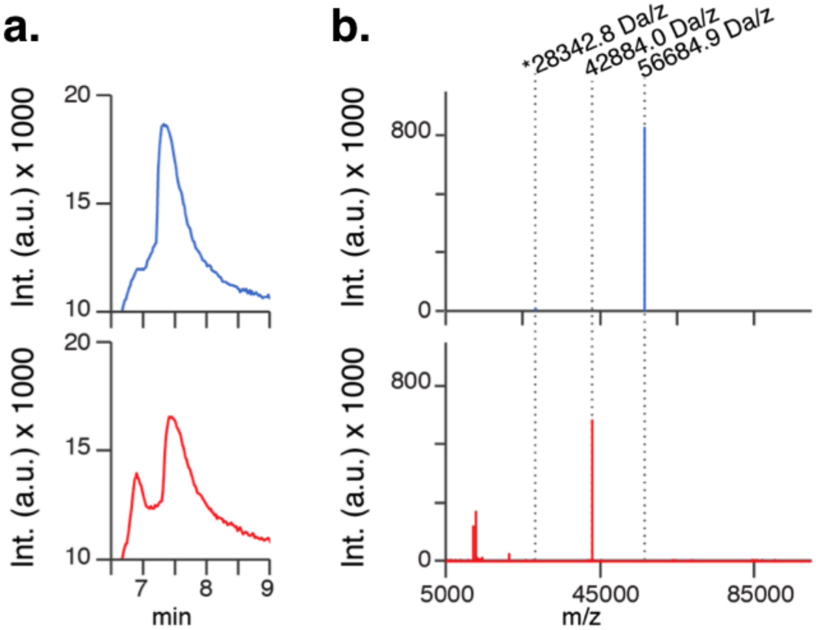


Figure 1.20. MBP release from MBP-Q44-httEx1 fusion protein monitored by ESI-TOF MS

(a) UV absorbance following of LC-MS elutions for MBP-Q44-httEx1 before (top) and after (bottom) proteolytic cleavage. (b) ESI-TOF mass spectra corresponding to the entirety of the peaks depicted in panel a.

ESI-TOF MS was also integral for several proteolytic cleavage assays described in Chapter 3.0. I used ESI-TOF MS coupled with high performance liquid chromatography (HPLC) to quantify the rate of httEx1 release from the fusion protein and identify additional cleavage products resulting from proteolysis. An HPLC column is used to separate protein within a sample, and protein elutions are monitored by UV absorbance. This is shown in Figure 1.20, a, for MBP-Q44-httEx1 both before (top) and after (bottom) cleavage. Mass spectra measured over the entirety of the elution peaks are shown in Figure 1.20, a. Prior to cleavage, MBP-Q44-httEx1 is observed as a single peak at 56684.9 Da/z (Figure 1.20, b, top), and in some samples also at 28242.8 Da/z. The latter represents the MW of MBP-Q44-httEx1 ions with a charge of $z = +2$. MBP is observed after cleavage at 42884.0 Da/z (Figure 1.20, bottom) and occasionally at 21442 Da/z when $z = +2$. Additional cleavage products may also be observed by ESI-TOF MS, including httEx1.

1.5.6.3 High performance liquid chromatography (HPLC) based sedimentation assays

I used HPLC to quantify the kinetics of peptide and protein aggregation [3, 54, 202]. I thank Dr. Ravi Kodali for his mentorship and help with my initial sedimentation assays. Briefly, aggregation of amyloid-like fibrils was initiated, often by proteolytic cleavage, in the presence and absence of pre-formed amyloid seeds. At various time points, aggregates were removed from aggregating reaction mixtures by sedimentation via (ultra)centrifugation and aliquots of the supernatant were then diluted into formic acid and loaded onto an HPLC column. The column is necessary to separate the monomers of amyloidogenic peptides and protein from proteases, soluble cleavage products, and other contaminants. The elution of amyloidogenic monomers is monitored by UV absorbance at 215 nm (A_{215}), which reports on the relative concentration of peptide backbone in the sample in the absence of tryptophan residues in the protein sequence. An ESI-TOF MS instrument was used to automatically determine the molecular weight of the molecules present in each elution. Manual integration of the A_{215} peaks compared to negative and positive controls allows for the quantification of monomeric species remaining in the aggregating sample over time.

1.5.6.4 Thioflavin T (ThT) fluorescence

Thioflavin T (ThT) is a small dye that usually fluoresces in the presence of amyloids but not in the presence of non-amyloid (e.g. amorphous) aggregates. Upon binding to amyloids, the dynamics of the ThT dye changes, and this change induces fluorescence [203, 204]. The intensity of the induced ThT fluorescence correlates linearly with the concentration of amyloid fibrils, therefore ThT assays are highly quantitative [205]. I used this dye to measure the kinetics of amyloid formation for both httEx1 and P23T hγD [3, 5]. I found that kinetics of fibril formation is strongly dependent on the length of the expanded polyQ domain and the concentrations of protease and substrate, with only minimal dependence on temperature.

As discussed above, amyloid formation follows sigmoidal kinetics (Figure 1.2). As monomers begin to aggregate and form amyloid fibrils, ThT fluorescence typically increases in a sigmoidal fashion. This indicates that the ThT dye preferentially binds to a structural motif present in both the nuclei and fibrils and does not favor binding to species prior to nuclei formation. In Figure 1.21, ThT fluorescence is shown for two timepoints following httEx1 release via Factor Xa (FXA) proteolytic cleavage of the MBP fusion protein. Importantly, release of httEx1 reached completion within 15 minutes for this particular assay (as verified by SDS-PAGE). After 1 hour from the addition of FXA protease, httEx1 (mostly monomeric, some oligomers potentially present) does not exhibit significant ThT fluorescence compared to control samples without FXA proteolysis. After 3 days, mature fibrils bind readily to the ThT dye and induce fluorescence.

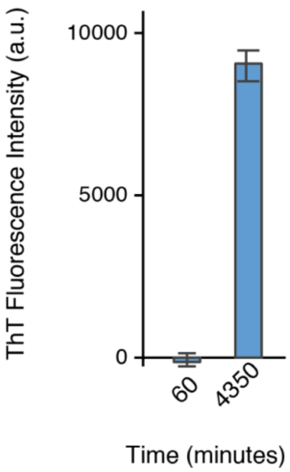


Figure 1.21. ThT fluorescence of httEx1 monomer and mature fibrils

After 60 minutes from FXA proteolysis, no significant increase in fluorescence is observed compared to negative control. After 4350 minutes (3 days), mature fibrils have formed, and significant ThT fluorescence is observed.

1.5.6.5 Amyloid seeding assays

I performed several assays to test the ability of amyloid fibrils to propagate through seeding. During my work on httEx1 aggregation (as discussed in Chapter 1.5 and Chapter 3.0), we observed different polymorphs for Q44-httEx1 aggregated at 22 and 37 °C [3]. Both fibril types are capable of seeding expanded polyQ protein (Figure 1.22) and shorter model polyQ peptides (Figure 1.23) [3]. This is shown with a ThT fluorescence assay (Figure 1.22, a,b; Figure 1.23, a,c) that monitors the production of amyloid over time at 22 °C [3]. When no seeds are present, amyloid formation begins after a long lag phase. In the presence of 20mol-% of pre-aggregated seeds, the lag phase is abolished. The rate of amyloid formation in the elongation phase is much faster in the presence of seeds. As expected, amyloid formation was found to be correlated with aggregate formation detected through HPLC-based sedimentation assays (Figure 1.22, c,d; Figure 1.23, b,d), performed in parallel to the ThT assays [3], as discussed above in section 1.5.6.3.

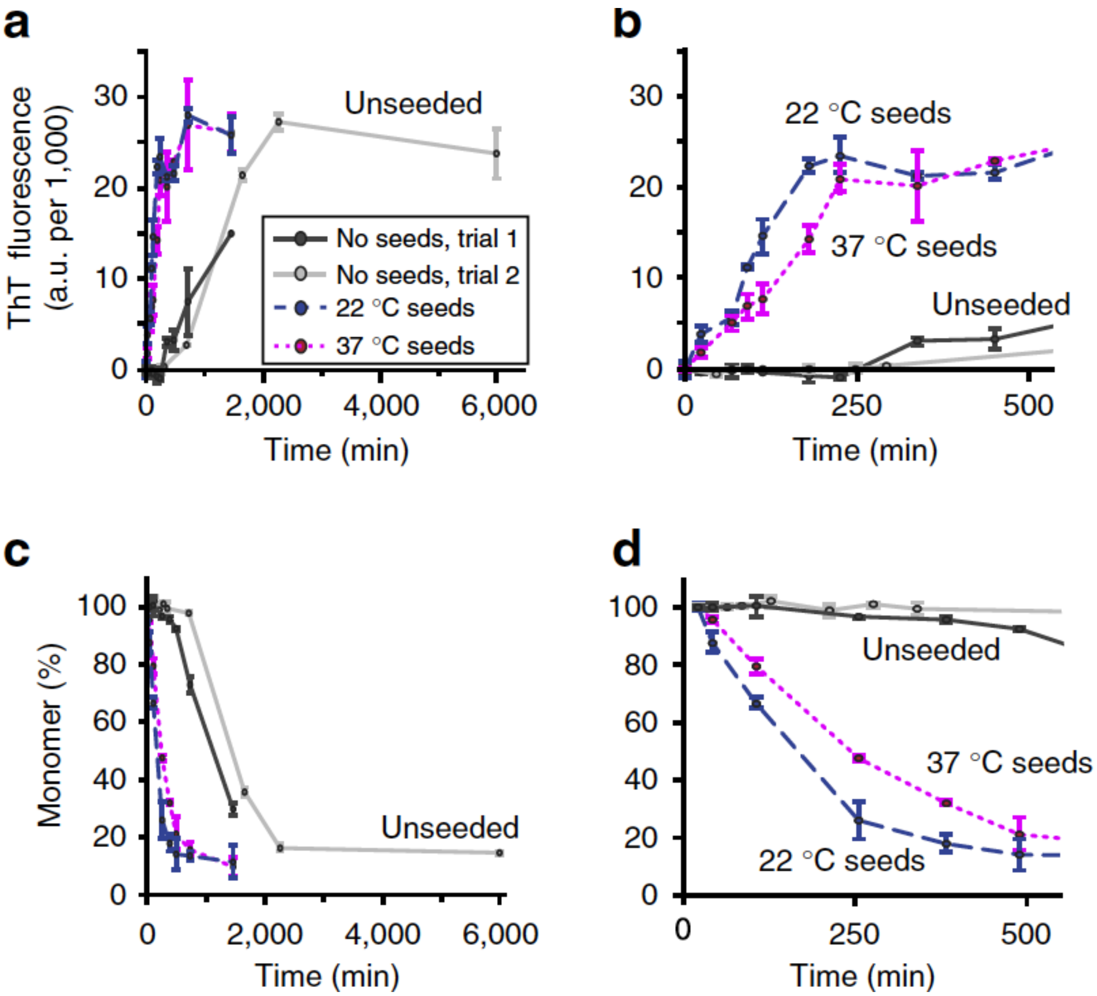


Figure 1.22. PolyQ protein recruitment through seeding

(a) Aggregation kinetics at 22 °C in the absence (solid black and grey lines) and presence (dashed lines) of pre-made seed aggregates, detected as ThT fluorescence at indicated time points after complete trypsin cleavage of the htt exon1 fusion protein. Dark blue and magenta dashed lines reflect the aggregation in presence of 20 mol-% htt exon1 aggregates formed at 22 and 37 °C, respectively. The unseeded reactions have lag phases exceeding 4 h, which are eliminated by the seeds. Error bars indicate s.d., with n=2–3. (b) Expansion of the first 500 min. (c,d) Results of a single (n=1) HPLC measurement of the residual monomer concentration after aggregate sedimentation, applied to the same samples, as a complementary measure of aggregation. Error bars reflect the estimated peak integration error. Modified and reprinted from Lin, H. K., ***Boatz, J. C.****, Krabbendam, I. E., Kodali, R., Hou, Z., Wetzel, R., Dolga, A. M., Poirier, M. A., van der Wel, P. C. A. Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core, Nature Communications 8:15462. Copyright (2017) Lin, H. K., Boatz, J. C., et. al. under Creative Commons Attribution 4.0 International License.*

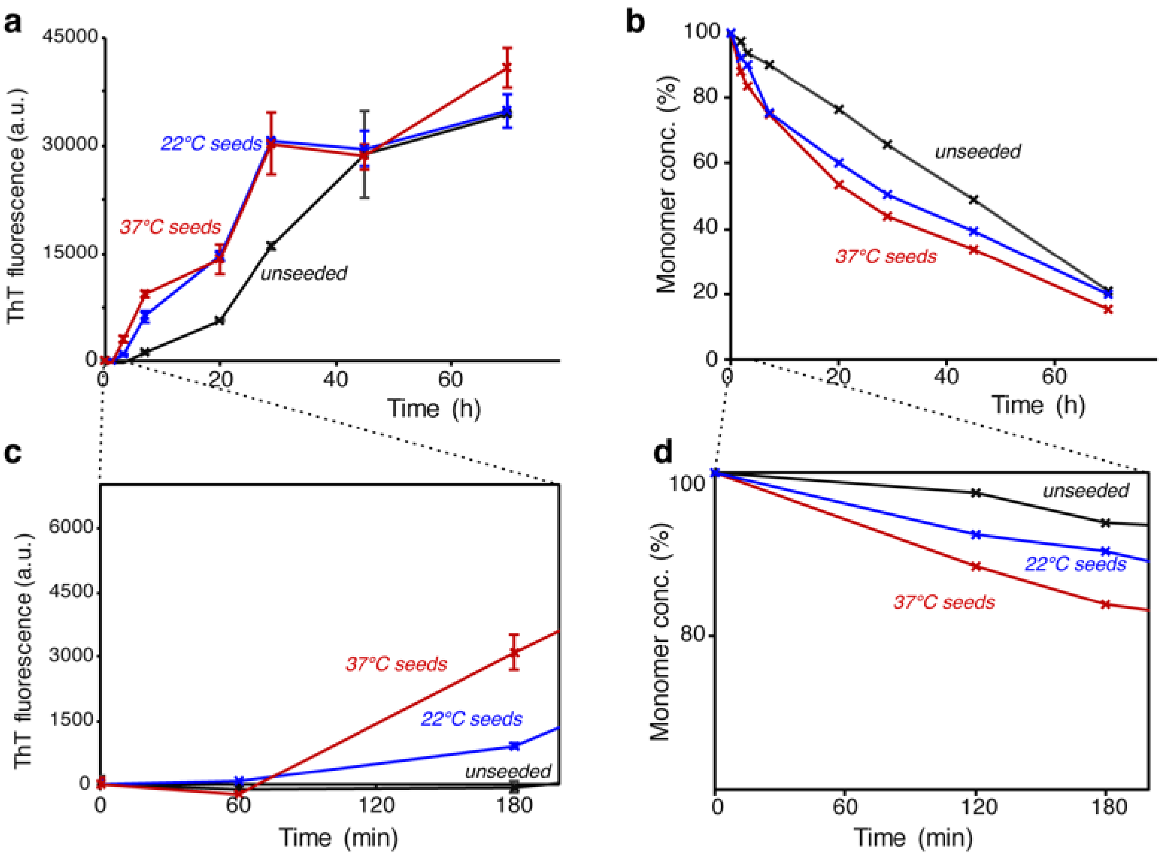


Figure 1.23. Seeding assay results on htt^{NT} $Q_{23}P_{10}K_2$ peptide

*(a) Seeding effects on the aggregation of htt^{NT} $Q_{23}P_{10}K_2$ peptide [54] at 37 °C, monitored by ThT fluorescence measurements. Reactions were performed without seeds (black; solid), and with 20-mol-% htt exon1 seeds prepared either at 22 °C (dark blue; solid), or 37 °C (magenta; dashed). Error bars indicate s.d. with n=2 or n=3. (b) In the same samples, sedimentation followed by reverse-phase HPLC was used to monitor the percent monomer remaining in the supernatant. (c,d) Expansion of the first 200 minutes of panels (a) and (b), respectively. Both types of polymorphs are capable of cross-seeding the aggregation of this peptide featuring a shorter polyQ length that is below the HD threshold. In the earliest stages (see panels c,d) the fibrils formed at 37 °C appear to be more effective at seeding this aggregation reaction. Reprinted from Lin, H. K., **Boatz, J. C.**, Krabbendam, I. E., Kodali, R., Hou, Z., Wetzel, R., Dolga, A. M., Poirier, M. A., van der Wel, P. C. A. Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core, Nature Communications 8:15462. Copyright (2017) Lin, H. K., Boatz, J. C., et. al. under Creative Commons Attribution 4.0 International License.*

1.5.7 Toxicity studies

To test whether our polymorphic protein fibrils were cytotoxic, I prepared several samples for toxicity studies, which were performed and analyzed by my collaborators Inge Krabbendam and Dr. Amalia M. Dolga at the University of Groningen [3]. Immortalized murine hippocampal and human differentiated dopaminergic neuronal cell lines were exposed to each Q44-httEx1 temperature polymorph to probe potential cytotoxic effects. The results are plotted in Figure 1.24, e,f [3]. Cell death was not observed at 0.5 μM concentrations in either cell type. At 1 μM, cell death was induced in the dopaminergic cell line but not in the immortalized cells, however at 5 μM cell death was induced in both cell lines. There appears to be a slightly larger cytotoxic affect in the presence of 37 °C fibrils, as indicated by differences in the morphology of treated dopaminergic neurons (Figure 1.24, a-d) [3].

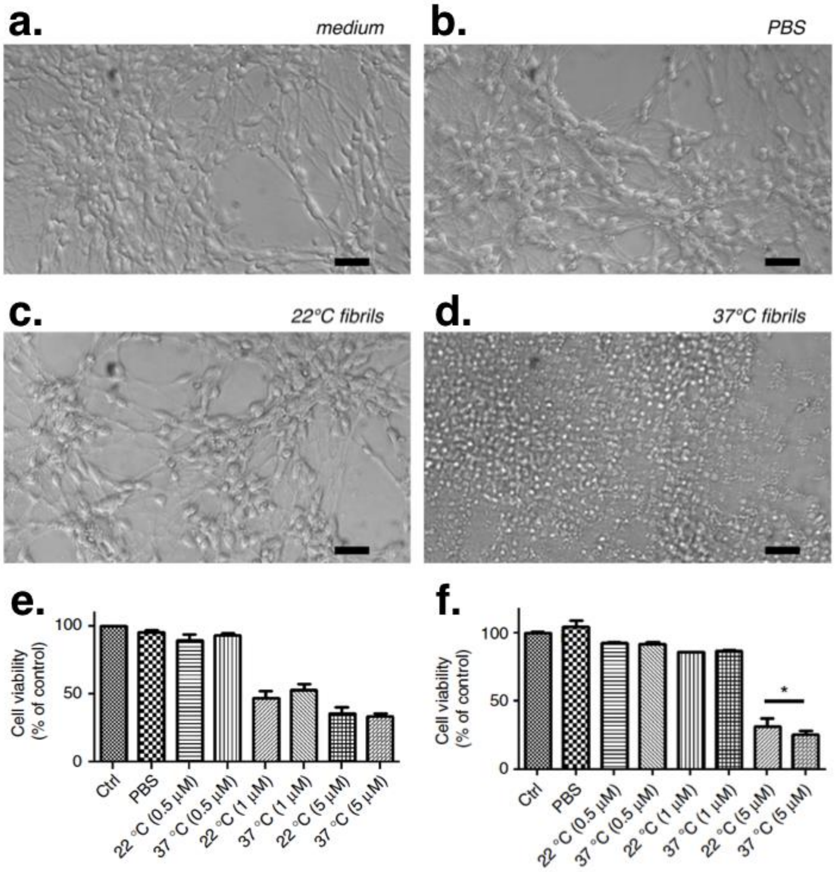


Figure 1.24. Treated and untreated human dopaminergic neurons

*(a) Control neurons compared to treated neurons 24 h after administering (b) PBS buffer only, (c) 5 μM of the thick 22 °C fibrils in PBS, and (d) 5 μM of the thinner 37 °C fibrils in PBS. Scale bars are 50 μm. (e) Cellular viability of human dopaminergic neuronal cells upon exposure to varying concentrations of pre-formed fibrils prepared at 22 and 37 °C. The data reflect MTT reduction assays performed after 24 h (n=2; two biological replicates with three technical replicates each—shown is the mean with s.d. compared to non-treated controls set at 100%). (f) Cell viability assay data for a 24 h exposure of immortalized HT-22 neurons (n=2; two biological replicates with 6 technical replicates each–shown is the mean with s.d. compared to non-treated controls set at 100%; *P<0.05, Mann–Whitney non-parametric test). Modified and reprinted from Lin, H. K.,* ***Boatz, J. C.****, Krabbendam, I. E., Kodali, R., Hou, Z., Wetzel, R., Dolga, A. M., Poirier, M. A., van der Wel, P. C. A. Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core, Nature Communications 8:15462. Copyright (2017) Lin, H. K., Boatz, J. C., et. al. under Creative Commons Attribution 4.0 International License.*

2.0 STRUCTURAL INSIGHTS ON EXPANDED HUNTINGTIN EXON 1

Here, I summarize conclusions from recent results that the Van der Wel lab has published, which report high-resolution data on the structure and dynamics of the individual domains of Q44-httEx1 polymorphs. It is vital to understand the results and conclusions presented in these manuscripts prior to reading Chapter 3.0. Together, these results demonstrate that the polyQ amyloid core arranges into an interdigitated antiparallel β-sheet structure that is robust between model polyQ peptides and Q44-httEx1 polymorphs. The flanking domains were found to have the same structure in two temperature-dependent Q44-httEx1 polymorphs, however interactions between flanking domains and the exposure of the flanking domains to the solvent were found to be important factors that dictate the polymorphism.

2.1 SUMMARY: HUNTINGTIN EXON 1 FIBRILS FEATURE AN INTERDIGITATED β-HAIRPIN-BASED POLYGLUTAMINE CORE

Hoop, C. L., Lin H. K., Kar K., Magyarfalvi, G., Lamley, J. M., **Boatz, J. C.**, Mandal, A., Lewandowski, J. R., Wetzel, R., van der Wel, P. C. A. *Huntington exon 1 fibrils features an interdigitated β-hairpin-based polyglutamine core,* PNAS 113(6):1546-1551. Copyright (2016) Hoop, C. L., et. al.

> I contributed data towards two figures in this manuscript, reprinted with permission as Figure 1.17, a, and Figure 1.12. Specifically, I obtained X-ray powder diffraction patterns (section 1.5.4) on $K_2Q_{31}K_2$ peptide fibrils, and I monitored the time course of Q44-httEx1 fibril formation by TEM (section 1.5.2). I also obtained negative stain TEM images on U-^{13}C,^{15}N labeled Q44-httEx1 fibrils.

The polyQ amyloid core of Q44-httEx1 fibrils forms antiparallel β-sheets and the extended and rigid glutamine side chains form two distinct structure types ('a' and 'b') in equal amounts. Each β-strand is found to present only one glutamine side chain type. The polyQ domain of httEx1 displays the same chemical shift signature representative of amyloid structure in several model polyQ peptides; peak patterns for 'a'/'b' type glutamine remain the same in the presence and absence of model flanking domains, with only one or two sequential residues labeled, and regardless of the length of the polyQ domain. Together, these data are indicative of a conserved polyQ structure between multiple polyQ model peptides and Q44-httEx1 polymorphs.

Glutamine side chain Nε and backbone Cα atoms from different molecules were found to be close in space, revealing an interdigitated glutamine steric-zipper between strands. These results, together with solvent exposure studies, are consistent with an intramolecular β-hairpin structure made up of both 'a' and 'b' strands in the polyQ domain of Q44-httEx1 fibrils. A third type 'c' glutamine likely populates the β-turns within the hairpin, which are typically four or more residues in length. Because signals from 'c' type glutamine make up at most 10% of the total glutamine signal in the Q44-httEx1 fibrils, it is predicted that there is a single β-turn per molecule. This implies that each β-strand in Q44-httEx1 is 20-residues in length. Kinetics studies on model polyQ peptides suggest that the minimum polyQ length required for β-hairpin formation is 26 residues, therefore the minimum β-strand length in polyQ β-hairpins is predicted to be 11 residues long [1, 146, 206].

2.1.1 Proposed stochastic formation of β-hairpins

Peak doubling is also observed whenever a single glutamine residue is labeled in model peptides; this is consistent with the idea of stochastic formation during nucleation or fibril elongation. For Q44-httEx1 fibrils, this implies that in half of the molecules the first 20 glutamine residues (Q18-Q37) adopt a type 'a' structure and the last 20 glutamine residues (Q42-Q61) adopt a type 'b' structure. The remaining half of the molecular population adopts a type 'b' structure in residues Q18-Q47 and a type 'a' structure in residues Q42-Q61. The observed peak doubling appears independent of polyQ length and is also observed in $D_2Q_{15}K_2$ peptides that are unlikely to form a β-hairpin but do assemble into fibrils in an antiparallel fashion.

2.2 SUMMARY: FIBRIL POLYMORPHISM AFFECTS IMMOBILIZED NON-AMYLOID FLANKING DOMAINS OF HUNTINGTIN EXON1 RATHER THAN ITS POLYGLUTAMINE CORE

Lin, H. K., **Boatz, J. C.**, Krabbendam, I. E., Kodali, R., Hou, Z., Wetzel, R., Dolga, A. M., Poirier, M. A., van der Wel, P. C. A. *Fibril polymorphism affects* *immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core*, Nature Communications 8:15462. Copyright (2017) Lin, H. K., et. al.

I directly contributed towards two main text figures and three supplementary figures in the published manuscripts, some of which are reprinted with permission as Figure 1.22, Figure 1.23, and Figure 1.13. I produced protein and prepared several fibril samples (section 1.5.5), monitored aggregation by HPLC sedimentation assays (section 1.5.6.3), monitored amyloid formation by ThT fluorescence (section 1.5.6.4), performed several seeding assays (section 1.5.6.5), and obtained micrographs of fibrils by negative stain transmission electron microscopy (section 1.5.2). I thank Dr. Ravindra Kodali for his help and guidance with the assays utilizing HPLC and ThT fluorescence. I also prepared samples that were used in toxicity studies (section 1.5.7) performed by my collaborators Inge Krabbendam and Dr. Amalia M. Dolga at the University of Groningen, and for FTIR measurements (section 1.5.3) performed by Dr. Ravindra Kodali.

Temperature during fibrillation affects the width of mature Q44-httEx1 fibrils as seen in my TEM data (see section 1.5.2); 15-16 and 6-7 nm fibrils are formed at 22 and 37 °C, respectively. Both polymorphs are capable of propagating through seeding and are cytotoxic to neuronal cell lines. There is a small but noticeable heightened cytotoxicity of the 37 °C fibrils on dopaminergic neurons. Subtle differences are observed by FTIR in signals that originate from the PRD. In both fibril types the polyQ signature is the same. The htt^{NT} and PRD feature α- and PPII helices that are immobilized, and the polyQ-proximal sections of each flanking domain displays restricted motion.

The temperature-dependent polymorphs present variable dynamics and solvent accessibility in the PRD that affects the binding ability of antibodies. The thinner fibrils formed at 37 °C display several highly flexible solvent-exposed residues in the PRD and readily bind to MW7 and MW8 antibodies. Conversely, the fibrils formed at 22 °C display restricted motion and less solvent-accessibility in the PRD, which results in decreased binding to MW7. This contrasts with prior studies which attributed polymorphism to differences in the structure of the polyQ domain [47].

2.2.1 Proposed model of htt exon1 fibrils

The results presented in this manuscript and prior studies from the Van der Wel lab were best explained by a structural model of Q44-httEx1 fibrils that is presented in Figure 2.1. Intermolecular β-hairpins makeup the polyQ domain, which places the flanking domains near each other in space. The htt^{NT} forms an α-helix and a short linker connects the polyQ domain to a PPII helix within the PRD. This arrangement additionally constrains the motion of the polyQ-proximal segments of the flanking domains and limits the accessibility of the flanking domains to the solvent. The polyQ amyloid core features β-strands consisting of 20 residues each and spanning 6-7 nm, consistent with the width of the 37 °C fibrils observed by TEM. The 22 °C fibrils are proposed to consist of two interacting filaments that primarily interact through the PRD, which results in restricted flanking domain motion and reduces solvent exposure. The filaments present in the 37 °C sample do not interact, leaving the flanking domains solvent accessible with less constrained motion. Together these results show that differences in the solvent exposure and interactions of the polyQ flanking domains are important for Q44-httEx1 fibrillar structure and function.

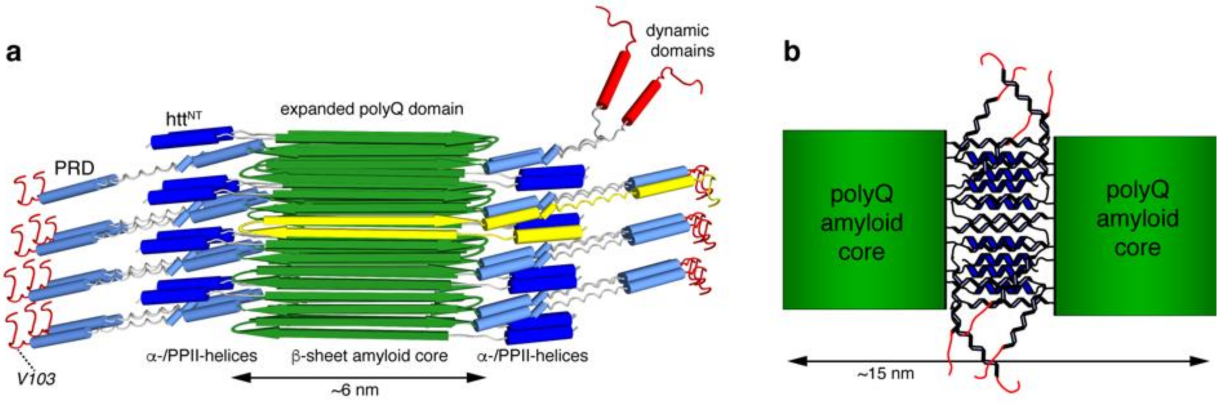


Figure 2.1. Schematic proposed model of htt exon1 fibrils

(a) The htt^{NT} α-helices (dark blue) and PRD PPII helices (light blue) are immobilized and tightly clustered on the perimeter of the rigid amyloid core (green β-strands). C-terminal domains show increased dynamics, either in the form of the unstructured C-terminal tail or a subpopulation of more exposed PRDs (top right; red). An individual protein monomer with its β-hairpin-based polyQ core is shown with lighter (yellow) β-strands. (b) Schematic illustration of interfilament flanking domain interactions that we propose to explain the larger TEM-based widths of the fibrils formed at 22 °C, as well as the observed differences in accessibility and immobilization of the PRD. Reprinted from Lin, H. K., ***Boatz, J. C.****, Krabbendam, I. E., Kodali, R., Hou, Z., Wetzel, R., Dolga, A. M., Poirier, M. A., van der Wel, P. C. A. Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core, Nature Communications 8:15462. Copyright (2017) Lin, H. K., Boatz, J. C., et. al. under Creative Commons Attribution 4.0 International License.*

3.0 ROLES OF CONCENTRATION AND THE N-TERMINAL FLANKING DOMAIN IN MUTANT HUNTINGTIN EXON 1 FIBER POLYMORPHISM

Jennifer C. Boatz, Talia Piretra, Irina Matlahov, James F. Conway, Patrick C. A. van der Wel. (2018) Roles of concentration and the N-terminal flanking domain in mutant huntingtin exon 1 fiber polymorphism, *manuscript in preparation.*

Data availability: Assigned chemical shifts of Q44-httEx1 fibrils were reported previously [3] and deposited into the Biological Magnetic Resonance Data Bank (BMRB) as entry 27045. All other data are available from the corresponding author.

> I performed all MAS ssNMR, X-ray diffraction, mass spectrometry, and fluorescence/seeding experiments and analyzed the associated data. I also collected all TEM micrographs, which Talia Piretra, Irina Matlahov, James F. Conway, and I analyzed. Talia Piretra and I performed and analyzed all monomer release (cleavage) assays. All samples were prepared by me, Talia Piretra, and Irina Matlahov. I designed all experiments with Patrick C. A. van der Wel, I created all figures, and I wrote the manuscript with input from all co-authors, especially from Patrick C. A. van der Wel and James F. Conway.

Huntington’s disease is a progressive neurodegenerative disease caused by expansion of the polyglutamine domain within the first exon of the huntingtin protein. The extent of polyglutamine expansion in huntingtin directly correlates with disease progression and the formation of amyloid-like fibrillar deposits of huntingtin and huntingtin fragments within the brain. Several amyloid polymorphs have been identified at the microscopic level, both *in vivo* from cerebral tissue and *in vitro*. Amyloid polymorphism can exhibit a dramatic influence on cytotoxicity. Understanding the conditions and mechanisms that dictate the formation of polymorphs in Huntington’s disease is therefore of considerable importance. We examine the conditions that govern huntingtin exon 1 polymorphism, including the role of the highly conserved huntingtin N-terminal domain that has high biological significance.

3.1 INTRODUCTION

Huntington's disease (HD) is one of several heritable polyglutamine expansion diseases that are characterized by the abnormal expansion of a CAG trinucleotide repeat that codes for a polyglutamine stretch (polyQ) in the mutant protein [19]. In HD, polyQ expansion occurs within the first exon of the huntingtin protein (htt; httEx1); see Figure 3.1, a,b. PolyQ expansion results in the aggregation of htt N-terminal fragments (including httEx1) into insoluble amyloid fibrils that accumulate in neuronal inclusion bodies. The formation of htt fibrils is associated with a toxic gain-of-function that ultimately results in neuronal cell death [100, 207].

HttEx1 is composed of three domains: a 17 residue N-terminal domain (htt^{NT}), a polyQ domain, and the proline-rich domain (PRD). The expanded polyQ domain forms antiparallel β-sheets that make up the amyloid core in mature fibrils [1, 3]. The htt^{NT} domain and the PRD flank the polyQ domain in the httEx1 sequence and have countering influences on the kinetics of httEx1 aggregation [208, 209]. The PRD contains PPII helices and a random coil region [3, 91, 210-212]. The α-helical htt^{NT} domain is very important for htt function, acting as a binding partner for membranes, chaperones including TRiC, and several other proteins with implications in membrane trafficking, cell signaling, and the regulation of gene expression and transcription [3, 90, 91, 200, 210, 213, 214]. Additionally, the htt^{NT} domain is the target of several post-translational modifications that directly influence htt function [199, 215].

As seen in other amyloid-forming proteins, httEx1 fibrils are varied in their morphologies at the microscopic level, and individual polymorphs are thought to be capable of propagating through seeding [3, 47, 56, 104, 216, 217]. Notably, flanking domains are considered crucial for the biological properties of amyloid polymorphs and have been shown to directly affect fibril polymorphism in httEx1 and in other amyloid systems [218-222].

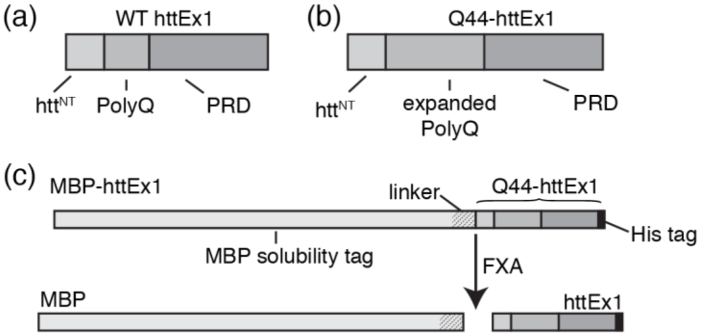


Figure 3.1. Huntingtin exon 1 and MBP based fusion protein

(a) Schematic of wild type httEx1 that does not contain an expanded polyQ domain, and (b) of Q44-httEx1, which contains an expanded polyQ domain. (c) Top: representation of the MBP-httEx1 fusion protein involved in these studies. Bottom: Cleavage with Factor Xa (FXA) protease results in the release of httEx1 and free MBP (bottom). The width of each domain indicated in (a)-(c) is to scale with the length of the primary protein sequence.

We previously reported two temperature-dependent polymorphs of Q44-httEx1 fibrils that can be distinguished by transmission electron microscopy (TEM) [3]. Both polymorphs propagate through seeding and are cytotoxic to neuronal cell lines, which contrasts with other reports showing that polymorphism of httEx1 fibrils can dramatically affect cytotoxicity [3, 47]. It was expected that these differences between the polymorphs are a direct result of polymorphism in the molecular structure or assembly of httEx1 within mature fibrils. Intriguingly, the polyQ fingerprint derived from magic angle spinning solid state NMR (MAS ssNMR) is the same for both temperature-dependent polymorphs that we observed, indicating that the polyQ core is structurally the same [3]. The htt^{NT} and PPII helices in the PRD are immobilized with restricted motion proximal to the polyQ domain [3]. The PRD as a whole presents variable dynamics and solvent accessibility between polymorphs, which affects the ability of the polymorphs to bind antibodies [3]. These results show that the flanking domains are important for Q44-httEx1 fibril polymorphism [3].

The amyloid formation progresses through sigmoidal kinetics [8, 48, 49]. PolyQ peptides that resemble elongated httEx1 progress in the lag phase while forming early oligomeric aggregates prior to the formation of β-sheet rich nuclei [208]. Fibril elongation follows the formation of nuclei and proceeds as rapid polymerization along the fibril axis until the non-fibrillar population of httEx1 is mostly depleted, resulting in an observed plateau phase. Kinetics of amyloid formation are modified by the presence of "seeds", which are preformed fibrils and fibril fragments that can propagate their structure by acting as nuclei for monomers and multimers that have not yet aggregated [8]. Seeds dramatically shorten the lag phase and can be used to mimic conditions present in the brain following the formation of amyloid fibrils [8]. For example, seeding *in vivo* is considered to be responsible for the observation that a single $A\beta_{1\text{-}40}$ polymorph (or 'strain') propogates within an individual Alzheimer's brain despite the fact that multiple strains are observed between patients [44]. The polymorphs of $A\beta_{1\text{-}40}$ can be readily distinguished through TEM and MAS ssNMR [41-46].

Magic angle spinning solid state NMR (MAS ssNMR) is a structural biology tool that is ideally suited to reveal structure and dynamic information of amyloid fibrils at atomic resolution, which is otherwise challenging to obtain by solution NMR nor X-ray crystallography due to innate insolubility and lack of crystallographic order within amyloid fibrils [1, 37, 44, 92, 145-147]. NMR chemical shifts, measured in parts per million (ppm), are exquisitely sensitive to protein structure; thus, the isotopic chemical shifts observed for individual nuclei by MAS ssNMR can be used to rapidly identify protein secondary structure and identify structural changes between two different polymorphs [166].

Our starting point for this work was to understand the temperature-dependent polymorphism we previously observed [3]. Here, we find that polymorphs can be formed at 37 °C that we had not observed previously, indicating that additional experimental conditions that dictate fibril polymorphism. Here, we also investigate the role of the N-terminal flanking and observe that the concentration and cleavage rate of the fusion protein play a critical role.

3.2 METHODS

3.2.1 Protein expression and purification

We employed a previously described MBP-fusion featuing mutant htt exon1 with 44 consecutive glutamine residues within the polyQ domain, subcloned into a pMALc2x plasmid [3, 223]. The fusion protein MBP-Q44-httEx1 was overexpressed in *E. coli* BL21(DE3)pLysS cells (Invitrogen, Grand Island, NY) using an optimized isotopic labeling protocol as previously described, for samples at isotopic natural abundance as well as samples that were uniformly ^{13}C,^{15}N labeled [3]. Cells were harvested by pelleting at 7,000g and then stored at -20 °C overnight. Lysis was done by microfluidization (Microfluidics, Newton, MA) until the lysate appeared nearly clear. Debris was removed by centrifugation at 38,720g for 1 hour and subsequently by filtering over 0.22 μm PES membranes. 1 mM PMSF (phenylmethanesulfonyl fluoride) was added to the resuspended cells just prior to lysis and replenished every 0.5 hour after lysis until the lysate was filtered. The soluble fusion protein was purified using a nickel column and imidazole gradient. The molecular weight, isotopic labeling, and purity were verified by ESI-TOF MS and SDS-PAGE [92]. After exchanging into imidazole-free buffer, the concentration of the fusion protein was determined by its average absorbance (n = 3) at 280 nm, which was measured in a 100 μL quartz cuvette with a 1 cm path length by a DU 800 UV-visible spectrophotometer (Beckman Coulter, Brea, CA). The extinction coefficient of the fusion protein is estimated to be 66350 M^{-1} cm^{-1} (determined using the ProtParam tool by ExPASy) [224].

3.2.2 Fibril formation

MBP-Q44-httEx1 was cleaved to release Q44-httEx1 by treating with Factor Xa (FXA) protease (Promega, Madison, WI), or cleaved to release ΔN15-Q44-httEx1 by treating with N-tosyl-L-phenylalanine chloromethyl ketone treated trypsin lyophilized powder from bovine pancreas (SKU T1426 from Sigma-Aldrich, St. Louis, MO) re-dissolved in PBS buffer with 3 μM HCl. The progression of protein cleavage was monitored by SDS-PAGE (Bio-Rad Mini-Protein Precast TGX Gels 12%) or by ESI-TOF MS as described below [3]. For the TEM studies on Q44-httEx1, MBP-Q44-httEx1 was treated with FXA using MBP-Q44-httEx1 concentration and fusion protein to protease molar ratios of: 98.9 μM/4.25:1, 7.2 μM/25:1, 28.6 μM/25:1, 98. 9 μM/42.5:1, 28.6 μM/62.5:1, 14.3 μM/62.5:1, 14.3 μM/150:1, 78.8 μM/170:1, 28.6 μM/250:1, 88.9 μM/382:1, 10.1 μM/510:1, 98.9 μM/847:1, 14.3 μM/1500:1 as indicated in Figure 3.3 and Figure A.4. The protein was allowed to aggregate at 37 °C over 3 days. For ThT studies on ΔN15-Q44-httEx1, 11.6 μM MBP-Q44-httEx1 was aggregated for 13 days at either 22 or 37 °C, with the MBP-Q44-httEx1 to trypsin ratios of 3:1. ThT seeding assays, TEM imaging, and MAS ssNMR studies were performed on samples of Q44-httEx1 that were prepared at 22 °C from 11.6

μM MBP-Q44-httEx1. Unless otherwise indicated, in all ΔN15-Q44-httEx1 preparations the trypsin cleavage reaction was quenched after 10 minutes with addition of PMSF to a final concentration of 26 nM.

3.2.3 Transmission electron microscopy

Prior to EM imaging, fibrils were pelleted and washed with PBS buffer three times to remove free MBP and un-aggregated monomers and oligomers. Aliquots of fibril suspensions were diluted into PBS buffer and then 5 μL were deposited and adsorbed onto freshly glow-discharged 400 mesh size carbon-coated copper grids for 30 seconds to 2 minutes, and then excess buffer was removed by blotting. Staining agents used were either 1% (w/v) uranyl acetate (UA) or 1% phosphotungstic acid (PTA), as indicated. Negative stain was applied to the grids for approximately 5 seconds, and then removed by blotting, and grids were air dried for 5 minutes. Images of the negatively-stained samples were obtained by transmission electron microscopy on a Tecnai T12 TEM (FEI, Hillsboro, OR) operating at 120 kV and equipped with an UltraScan 1000XP CCD camera (Gatan, Pleasanton, CA). Fibril widths were measured from micrographs using the straight Free-Hand tool of ImageJ (NIH, Bethesda, MD) [187]. Three measurements were obtained per fibril transverse to each fiber axis. In fibrils where the width varied significantly, measurements were obtained at both the thickest and thinnest part of the fibril and also at an intermediate area. In select cases, widths were verified on isolated and vertically aligned fibrils using the Plot Profile tool of ImageJ, which plots average gray scale intensity values across a rectangular region [187]. The width of the resulting peak was measured at its estimated half-height. Fourier transform band pass filters were applied to aid in half-height estimations, and to images of fibril bundles to aid in visibility of individual fibrils, as indicated [187].

3.2.4 X-ray powder diffraction

Fibrils were pelleted by ultracentrifugation at 100,000 g for 4 hours (Beckman Coulter Optima MAX Ultracentrifuge, with TLA 120.2 rotor) and then the supernatant was removed. The hydrated pellet was packed into a glass capillary (0.7 mm) using a spatula and PBS buffer was inserted on both ends of the capillary by syringe. Capillaries were then sealed with wax. A Rigaku FR-E generator (2kW, spot size 0.07 mm) with a source wavelength of 1.541780 Å was placed 65 mm from the sample and diffraction data was collected on a Rigaku Saturn 944 CCD camera (Tokyo, Japan) operating at -45 °C. The sample remained at room temperature during data collection. Diffraction data was analyzed in Structure Studio (Rigaku).

3.2.5 Thioflavin T fluorescence and seeding assay

Thioflavin T (ThT) fluorescence assays were obtained on seeded and unseeded samples of Q44-httEx1 following FXA cleavage as described previously [3]. Fibrils of Q44-httEx1 produced at 22 °C were used as seeds. The seeds were washed with PBS buffer and resuspended by vortex for thirty minutes prior to the seeding assays. For the ThT assays, 11.6 μM of fresh MBP-Q44-httEx1 was cleaved with FXA with a fusion protein to protease ratio of 25:1. Sonicated seeds were added to the reaction mixture immediately at a molar ratio of 20% to Q44-httEx1. The mixture was then vortexed and allowed to aggregate at 22 °C. ESI-TOF MS and HPLC measurements obtained from the supernatant were used to verify that cleavage was complete after 30 minutes [3]. Samples were mixed by vortex at specific time points and aliquots of sample suspensions were immediately diluted into ThT stock solution for a total volume of 400 μL. Samples were excited at 445 nm through a 2 nm excitation slit on a FluoroMax-4 spectrofluorometer (Horiba; Kyoto, Japan). Emission was recorded at 489 nm through a 4 nm emission slit. Two measurements were obtained for each sample per timepoint. Curves were fit using AmyloFit using the Nucleated Elongation Unseeded and Nucleated Elongation seeded models [225].

3.2.6 Cleavage kinetics assays by SDS-PAGE and ESI-TOF MS

The kinetics of monomer release by FXA or trypsin cleavage on MBP-Q44-httEx1 was monitored by SDS-PAGE for samples indicated in Figure A.2 and Figure A.1. Samples (15 μL) were removed from the reaction mixture at the indicated time points and mixed with an equal volume of SDS-PAGE loading dye to terminate the reaction. An SDS-PAGE gel (Bio-Rad Mini-Protein Precast TGX Gels 12%) stained with Coomassie R-250 dye analyzed by the Gel analyzer tool in ImageJ was used to analyze the results [187]. Kinetics of cleavage were also measured in samples with 1:1 molar ratio of MBP-Q44-httEx1 to either FXA protease or trypsin protease, as indicated in Figure 3.6. Aliquots of a few μL were removed and quenched with formic acid at 30 second, 1 minute, 5 minute, and 6 hour time points. ESI-TOF MS was used to analyze the kinetics of cleavage and identify cleavage products, as described below.

3.2.7 Potential trypsin cleavage products of MBP-Q44-httEx1

ExPASy's PeptideCutter tool [224] was used to determine all 42 of the possible trypsin cleavage sites within MBP-Q44-httEx1. These potential cleavage sites were marked onto the predicted structure of free MBP adapted from the structure of fusion protein MBP-L30 (PBD ID 1NMU) using UCSF Chimera [226]. The molecular weights of all fragments were calculated using ExPASy's ProtParam tool [224] assuming 42 active cleavage sites. The molecular weights of all possible fragments assuming selective cleavage sites were calculated by combining the molecular weights of individual fragments and adding 18 Da (the molecular weight of water) for each peptide bond formation. A total of 945 possible cleavage products were identified, 760 of which have molecular weights above 5000 Da.

3.2.8 Mass Spectrometry

The mass of the MBP-Q44-httEx1 fusion protein (56.685 kDa) was confirmed within ± 1 Da by ESI-TOF MS using a MaXis II ESI-QTOF mass spectrometer (Bruker Daltonics, Billerica, MA). Mass spectra were deconvoluted by maximum entropy using Compass Data Analysis Hystar 4.1 software (Bruker Daltonics, Billerica, MA), with lower and upper limits of 5000 and 100,000 m/z. MBP-Q44-httEx1 was deemed usable for additional experiments if the MS deconvolution curve indicated the presence of a single species without any indication of prior protease cleavage.

To identify the trypsin cleavage products of MBP-Q44-httEx1, the fusion protein was cleaved in 7 independent trials and the mass spectra were obtained within the first six hours following the addition of trypsin. PMSF was not added to any of the reaction mixtures. All peaks observed in the resulting deconvoluted mass spectra were screened against the expected molecular weights of 760 possible cleavage products within ± 2 Da. An additional 4 independent trials were analyzed after 1 day of aggregation.

3.2.9 Magic Angle Spinning solid state NMR spectroscopy

Isotopically labeled samples were prepared and then packed by pelleting a hydrated suspension of purified protein aggregates into 3.2 mm zirconia thin wall MAS rotors (Bruker Biospin, Billerica, MA) using a home-built ultracentrifugal packing device under centrifugation at ~ 130,000 g in a Beckman Coulter Optima L-100 XP ultracentrifuge equipped with a SW-32 TI rotor [4]. For ΔN15-Q44-httEx1, 7 mg were packed into the MAS rotor. Caps were sealed to the rotor with epoxy glue to maintain hydration and the samples were subsequently studied by MAS ssNMR in an unfrozen state. All MAS ssNMR experiments were performed using a wide-bore Bruker Avance I NMR spectrometer operating at a ^{1}H Larmor frequency of 600 MHz (14.1 Tesla) and equipped with a 3.2 mm MAS probe with an HCN 'EFree' coil (Bruker Biospin). Immobilized and rigid parts of the fibrils were studied using 1D ^{13}C and 2D ^{13}C-^{13}C Cross polarization (CP) based experiments at 12.5 kHz MAS. ^{13}C-^{13}C correlations were obtained from 2D experiments with a ramped ^{1}H-^{13}C CP step and 8ms dipolar-assisted rotational resonance (DARR) ^{13}C-^{13}C mixing [158]. Dipolar-recoupling curves based on ^{13}C-^{13}C 2D Proton-Driven Spin Diffusion (PDSD) experiments were acquired with mixing times of 0, 25, 250, and 500 ms. The volumes of 2D peaks were integrated using the Sparky NMR software package and normalized to the corresponding peak volume at the diagonal in the PDSD experiment that employed 0 ms mixing [3]. Mobile parts of the fibrils were identified from 1D J-coupling-based ^{13}C spectra which were acquired using rotor-synchronized refocused insensitive nuclei enhanced polarization transfer (INEPT) ^{1}H-^{13}C transfers at 8.333 kHz MAS rate. 2D spectra showing ^{13}C-^{13}C correlations between highly mobile carbons were obtained by combining INEPT ^{1}H-^{13}C transfers with ^{13}C-^{13}C transfers using the $P9^{1}_{3}$ total through bond correlation spectroscopy (TOBSY) pulse sequence [161].

Experimental details for each MAS ssNMR experiment performed are available in Table A.4. For all 1D experiments, two-pulse phase modulation ^{1}H decoupling (typically at 83 kHz) was applied during acquisition, and for all 2D experiments the same ^{1}H decoupling was applied during evolution and acquisition [227]. All spectra were acquired with Bruker Topspin, processed in NMRPipe [228], and analyzed with the CcpNmr Analysis program developed by the Collaborative Computation Project for the NMR community (CCPN). The chemical shifts of ^{13}C and ^{15}N were indirectly referenced to 4,4-dimethyl-4-silapentane-1-1 sulfonic acid and liquid ammonia based on external measurements of the ^{13}C signals of adamantane [92, 229].

3.3 RESULTS

3.3.1 Concentration of MBP-Q44-httEx1 affects width of mature fibrils

We previously reported that fibrillation temperature affects the average width of mature Q44-httEx1 fibrils [3]. ‘Thin’ fibrils (average 5 - 7 nm) are generated at 37 °C and ‘thicker’ fibrils (average 15 – 16 nm) are generated at 22 °C [3]. We performed experiments to probe in more detail the parameters that affect this polymorphism. At increased MBP-Q44-httEx1 concentrations, we observe the formation of even thicker fibrils (> 16 nm) at 37 °C. In order to verify this apparent concentration dependent effect on fibril polymorphism, we performed TEM on three Q44-httEx1 samples generated at 37 °C from 14.3, 28.6, and 98.9 μM MBP-Q44-httEx1. Thin fibrils are generated at 14.3 μM (Figure 3.2, a) [3]. At 28.6 μM, most fibrils are between 5 - 7 nm in width, however a small population of 14 - 17 nm fibrils are observed (Figure 3.2, b). At 98.9 μM MBP-Q44-httEx1, two predominant populations of Q44-httEx1 fibrils are observed with average widths of 5 - 7 and 21 - 23 nm (Figure 3.2, c). These data show that the width of Q44-httEx1 fibrils generated at 37 °C is concentration dependent, and the populations of thicker fibrils increases with the protein concentration.

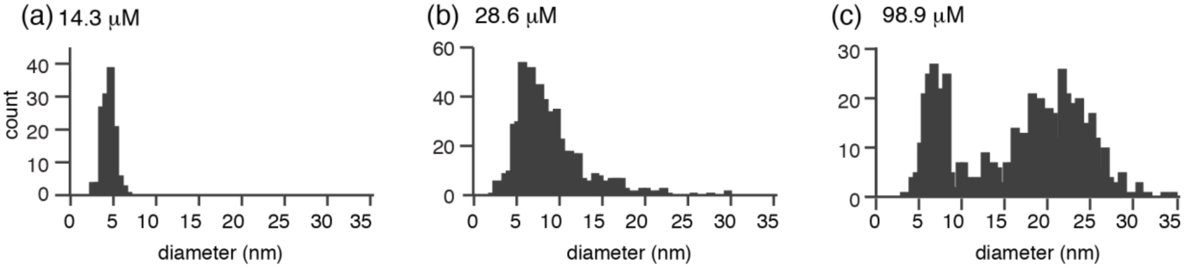


Figure 3.2. Dependence of Q44-httEx1 fibril width on the starting concentrations of MBP-Q44-httEx1 prior to cleavage by FXA at 37 °C

Histogram of fibril widths observed from samples prepared at: (a) 14.3 μM MBP-Q44-httEx1 and FP:P = 62.5, (b) 28.6 μM MBP-Q44-httEx1 and FP:P = 62.5, and (c) 98.9 μM MBP-Q44-httEx1 and FP:P = 42.5.

3.3.2 Kinetics of Q44-httEx1 monomer release by Factor Xa protease

One complication in the above experiment is that they involve aggregation of Q44-httEx1 released from a fusion protein by proteolytic cleavage. As such, the concentration of Q44-httEx1 is actually changing during the assay. Because MBP-Q44-httEx1 is not incorporated into the Q44-httEx1 fibrils [3], we hypothesized that the width of the Q44-httEx1 fibrils is not dependent on MBP-Q44-httEx1 concentration *per se*, but instead dependent on the concentration of free Q44-httEx1 monomer available during fibril formation. To address this, we first analyzed in detail the kinetics of monomer release from MBP-Q44-httEx1. To do so, we monitored the disappearance of the fusion protein and production of free MBP over time from SDS-PAGE gels. Monomer release rates were calculated using the Gel Analyzer tool in ImageJ [187]. The release rates of MBP-Q44-httEx1 by FXA show only negligible difference at 22 and 37 °C (Figure A.2, a).

The molar ratio of the fusion protein (FP) to protease (P). When FP:P is 15,000:1 (FXA), no noticeable monomer release is detectable after 41 hours (Figure A.2, a). In contrast, moderate but incomplete release is detected when FP:P (FXA) is 1500:1 and 150:1, respectively, after 41 hours (Figure A.2, a) [187]. At FP:P (FXA) of 1500:1 and 14.3 μM MBP-Q44-httEx1, significant fibril formation is observed by TEM after 3 days. Interestingly, at this point the monomer release is only ~15% complete after 41 hours and the estimated released Q44-httEx1 monomer concentration at this timepoint is 2.15 μM.

Decreasing the FP:P (FXA) to 2:1 results in >50% release within 5 minutes (Figure A.2, d, lane 2) and complete monomer release in less than 16.5 minutes (Figure A.2, c). At FP:P (FXA) to 5:1 and 10:1, incomplete release is observed at 15 minutes (Figure A.2, d) but complete release is expected within 30 minutes. We note that at 14.3 μM MBP-Q44-httEx1 and FP:P (FXA) of 5:1, the lag phase lasts for approximately 6 hours (Figure A.3, b). Therefore, the concentration of Q44-httEx1 monomer can reasonably reach 14.3 μM early during the lag phase. Interestingly, the rate of monomer release also appears to be dependent on the length of the polyQ domain; release from MBP-Q20-httEx1 is significantly faster rate than release from MBP-Q44-httEx1 (Figure A.2, e).

3.3.3 The rate of Q44-httEx1 release from MBP-Q44-httEx1 affects sample heterogeneity

After developing a better understanding of Q44-httEx1 monomer release kinetics, we performed a systematic analysis of the fibril morphology at several FP:P (FXA) values across a range of initial MBP-Q44-httEx1 concentrations. Histograms depicting the distribution of fibril widths observed for each sample are shown in Figure A.4. Trends derived from these histograms are presented in the scatterplot in Figure 3.3. The FP:P ratio and the concentration of MBP-Q44-httEx1 both dramatically affect the microscopic morphology of the resulting mature Q44-httEx1 fibrils.

‘Thin’ fibrils (defined as having an average width <9 nm) form at 7.2- 28.6 μM MBP-Q44-httEx1 when FP:P is between 250:1 and 25:1 (quadrant III, blue circles). ‘Thick’ fibrils (defined as having an average width > 16 nm) are produced from 78.8 - 98.9 μM MBP-Q44-httEx1. Apparent branching is observed in Figure 3.4, a,b (additional examples are available in Figure A.5, a,b). A substantial population of thin fibrils is also observed in these samples. ‘Intermediate’ fibrils (defined as having an average width between 9 and 16 nm) are observed at 10.1 and 14.3 μM MBP-Q44-httEx1 with FP:P of 510:1 and 1500:1, respectively (quadrant I, red triangles). Interestingly, thin fibrils but not thick fibrils are produced at 98.9 μM MBP-Q44-httEx1 when FP:P is 847:1 (quadrant II).

A notable feature of the thicker fibrils is that they show apparent branching (Figure 3.4, a,b, highlighted). Striations likely indicating individual filaments are often observed across the fibril axis of many of the thick fibrils, as indicated in an isolated and vertically aligned section of a thick fibril (Figure 3.4, d). These striations are confirmed by plotting the gray value across the fibril axis of the isolated section; three peaks are observed (Figure 3.4, e). Unlike the thick fibrils, the surface of both the thin and intermediate fibrils frequently appears fuzzy and this fuzziness may obstruct the visibility of individual filaments within these fibrils.

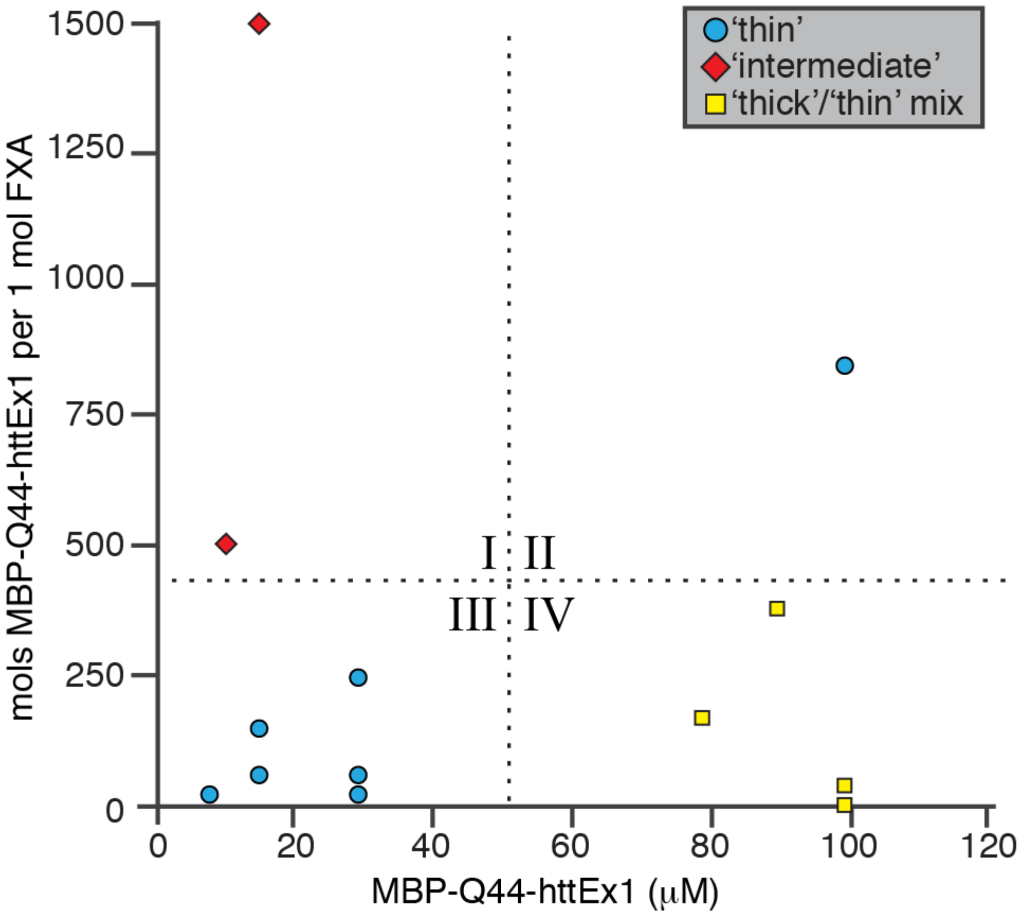


Figure 3.3. Classification of fibril polymorphs by average width in relation to the initial concentration of MBP-Q44-httEx1 and the Fusion Protein to Protease ratio

Scatterplot of conditions used during fibril formation, comparing the effect of the concentration of MBP-Q44-httEx1 vs the moles of MBP-Q44-httEx1 available per mole of FXA protease. Blue circles indicate conditions where a majority of the fibrils were found to be 'thin' by TEM, red diamonds indicate conditions where a majority of the fibrils were found to be 'thick', and yellow squares represent conditions where the mixture of 'thin' and 'thick' fibrils was considerably heterogeneous (see raw data in Figure A.4). Dashed lines are used for visual purposes only to divide the results into quadrants.

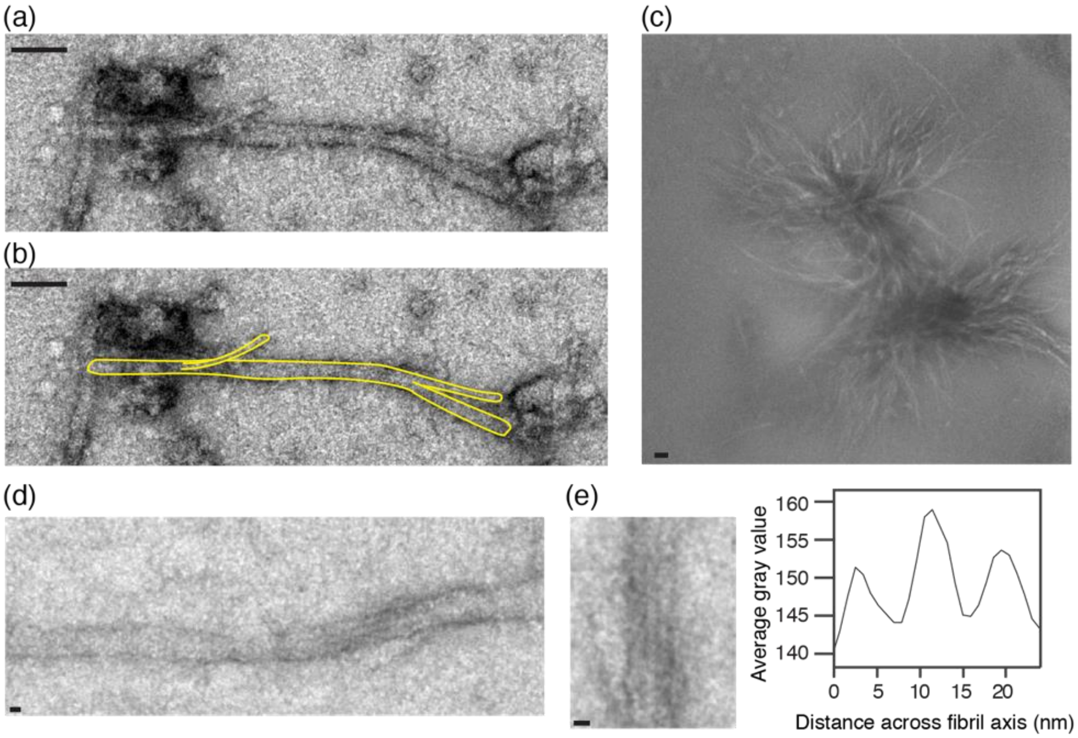


Figure 3.4. Fibril heterogeneity as observed by TEM

(a,b) Thick fibril with apparent branching, with and without highlights. Scale bars are 100 nm(c) Fibril bundle, with band pass filter applied to the image to balance contrast. The original image is available in Figure A.5. Scale bar is 100 nm. (d) Thick fibril with visible striations. (e) Striations from (d) are confirmed by plotting the average gray value vs the distance across the fibril axis. Scale bar in panels (d) and (e) are 10 nm.

Large bundles of fibrils with frayed ends are observed in many of the samples, seemingly independent of MBP-Q44-httEx1 concentration and FP:P ratio (Figure 3.4, c). Fibril bundles appear extremely dark when imaged, so a Fourier transform band pass filter was applied to the image to balance contrast and deemphasize the density of the bundles while still allowing individual fibrils splayed from the bundles to be visible. The original image and an alternatively filtered image with enhanced contrast is available in Figure A.5, c. Fibrils of several lengths were observed, however long fibrils on the order of 1 μm in length were primarily observed in bundles while most unbundled fibrils appeared to be on the order of 0.1 to 0.5 μm.

3.3.4 Heterogeneity between and within Q44-httEx1 fibrils

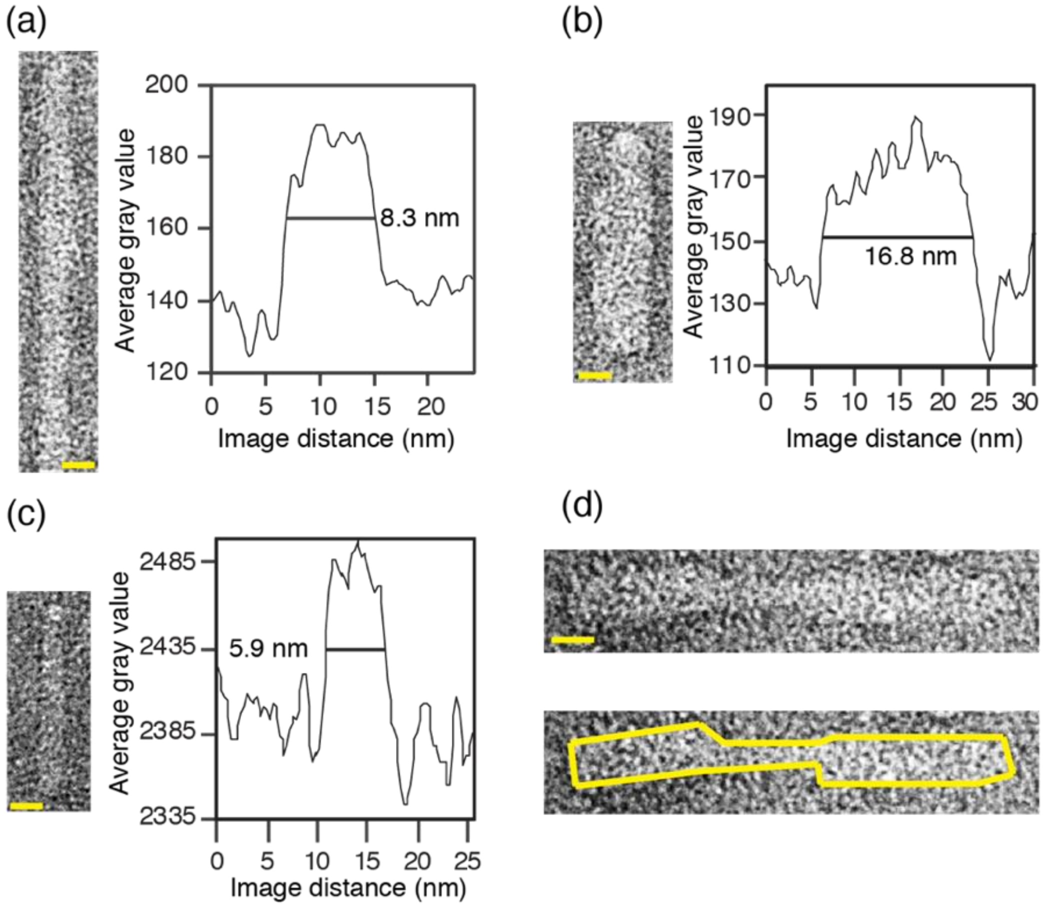


Figure 3.5. TEM images of Q44-httEx1 fibrils formed at 28.6 μM MBP-Q44-httEx1 and Fusion Protein to Protease ratio of 62.5:1

(a-c) Width length analysis of individual fibrils shows there is heterogeneity of fibril widths within the sample. (d) Fibril with varying width along the fiber axis.

The Q44-httEx1 fibrils within a single sample demonstrate heterogeneity in fibril width, as shown in Figure 3.5. Although the majority of the fibrils observed in this sample are 'thin', fibrils with widths up to 30 nm were observed (Figure A.4). An accurate estimate of individual fibril widths from this sample can be achieved by isolating and vertically aligning individual fibrils and plotting the average gray value across the Y axis of the image vs the image width (Figure 3.5, a-c). Low gray values indicate areas of the grid where stain is present. Two examples of thin fibrils with average widths of 8.3 and 5.9 nm are shown in Figure 3.5, a, and Figure 3.5, c. An intermediate fibril with a width of 16.8 nm is shown in Figure 3.5, b. In Figure 3.5, d, we show an example of a fibril with varying width across the fibril axis; the center of the fibril is thicker than either end. We noted several instances of the width varying across the fibril axis under most sample conditions. The increase in width at either end appears in quantified steps, as opposed to a gradual change.

3.3.5 Trypsin protease releases ΔN15-Q44-httEx1 from MBP-Q44-httEx1 with high activity

Prior work has proposed the existence of two competing aggregation mechanisms, driven either by the polyQ of the htt^{NT} domain [218-222]. Given that these competing mechanisms were concentration dependent, we considered the possibility that this plays a role in polymorphism. The htt^{NT} domain is known to accelerate the aggregation kinetics of Q44-httEx1, presumably by htt^{NT} driven oligomerization that forces an increased local concentration of expanded polyQ [91, 106, 230]. We set out to study how eliminating the htt^{NT} domain would affect fibril polymorphism.

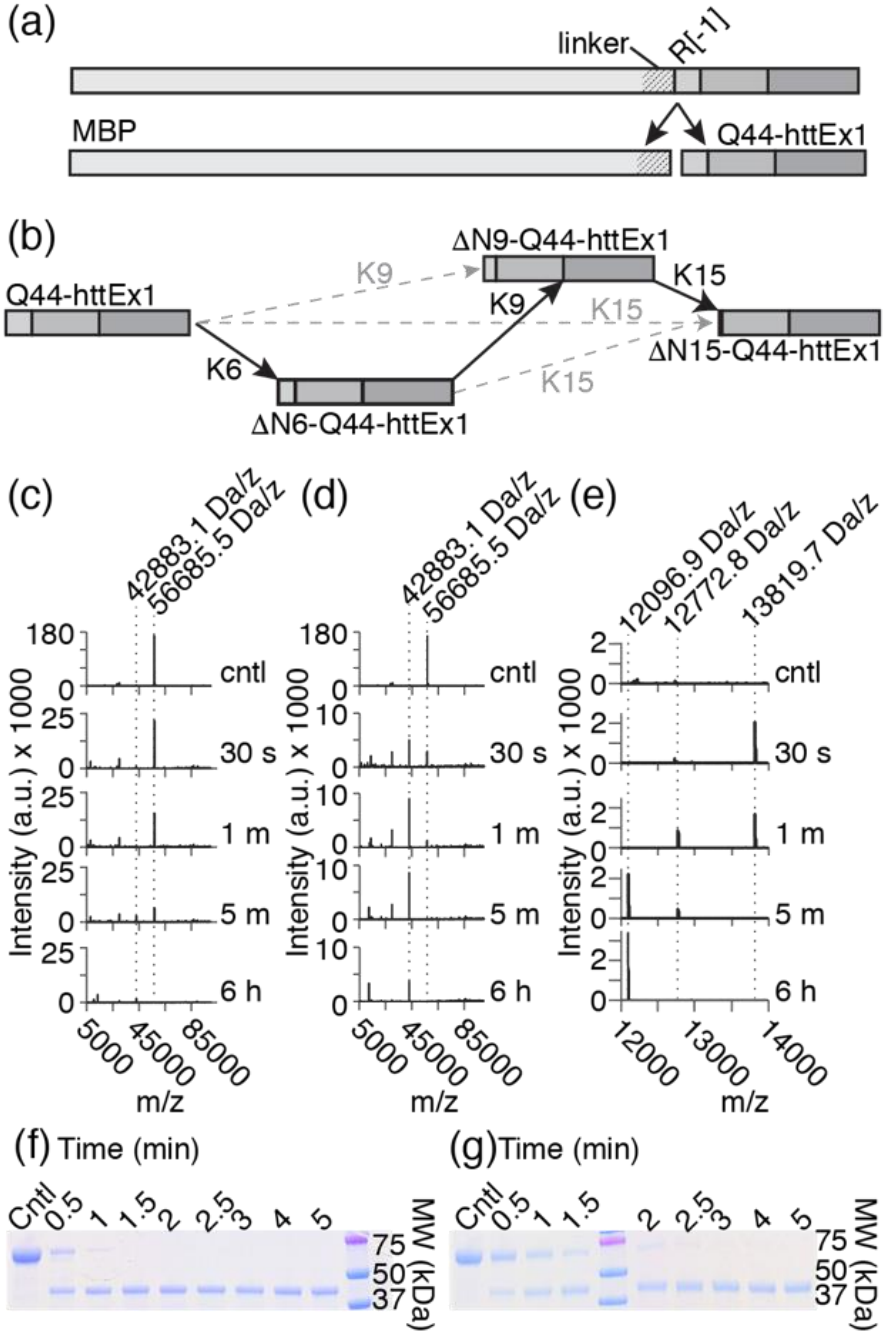


Figure 3.6. Trypsin protease cleavage of MBP-Q44-httEx1.

(a) MBP-Q44-httEx1 is cleaved at R[-1] into free MBP and Q44-httEx1. (b) The htt exon1 monomer is subsequently cleaved into ΔN6-Q44-httEx1 and ΔN9-Q44-httEx1 at K399 and K402, respectively. ΔN6-Q44-httEx1 and ΔN9-Q44-httEx1 are cleaved at K408 to form ΔN15-Q44-httEx1. (c,d) ESI-TOF mass spectra of 1 μM MBP-Q44-httEx1 (56685.5 Da) cleavage via 1 μM FXA (c) or 1 μM trypsin (d), at 30 sec, 1 min, 5 min, and 6 hours. In (c), MBP-

Q44-httEx1 is observed at 30 sec, 1 min, and 5 min but is not observed at 6 hours, and free MBP (42883.1 Da) is observed at 5 min. In (d), release of free MBP from MBP-Q44-httEx1 is complete by 5 min and there is noticeable release of free MBP detected within the first 30 sec. (e) Zoomed in mass spectra of (d) between 12000 and 14000 m/z. Q44-httEx1 and ΔN9-Q44-httEx1 are observed after 30 sec trypsin cleavage. By 5 min, Q44-httEx1 is no longer detectable, and ΔN15-Q44-httEx1 is observed. By 6 hours, only ΔN15-Q44-httEx1 is observed. (f) Cleavage of 10 μM MBP-Q44-httEx1 with 10 μM trypsin protease: MBP-Q44-httEx1 to trypsin molar ratio of 1:1. (g) Cleavage of 10 μM MBP-Q44-httEx1 with 0.16 μM trypsin protease: MBP-Q44-httEx1 to trypsin molar ratio of 625:10. Full gels for (f) and (g) are available in Figure A.1.

To do so, we considered the use of trypsin cleavage sites near the htt^{NT}/polyQ boundary. To develop a better understanding of the trypsin cleavage protease we investigated the cleavage kinetics and the cleavage products. In Figure 3.6, the trypsin cleavage is examined by SDS-PAGE and ESI-TOF MS. The MBP-Q44-httEx1 fusion protein is mostly depleted after 30 seconds and is no longer observed at 5 minutes. Free MBP and Q44-httEx1 (13819.7 Da) are initially observed at 30 seconds in Figure 3.6, e, but soon the Q44-httEx1 peak decreases. At 5 minutes, Q44-httEx1 is no longer observed and ΔN15-Q44-httEx1 is instead observed at 12096.9 Da, which becomes the dominant species at the 6 hour time point. Additional data and more detailed analysis is in section A.1 and Figure A.1. No additional Q44-httEx1 products are observed, however a total of five MBP cleavage products are observed by SDS-PAGE (Figure A.1, e) and ESI-TOF (Figure A.2). An in-depth analysis of these cleavage products is available in section A.2.

Altogether, our data reveal that trypsin preferably cleaves MBP-Q44-httEx1 at residue R[-1] (the last residue in the fusion protein prior to Q44-httEx1) to release Q44-httEx1 (Figure 3.6, a), but it then sequentially and efficiently cleaves Q44-httEx1 from the htt^{NT} N-terminus to C-terminus at K6, K9, and K15 (Figure 3.6, b). Thus, under these conditions, trypsin cleavage allows us to rapidly generate ΔN15-Q44-httEx1, which is expected to be unable to assemble via the htt^{NT} pathway.

3.3.6 Aggregates of ΔN15-Q44-httEx1 fibrils

We removed the first 15 amino acids of the 17 residue htt^{NT} domain with trypsin and monitored amyloid formation with Thioflavin T (ThT) in the presence and absence of Q44-httEx1 seeds that were originally generated by FXA proteolysis [3, 63]. We observed that that the ThT fluorescence is ~2.5x lower when FXA protease is utilized for this seeding assay in place of trypsin (Figure A.3, a) and seeding behavior is retained (Figure A.3, b). This suggests a structural change upon the removal of the htt^{NT} domain that allows the ThT dye to more readily bind to the rigid polyQ amyloid core.

TEM images of the resulting ΔN15-Q44-httEx1 fibrils demonstrate that the majority of fibrils have widths between 8-15 nm and an average width is 11.6 nm (Figure 3.7, b). There is a small population of fibrils with a width exceeding 15 nm. The width of the ΔN15-Q44-httEx1 fibrils also varies along the fiber axis in some fibrils, however at much lower frequency than observed for Q44-httEx1. Both isolated fibrils and bundled fibrils with frayed ends are observed. Fibrils associated with bundles are approximately 1 μm in length, similar to what was observed for Q44-httEx1 fibrils. Fibrils that are not associated with bundles are shorter, on the order of 0.1 μm in length. Striations were visible in a small number of fibrils; however, it is difficult to say whether these striations are indicative of multiple filaments interacting within one fibril (Figure

3.7, e). Apparent branching is not observed, potentially highlighting the role for the htt^{NT} domain in the process.

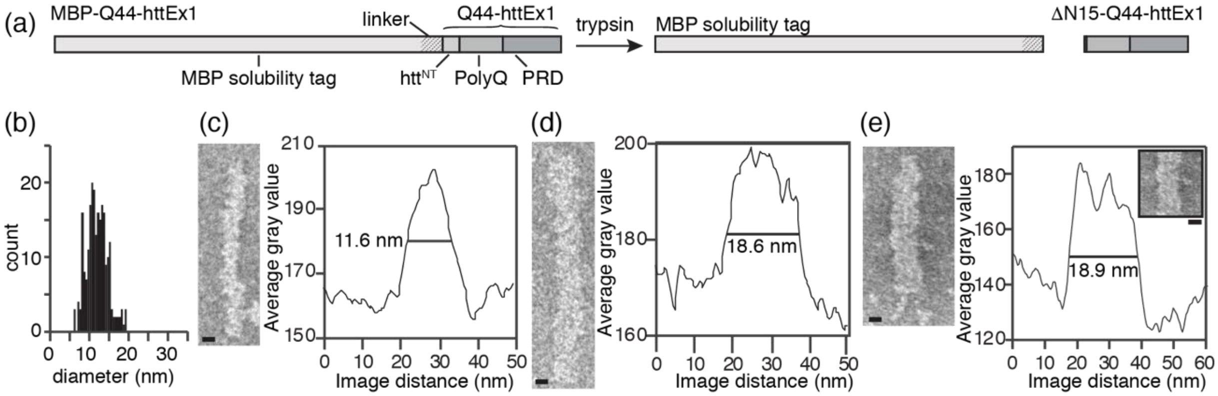


Figure 3.7. TEM of [U-^{13}C-^{15}N] ΔN15-Q44-httEx1 fibrils

(a) MBP-Q44-httEx1 releases free MBP and ΔN15-Q44-httEx1 upon cleavage by trypsin protease. (b) Distribution of widths measured for [U-^{13}C-^{15}N] ΔN15-Q44-httEx1 fibrils. The average width measurement was between 11-12 nm, however there was a wide distribution of widths observed. (c-e) Width length analysis (right) of individual fibers (left) reveals heterogeneity in widths between fibrils. The analysis in (e) was done on the straightened inset.

3.3.7 Comparison between Q44-httEx1 and ΔN15-Q44-httEx1 fibrils using MAS ssNMR

Next, we used MAS ssNMR to probe the structure and dynamics of these amyloid fibrils. 1D ^{13}C cross-polarization (CP) experiments were used to compare the rigid residues of Q44-httEx1 and ΔN15-Q44-httEx1 fibrils, both aggregated at 22 °C (Figure 3.8, a). The spectra area very similar, however more peaks are visible in the spectrum of Q44-httEx1. Missing peaks previously attributed to htt^{NT} are consistent with the removal of the first 15 residues of the 17 amino acid htt^{NT} in the ΔN15-Q44-httEx1 fibrils. The aliphatic peaks (between 10 - 80 ppm) that are present in both samples overlap in their chemical shifts. The most intense peaks in this region also appear with similar intensities in the two samples.

To get a better understanding of what peaks are missing in the ΔN15-Q44-httEx1 spectrum, the ΔN15-Q44-httEx1 CP spectrum was subtracted from the Q44-httEx1 spectrum (Figure 3.8, a, overlay in green). When normalized to the signals in the Q44-httEx1 spectrum and compared directly to the CP spectra of both fibril types, key differences are observed. There is a noticeable difference observed in the carbonyl/carboxyl region (between 165 and 190 ppm), primarily a result of the htt^{NT} truncation. A small positive shoulder associated with the carbonyl chemical shift for residue A2 is marked in the subtraction spectrum, indicating that A2 is absent in ΔN15-Q44-httEx1 (Figure 3.8, b) [1, 3]. It is difficult to assign the remaining peaks in this region due to signal overlap. More insight can be interpreted from the aliphatic region of the subtraction spectrum. Residue types A2, E5, K6, and A10 are marked as positive peaks. A broad positive peak can also be assigned to the typical Cα chemical shift, indicative of a loss of Cα nuclei from the htt^{NT} domain in the ΔN15-Q44-httEx1 sample. A negative peak is assigned to Pro residues.

Solution-like experiments based on the INEPT scheme were used to identify flexible residues (Figure 3.8, c). INEPT experiments capture signals from residues within the aggregates with similar mobility to soluble protein. Like the CP spectra, the peaks that appear in the aliphatic region of both INEPT spectra overlap in their chemical shift and peak intensities of these peaks are relatively similar. Intensity for the Pro residues do not differ significantly. We note a decrease in intensity in the peaks in the aromatic region of the spectra (80 - 165 ppm). Peaks in this region can only be assigned to two Phe residues, both of which are present in the htt^{NT} domain, and His residues primarily present in the His-tag.

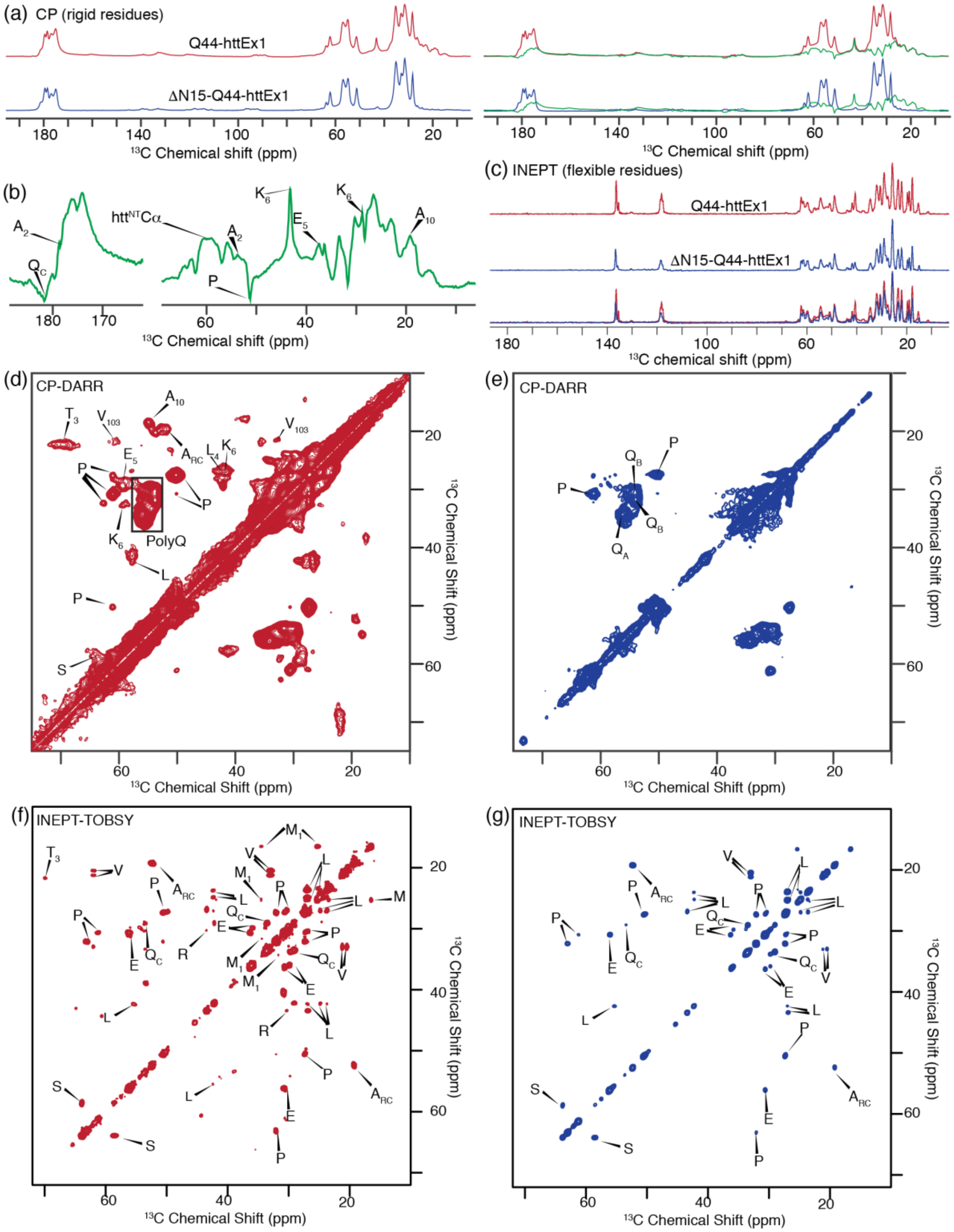


Figure 3.8. Comparison of ΔN15-Q44-httEx1 and Q44-httEx1 by Magic Angle Spinning solid state NMR

(a) Top, left: 1D cross polarization spectrum of Q44-httEx1 [3] and (bottom, left) ΔN15-Q44-httEx1, and (right) their comparison to the subtraction between ΔN15-Q44-httEx1 and Q44-httEx1, normalized to the peak intensities

of the CP spectrum for Q44-httEx1. (b) Subtraction ΔN15-Q44-httEx1 and Q44-httEx1, zoomed in on the carbonyl (left) and aliphatic (right) regions. (c) From top to bottom: ^{13}C *INEPT of Q44-httEx1, ΔN15-Q44-httEx1, and overlaid. (d) DARR spectrum of Q44-httEx1 [3] and (e) ΔN15-Q44-httEx1. (f)* ^{13}C*-*^{13}C *INEPT-TOBSY spectrum of 'thick' Q44-httEx1 fibrils formed at 22 °C [3]. (g)* ^{13}C*-*^{13}C *INEPT-TOBSY of ΔN15-Q44-httEx1. The spectra of Q44-httEx1 in panels (a), (d), and (f) are reprinted with permission from Lin, H. K.,* ***Boatz, J. C.****, Krabbendam, I. E., Kodali, R., Hou, Z., Wetzel, R., Dolga, A. M., Poirier, M. A., van der Wel, P. C. A. Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core, Nature Communications 8:15462. Copyright (2017) Lin, H. K., et. al. [3]*

More insight into structural differences between the ΔN15-Q44-httEx1 and Q44-httEx1 fibrils can be gained with two dimensional spectra, e.g. the ^{13}C-^{13}C dipolar assisted rotational resonance (DARR) spectra. For Q44-httEx1 (Figure 3.8, d), several peaks are assigned to the htt^{NT} domain [1, 3]. A large glutamine signal is observed that corresponds with the chemical shifts of glutamine types 'a', 'b', and 'c' [1, 3]. Several peaks are assigned to an α-helical htt^{NT} [1, 3, 210]. The most prominent cross peaks attributed to the PRD in the ^{13}C-^{13}C CP-DARR spectrum of Q44-httEx1 are attributed to a PPII helix structure, with additional signals from a random coil structure (Pro$_{RC}$) [1, 3, 210].

The 2D ^{13}C-^{13}C DARR spectrum of ΔN15-Q44-httEx1 (Figure 3.8, e) overwhelmingly consists of Gln and Pro cross-peaks; several cross-peaks assigned in the Q44-httEx1 DARR spectrum are absent in the DARR spectrum of ΔN15-Q44-httEx1. Notably, htt^{NT} signals are missing. Additionally, peaks representing an alanine residue with a random coil secondary structure (A$_{RC}$) in addition to two peaks representing V103 are absent. Both these residues are within the random coil region of the PRD in Q44-httEx1. Interestingly, the intensity of the peaks representing proline residues with a PPII secondary structure (Pro$_{PPII}$) is decreased for ΔN15-

Q44-httEx1 in the ^{13}C-^{13}C DARR. Peaks representing proline residues with a random coil secondary structure (Pro_{RC}) are missing completely. However, Pro_{RC} peaks do occur in a proton driven spin diffusion (PDSD) experiment with 250 ms of ^{13}C-^{13}C mixing (Figure A.6). PDSD experiments are extremely similar to DARR experiments, however subtle differences between PDSD and DARR spectra are common. The PDSD experiment shown in Figure A.6, optimized to observe long-range intermolecular contacts shows no differences compared to Q44-httEx1 [1, 3]. Thus, the immobilized core of the polyQ and PRD is largely unchanged in structure. However, some indications of dynamic changes are observed. As expected, signals assigned to the htt^{NT} are missing.

2D ^{13}C-^{13}C INEPT-based total through bond correlation spectroscopy (TOBSY) spectra were also obtained to aid in residue assignments. There are more peaks observed in the 2D TOBSY spectrum of Q44-httEx1 (Figure 3.8, f) than in ΔN15-Q44-httEx1 (Figure 3.8, g) [3]. The missing peaks are consistent with residues found in the htt^{NT} domain. Peaks assigned as Arg residues that cannot be assigned to the htt^{NT} are also missing in the ΔN15-Q44-httEx1 2D spectra when compared to the Q44-httEx1 spectra. We note that one small signal is observed in the TOBSY spectrum of ΔN15-Q44-httEx1 that we previously assigned to M1 for Q44-httEx1 [3]. This is likely the result of a slight incomplete trypsin cleavage, or an indication that this nuclei was incorrectly assigned previously [3].

3.3.8 Dynamics of the PRD of ΔN15-Q44-httEx1 fibrils

Previously, we noted dynamic differences in the immobilized core of Q44-httEx1 polymorphs, detected by CP-based experiments [3]. We performed an analogous series of PDSD experiments with variable mixing times from 0 - 500 ms in order to compare ΔN15-Q44-httEx1 to Q44-httEx1 [3]. Polarization transfer between directly bonded Cα-Cβ carbons is rapid in a rigid lattice, with maximum polarization transfer (10 - 20%) reached within 10 - 20 ms [3].

As before, such a fast buildup curve is observed for Cα-Cβ glutamine peaks within the amyloid core of Q44-httEx1 fibrils [3]. This indicates that the core of the fibrils experiences crystal-like rigidity. However, the one-bond polarization buildup curves for PPII and $\mathrm{Pro_{RC}}$ residues in the PRD of ΔN15-Q44-httEx1 experience a much slower buildup and lower peak maximum, indicating that the PRD experiences less rigidity (Figure A.7, c) [3]. A more dramatic difference was observed for two-bond Cα-Cγ polarization transfers, signifying that the $\mathrm{Pro_{RC}}$ residues experience more dynamics than those with a PPII structure. The buildup curves measured for the PRD of the ΔN15-Q44-httEx1 are available in Figure A.7, b. The polarization buildup profile of $\mathrm{Pro_{RC}}$ has a lower peak maximum than what was observed for Q44-httEx1. The PPII residues also appear to have a lower peak maximum, but this effect is much less pronounced. Additionally, peaks representing V103 and an Ala residue in the PRD are missing in the DARR spectrum. Peaks with low intensity associated with $\mathrm{Ala_{RC}}$ are observed in the PDSD spectrum, however peaks associated with V103 do not. Combined, these MAS ssNMR results show that the PRD is more dynamic in ΔN15-Q44-httEx1 than in Q44-httEx1.

3.4 DISCUSSION

3.4.1 Threshold concentration of Q44-httEx1 monomer necessary to induce the production of thick Q44-httEx1 fibrils

We found that the concentration of MBP-Q44-httEx1 and FP:P (FXA) ratio are both important for controlling the production of specific Q44-httEx1 polymorphs at 37 °C. We interpret this to reflect a dependence on the concentration of free Q44-httEx1 monomers in the reaction mixture. A threshold concentration of free Q44-httEx1 monomer is necessary to induce the formation of thick fibrils, rather than the thinner fibrils.

3.4.2 Comparison of ΔN15-Q44-httEx1 and Q44-httEx1 fibrils

Trypsin was found to significantly improve the cleavage rate of MBP-Q44-httEx1, allowing for near immediate release of ΔN15-Q44-httEx1 monomers. As expected, aggregation proceeded at a slower rate than observed in Q44-httEx1. Additionally, improved ThT fluorescence indicates a likely structural change between ΔN15-Q44-httEx1 and Q44-httEx1 fibrils. Trypsin cleavage of MBP-Q44-httEx1 results in the release of ΔN15-Q44-httEx1 monomers that contain only two residues from the htt^{NT} domain. We consider these residues insufficient for multimer formation to be driven by the htt^{NT} domain and we expect polyQ driven multimer formation to prevail.

The ΔN15-Q44-httEx1 fibrils were found to have an average width of 11.6 nm, unlike any of the fibril types observed for Q44-httEx1 polymorphs generated at 37 °C, and some thick fibrils (17 - 20 nm) are also observed. Isolated fibrils were on average several times shorter in length when compared to isolated Q44-httEx1 fibrils. The ΔN15-Q44-httEx1 fibrils were also observed to vary in width across the length of the fibril, display visible striations, and form bundles with frayed ends. However, the ΔN15-Q44-httEx1 fibrils were not observed to form branches. The population of thicker fibrils and fibrils with striations indicates that the ΔN15-Q44-httEx1 fibrils can be composed of multiple filaments. Because the htt^{NT} domain is missing, free monomers are expected to interact with each only other through the polyQ domain, leading to a symmetric distribution of PRD domains across the length of mature fibrils, which in turn discourages branching sites. Secondary nucleation events can also only occur between the PRDs of preformed filaments and free monomers. The propensity for fibrils to form secondary nucleation sites does not appear to be disrupted upon removal of the htt^{NT} domain; this is consistent with our prior findings that filaments interact through the PRDs [3].

The PRD domains within the ΔN15-Q44-httEx1 fibrils experience similar rigidity to what we previously reported in Q44-httEx1 fibrils formed at 22 °C [3]. These 22 °C Q44-httEx1 fibrils are predicted to be composed of two filaments each interacting through the PRD [3]. Accordingly, we interpret that the majority of the fibrils in the ΔN15-Q44-httEx1 sample are composed of filaments interacting through the PRD. If there was only one filament in the ΔN15-Q44-httEx1 fibrils, the PRD would experience high dynamics similar to the thin 37°C Q44-httEx1 fibrils and would not mimic the dynamics we observe in the 22 °C Q44-httEx1 fibrils. The average width of ΔN15-Q44-httEx1 fibrils is approximately 4 nm less than the average width of the 22 °C Q44-httEx1 fibrils, likely as a result of the removal of the htt^{NT} domain.

Strikingly similar peak patterns in the CP-DARR and INEPT-TOBSY spectra indicate that atomic level structure of the polyQ and PRD within ΔN15-Q44-httEx1 and Q44-httEx1 fibrils is the same. The most notable difference observed by MAS ssNMR between the ΔN15-Q44-httEx1 and 22 °C Q44-httEx1 samples is in the random coil region of PRD. Particularly, the $\mathrm{Pro}_{\mathrm{RC}}$ residues are not present in the DARR spectrum of ΔN15-Q44-httEx1 under the same conditions they are present in the DARR spectrum of Q44-httEx1 that was aggregated at 22 °C. These residues appear in PDSD spectrum at longer mixing time, indicating that the $\mathrm{Pro}_{\mathrm{RC}}$ residues are less rigid in the ΔN15-Q44-httEx1 sample. PDSD-derived polarization transfer curves support the idea that the random coil region of the PRD is more dynamic in the ΔN15-Q44-httEx1 fibrils, however the PPII helices experience similar dynamics between samples. Altogether, the MAS ssNMR results suggest that the ΔN15-Q44-httEx1 fibrils are very similar in structure and dynamics to the 22 °C Q44-httEx1 fibrils which we previously reported, with the most significant difference noted in the random coil region of the PRD [3].

3.4.3 The effect of monomer polymorphism on fibril polymorphism

Misfolded monomers of Q44-httEx1 may adopt a variety of structures. Prior literature describes both extended both native-like (Figure 3.9, b, left) [231] and tadpole-like (Figure 3.9, c, left) [232] compact monomers. Native-like monomers are thought to interact through the $\mathrm{htt}^{\mathrm{NT}}$ domains (Figure 3.9, b, right) [233], while tadpole-shaped monomers are expected to interact through the polyQ domain (Figure 3.9, c, right) [232]. We expect these monomer polymorphs to vary in their stability and propensities to form multimeric intermediates. We predict native-like monomers are stable at low monomer concentrations and tadpole-like monomers are transient; at higher concentrations, tadpole-like monomers are expected to be stabilized through intermolecular interactions [232]. We propose that the polyQ driven formation of multimers leads to the formation of fibrils with an extremely symmetric distribution of flanking domains along the fiber axis (Figure A.8, a, left), and that htt^{NT} driven formation of multimers leads to an extremely asymmetric distribution (Figure A.8, b, left).

3.4.4 Fibril polymorphism may be explained by secondary nucleation events on the fibril surface

Fibril surfaces can act as seeds to promote the conversion of monomers to nuclei through secondary nucleation events [8]. As long as monomers are continuously introduced into the system, secondary nucleation events are expected to be frequent once fibrillation begins. We interpret the fuzziness of the thin and intermediate Q44-httEx1 fibrils observed by TEM to be indicative of several secondary nucleation surface events followed by the growth of short protofilaments (Figure 3.9, d, left, and Figure A.8, a,b, left). Apparent fibril widening occurs at high fusion protein concentrations and low FP:P ratios where the rate of monomer release is high. This leads to an increased monomer concentration that we expect to promote the elongation of the nucleated protofilaments into mature filaments that remain associated with the fibril originally used as a nucleation site (Figure 3.9, d, center and Figure A.8, a,b, center). This results in fibrils made up of multiple interacting filaments from a mechanism that is fundamentally different than preformed filaments coming together to form a thicker fibril.

If the distribution of flanking domains on the nucleating filament is extremely symmetric, and tadpole shaped Q44-httEx1 monomers are predominant in solution, then we expect the elongation of the nucleated filament to be uninhibited by steric hindrance from the flanking domains (Figure A.8, a, center). Nucleation with elongation is also possible when the distribution of flanking domains is extremely asymmetric and the predominant Q44-httEx1 monomer type is native-like. However, steric hindrance is likely to inhibit elongation of protofilaments nucleated from asymmetric filaments after the addition of just a few monomers (Figure A.8, a,c). A mixture of both native-like and tadpole-like monomers likely contributes to the formation of fibrils, leading to a semi-random distribution of flanking domains over the length of a filament (Figure 3.9).

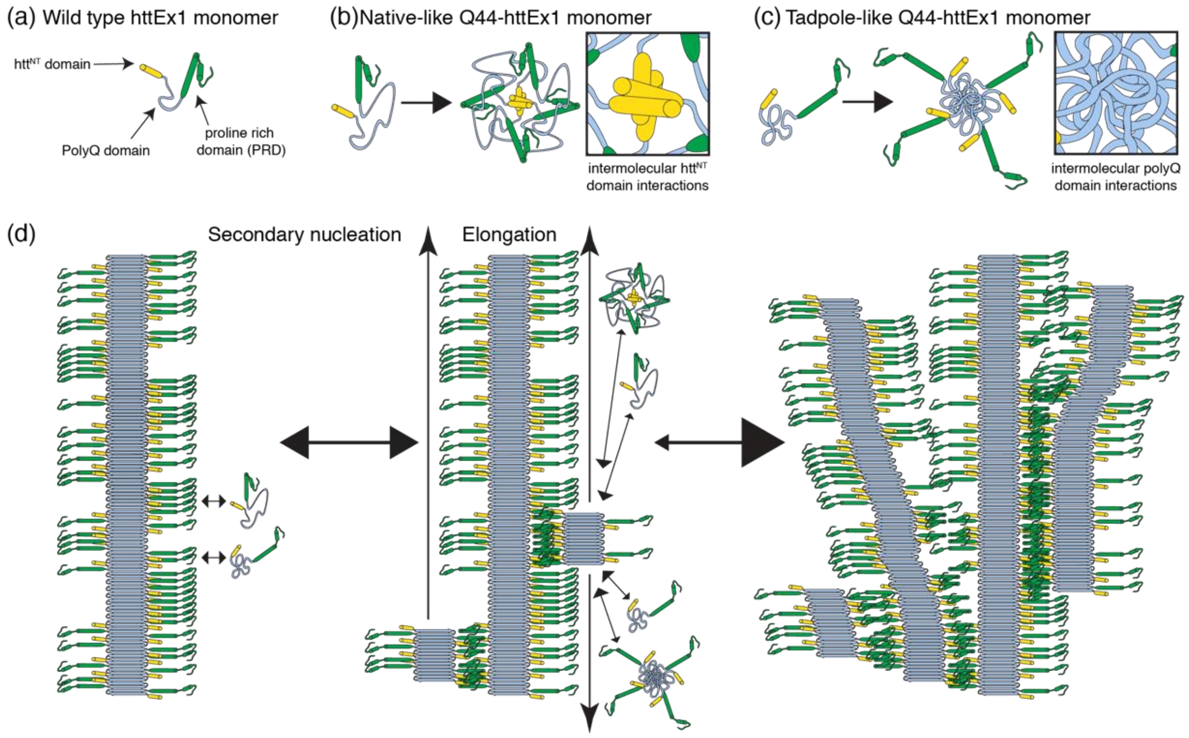


Figure 3.9. Hypothetical models of Q44-httEx1 fibril growth

(a) Cartoon of wild type httEx1 and its individual domains: htt^{NT} *(yellow), polyQ (blue), and PRD (green). (b) Clustering of native-like Q44-httEx1 monomers via the htt*^{NT} *domains [233]. (c) Clustering of 'tadpole-like' Q44-httEx1 monomers via the polyQ domains [232].*

Apparent fibril widening is expected when the concentration of Q44-httEx1 is high. This idea is supported by the fact that three filaments appear to be present in the thick fibrils, as indicated by the observation of striations by TEM. We note that the appearance of striations often “disappeared” when moving across the fiber axis and we interpret this to mean that filaments are very tightly interacting in these regions and less tightly interacting in the regions where striations are observed (Figure 3.9, d, right). The strength of the interactions between the filaments is dictated by the distribution of flanking domains over the length of the filament. Clusters where the flanking domains are extremely symmetric should promote filament-filament interactions, whereas asymmetric clusters inhibit the interactions. This idea supports the observation that the width of individual fibrils was often observed to vary across the length of the fibril. Secondary nucleation events can explain also the apparent fibril branching that was observed in the thick fibrils; we propose that branch points occur in locations on the nucleating fibril that contain an asymmetric cluster of flanking domains. We note that fibril branching has been reported for httEx1 fibrils previously [234].

The model explained above does not account for the apparent contradictory observation that fibrils with an intermediate width are preferentially formed at low MBP-Q44-httEx1 concentrations and high FP:P ratios. We would expect under this condition that thin fibrils composed of one filament each would be favored. Instead, the intermediate fibrils appear strikingly similar to the fibrils we previously reported that are formed at 22 °C, which we determined were likely to be composed of two filaments each [3]. The presence of MBP-Q44-httEx1 may also influence the distribution of native-like and tadpole-like monomer structures, which can in turn affect the propensity for secondary nucleation events to occur on fibril surfaces. As MBP is shown not to interact with httEx1, the interactions between Q44-httEx1 and MBP-Q44-httEx1 are not expected to involve the polyQ domain instead of the MBP solubility tag [235]. The propensity for httEx1 monomer to interact with the MBP-Q44-httEx1 surface may also be temperature dependent, which could explain the intermediate fibril morphology we observe in fibrils formed at 22 °C.

3.4.5 Conclusion

Together, our results indicate that polymorphism is driven not only by temperature, but also by the concentration of monomer in solution. The htt^{NT} domain does not appear to largely affect polymorphism and is not necessary for secondary nucleation events, however the htt^{NT} domain does appear to be important for initiating fibril branching. Overall, these results are important because they provide a better understanding of fibril polymorphism and yield some insight into the mechanisms that dictate fibril formation at low concentrations, as may exist *in vivo*. Amyloid fibril formation is expected to undergo complicated kinetics under conditions that mimic physiological sample mixtures, which can contain monomers, several intermediate species, intermediates, off-pathway aggregates, amyloid fibrils, and fibril fragments. These species may further modulate the kinetics of amyloid formation and the morphology and properties of the formed aggregates. This is an important topic that warrants further investigation.

3.5 ACKNOWLEDGEMENTS

We thank Drs. Mingyue Li, Abhishek Mandal, and Jinwoo Ahn and also Christine Monnie for helpful discussions related to experimental design and data analysis. We also acknowledge Audrey Valentine and James Nassur for their contributions to preliminary experiments towards this manuscript. We thank Michael Delk for technical assistance with the NMR spectrometers, and Dr. Alexander Makhov for his help and the use of the departmental electron microscope facility. This work was enabled by funding from the University of Pittsburgh, and National Institutes of Health (NIH grants R01 GM112678 to P.C.A.v.d.W., T32 GM088119 to J.C.B.), the Achievement Rewards for College Scientists (ARCS) Foundation (J.C.B.) and grant UL1 RR024153 from the National Center for Research Resources (NCRR).

3.6 AUTHOR CONTRIBUTIONS

J.C.B. and P.C.A.v.d.W. designed experiments; J.C.B., T.P., and I.M. conducted experiments; J.C.B., T.P., I.M., J.F.C., and P.C.A.v.d.W. analyzed data; J.C.B. wrote the paper.

4.0 CATARACT-ASSOCIATED P23T γ-D-CRYSTALLIN RETAINS A NATIVE-LIKE FOLD IN AMORPHOUS-LOOKING AGGREGATES FORMED AT PHYSIOLOGICAL PH

Adapted and reprinted from: **Boatz, J.C.**, Whitley M.J., Li M, Gronenborn A.M., and van der Wel, P.C.A. Cataract-associated P23T γD-crystallin retains a native-like fold in amorphous-looking aggregates formed at physiological pH, *Nature Communications 8*, 15137. Copyright (2017) Boatz, J. C., et. al. under Creative Commons Attribution 4.0 International License.

Data availability: The ssNMR chemical shifts of P23T hγD aggregated at pH 7 that were assigned are available in Table B.2, and have been submitted to the BMRB.

I prepared samples and performed all transmission electron microscopy, X-ray diffraction, thioflavin T, and aggregation experiments. I also conducted the magic angle spinning solid state NMR experiments with Dr. Patrick C. A. van der Wel. I directly contributed data that presented in Figure 4.2, Figure 4.3, Figure 4.4, Figure 4.5, Figure 4.6, Figure 4.7, Figure 4.8, Figure 4.9, and Figure 4.10, and I prepared Figure 4.1, Figure 4.2, Figure 4.3, Figure 4.4, Figure 4.5, Figure 4.6, Figure 4.7, and Figure 4.8. I designed the transmission electron microscopy and

> X-ray diffraction experiments performed on the samples prepared at pH 7, all experiments performed on the samples prepared at pH 3, and all thioflavin T experiments.

Cataracts cause vision loss through the large-scale aggregation of eye lens proteins as a result of aging or congenital mutations. The development of new treatments is hindered by uncertainty about the nature of the aggregates and their mechanism of formation. We describe the structure and morphology of aggregates formed by the P23T human γD-crystallin mutant associated with congenital cataracts. At physiological pH, the protein forms aggregates that look amorphous and disordered by electron microscopy, reminiscent of the reported formation of amorphous deposits by other crystallin mutants. Surprisingly, solid-state NMR reveals that these amorphous deposits have a high degree of structural homogeneity at the atomic level and that the aggregated protein retains a native-like conformation, with no evidence for large-scale misfolding. Non-physiological destabilizing conditions used in many *in vitro* aggregation studies are shown to yield qualitatively different, highly misfolded amyloid-like fibrils.

4.1 INTRODUCTION

Cataract formation is the leading cause of blindness, and age-related and congenital cataracts affect close to 100 million people across the globe [108]. The underlying cause of eye lens opacification is the formation of protein aggregates [107]. As such it is but one example of a broad family of protein-aggregation-related diseases, in which normally-soluble proteins end up as insoluble protein deposits in affected organs [236]. Current treatments of protein aggregation diseases (including cataracts) are limited to the invasive removal of affected tissue or treatment of symptoms, rather than a reversal, prevention or inhibition of the underlying aggregation process. A limiting factor in efforts to achieve the latter is our poor understanding of the molecular mechanism by which the protein aggregation occurs.

The α-, β-, and γ-crystallins, which constitute 90% of the total eye lens protein mass, are found in cataract-associated protein deposits. The α-crystallins are small heat shock proteins that perform protein quality control and counteract protein aggregation, misfolding, and cataract formation [118, 237]. The β- and γ-crystallins are thought to be responsible for generating and maintaining the proper refractive index of the eye lens [238]. To do so, the β/γ-crystallins are present at concentrations of 400 mg/mL in healthy lenses. Cataract formation occurs when the high solubility of β/γ-crystallins is compromised as a result of changes induced by mutations or chemical damage that accumulates with aging [107, 108]. However, it has proved difficult to relate specific mutations and structural changes to a molecular mechanism of cataract-associated protein aggregation [65-67, 69-71, 239]. X-ray crystallographic and solution NMR studies have yielded atomic structures for many of the structurally related β/γ-crystallins, including both wild-type (WT) proteins and proteins with either disease-associated mutations or mutations mimicking aging-induced damage [24, 121, 130, 240, 241]. Figure 4.1 shows the structure of the P23T mutant of human γD-crystallin (hγD), with the Greek key motifs that characterize the native fold of the γ-crystallins [24, 241]. As examined in detail in prior work [24, 69, 130, 242, 243], this mutant protein is structurally very similar to the WT protein. This is reflected in a backbone RMSDs of 0.4 and 0.5 Å for the N-terminal (NTD; cyan) and C-terminal domains (CTD; pink), respectively (comparing the X-ray structures), or solution NMR per-residue chemical shift differences (Δδ) of less than 0.2 ppm (except for the 0.4 ppm Δδ for the residue after the P23T mutation) [24, 241]. Nonetheless, the mutant protein has lost the remarkable solubility of WT hγD and is associated with autosomal dominant congenital cataracts that form early in a child's development, affecting families across the globe [24, 69, 116, 130, 242, 243].

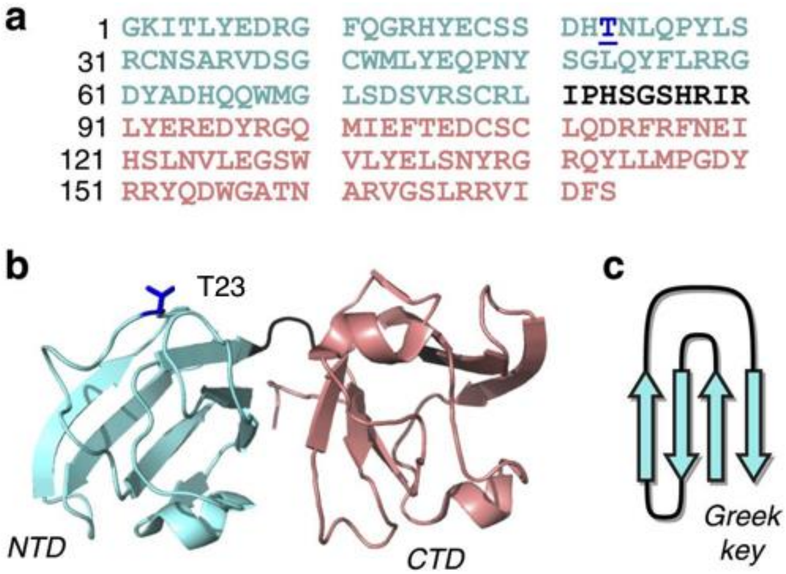


Figure 4.1. Structure of P23T hγD

(a) Primary sequence and (b) solution NMR structure [24]. The mutation site is underlined and highlighted dark blue. The NTD (cyan) and CTD (pink) make up two independently folded symmetric domains, separated by a short linker (black). (c) Schematic of the β-sheet-based Greek key motifs present in each domain.

Based on *in vitro* studies, two prominent mechanisms have been put forward to explain the aggregation propensity of mutated or modified crystallins associated with cataract formation. These either focus on misfolding-induced aggregation, in analogy to the amyloid formation associated with neurodegenerative diseases [236, 239], or suggest that crystallin solubility is lost in absence of extensive misfolding [65-67, 69-71]. Depending on the aggregation mechanism, mutated or damaged β/γ-crystallins are predicted to yield protein aggregates with distinct molecular structures. As a result, there is much interest in the structure of aggregated β/γ-crystallins. However, no atomic resolution structural data are currently available for any such aggregate, due to the fact that they are not suitable for solution NMR or X-ray crystallographic study. A further complicating factor is that even the reported morphology of the aggregates varies, a situation similar to the aggregate polymorphism commonly observed among amyloidogenic proteins [244-247]. Two distinct types of crystallin aggregates are widely reported *in vitro*, which are generally categorized as either 'amorphous' or 'fibrillar' [248-252]. This exact nomenclature reflecting two distinct types of protein aggregates is also reported for other aggregation-prone proteins [244-247]. Thus, this appears to be a more general phenomenon not constrained to cataract-related protein aggregation. Amorphous aggregates are often assumed to be disordered or poorly structured, while the amyloid-like fibrils are typically described as more ordered [248, 249]. It is worth noting that the nm-scale morphology of cataract-associated protein aggregates formed *in vivo* is also under debate, although the balance of evidence disfavors a large role for amyloid-like fibrils [107, 253, 254]. Previously reported arguments in favor of amyloid formation *in vivo* are largely based on the ability of mouse model cataracts to bind amyloid-binding dyes [253]. However, such dyes are not a foolproof measure of amyloid structure as both false positives and false negatives occur [46, 255]. This is borne out by studies that show dye binding by both non-cataract eye lenses and α-crystallins in a non-amyloid state [249, 255, 256].

To gain insight into the structural features of congenital-cataract-related protein aggregates we here investigate these aggregates by magic-angle-spinning (MAS) solid-state NMR (ssNMR) spectroscopy and complementary techniques. MAS ssNMR has emerged as the tool of choice for atomic-resolution studies of non-crystalline protein aggregates [1, 257, 258]. In particular, we study distinct aggregated states of the abovementioned P23T hγD that is associated with congenital cataracts (Figure 4.1). This mutant protein forms both worm-like fibrils and amorphous-looking aggregates, depending on the experimental conditions. We observe that the former is generated at low pH and display all the characteristics of highly misfolded amyloid-like fibrils. The amorphous-looking aggregates form at physiological pH and lack all of the amyloid hallmarks. Despite their appearance, these aggregates are shown by ssNMR to contain the aggregated protein in a well-ordered native-like conformation, with no evidence of dynamic or static disorder.

4.2 METHODS

4.2.1 Expression and purification of P23T hγD

The mutant P23T hγD gene (*CRYGD*) was cloned into a pET14b vector as described previously [24]. Transformed One Shot *Escherichia coli* BL21star (DE3) cells were grown in LB broth or M9 minimal media at 37 °C to an OD_{600} of 0.6, induced with 1 mM IPTG, and allowed to overexpress at 37 °C for four hours. Isotopically labeled protein was obtained by growing cells in modified M9 minimal medium containing 2 g/L U-^{13}C-D-glucose and 1 g/L ^{15}N-ammonium chloride (Cambridge Isotope Laboratories, Tewksbury, MA). Cell pellets were resuspended in QA buffer (50 mM Tris at pH 8.2, 1 mM EDTA, 1 mM dithiothreitol (DTT), and lysed via microfluidization (Microfluidics, Newton, MA). Nucleic acids were removed from the cell lysate by precipitation with 1% (w/v) polyethyleneimine, and cell debris was removed by centrifugation at 39,000 g for 20 minutes (Sorvall RC 3C Plus, H-6000A rotor). The protein was purified essentially as described previously [24, 241]. Briefly, clarified cell lysate was loaded onto a pre-equilibrated HiTrap Q XL anion exchange column (GE Healthcare) in QA buffer, and the flow-through fraction was collected and dialyzed overnight at 4 °C into SA buffer (10 mM MES at pH 6.0, 1 mM EDTA, 1 mM DTT, 2% v/v glycerol). Precipitates were removed from the dialyzed flow-through by centrifugation at 39,000 g for 20 minutes. The supernatant was loaded onto a pre-equilibrated HiTrap SP cation exchange column (GE Healthcare) in SA buffer and eluted using a linear gradient of 0 to 1 M NaCl. Fractions containing P23T hγD were collected and further purified by gel filtration on a Superdex 75 26/60 column (GE Healthcare) in 100 mM sodium phosphate buffer (pH 7), containing 5 mM DTT and 0.02% (wt/vol) NaN_3. The concentration of the final purified protein was determined with a NanoDrop 1000 spectrophotometer (Thermo Scientific) using an A_{280} molar extinction coefficient of 42.86 mM^{-1} cm^{-1}.

4.2.2 Preparation of P23T hγD aggregates

To prepare aggregates under acidic conditions, 10% and then 1% (w/w) HCl was added drop-wise to 1.28 mg/mL (62 μM) solutions of P23T hγD on ice, until the pH reached 3.0 [85]. The sample was immediately vortexed after the addition of each HCl droplet to insure proper mixing and also to avoid any local areas of low pH within the sample. After adjusting the pH to 3.0, the protein was allowed to aggregate overnight or longer, at 37 °C. To prepare aggregates under neutral conditions (pH 7), P23T hγD was concentrated to > 10 mg/mL (0.49 mM) over four 10-min and one 5-min steps using 3 kDa Centriprep centrifugal filters spinning at 3,200 g (Eppendorf 5810 R, A-4-81) at 4 °C in 100 mM sodium phosphate buffer (pH 7) containing 5 mM DTT and 0.02% (wt/vol) NaN_3. The protein concentration was measured occasionally using UV absorbance until the sample appeared cloudy by eye (Figure 4.2) or reached the desired concentration. Subsequently, the protein was allowed to aggregate overnight at 37°C [243]. [U-^{13}C-^{15}N]-labeled P23T hγD aggregates were formed using the same conditions and concentrations (in mg/mL) used to form the unlabeled aggregates.

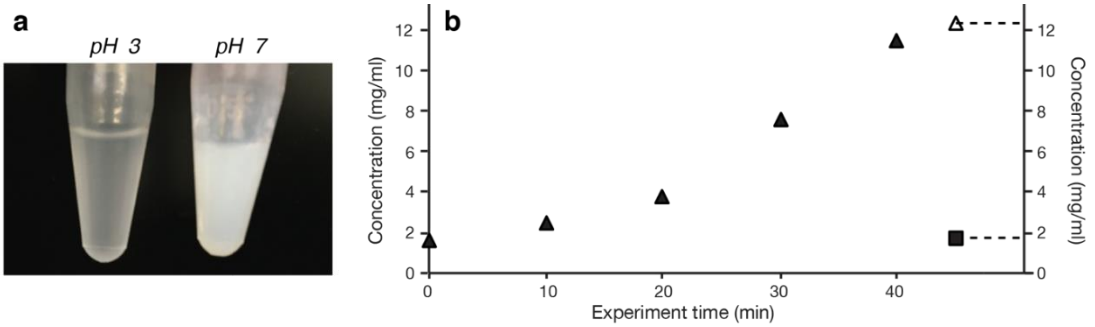


Figure 4.2. P23T hγD aggregation

(a) Appearance of P23T hγD aggregates. Left: 1.3 mg/mL suspension of fibrils formed at pH 3. Right: 9.2 mg/mL suspension of aggregates formed at pH 7. (b) Concentration-dependent aggregation of P23T hγD at neutral pH. Starting with an initial concentration of 1.8 mg/mL, the solution was gradually concentrated at 4 °C using a centrifugal concentrator. Triangles mark the total protein concentration determined by UV-VIS spectrometry after different times of centrifugal concentration. Solid triangles indicate time points prior to the onset of aggregation. The open triangle indicates the total protein concentration at the onset of aggregation (>12 mg/mL). The solid square shows the residual monomer concentration after the aggregation has completed (1.7 mg/mL). Note that the solubility of wild-type hγD is at least one to two orders of magnitude larger, with measurements ranging from 170 to more than 400 mg/mL depending on conditions.

4.2.3 Transmission electron microscopy

Aggregates of P23T hγD formed under acidic (pH 3) and neutral (pH 7) conditions were imaged using transmission electron microscopy. Samples (5 μL) were adsorbed on freshly glow discharged 400 mesh size carbon-coated copper grids for 15 seconds, and the excess was removed by blotting with filter paper. The sample grids were subsequently stained with 1% (w/v) sodium phosphotungstate for 5 seconds and blotted. Grids were imaged at 11,000-fold (pH 7) and 21,000-fold (pH 3) magnification using a Technai T12 transmission electron microscope

(FEI; Hillsboro, OR) operating at 120 kV and equipped with an UltraScan 1000 CCD camera (Gatan; Pleasanton, CA).

4.2.4 X-ray powder diffraction

Aggregates of P23T hγD formed under acidic (pH 3) and neutral (pH 7) conditions were pelleted by centrifugation. Aggregates formed under neutral conditions were pelleted with a tabletop centrifuge at 14,000 g for 20 minutes in a table-top VWR Galaxy 14D centrifuge, while aggregates formed under acidic conditions were pelleted by ultracentrifugation at 100,000 g for 4 hours (Beckman Coulter Optima MAX Ultracentrifuge, with TLA 120.2 rotor). The majority of the supernatant was removed, and the pellet was resuspended in the remaining buffer and packed into a glass capillary (0.7 mm) using a syringe. Capillaries were sealed with wax to retain hydration of the samples. Diffraction data was measured at room temperature with a Rigaku Saturn 944 CCD camera (Tokyo, Japan). A Rigaku FR-E generator (2 kW, spot size 0.07 mm) was used as the X-ray source. Aggregates formed under neutral conditions were exposed for 75 seconds, while aggregates formed under acidic conditions were exposed for 210 seconds. Diffraction data were processed and analyzed in Structure Studio (Rigaku).

4.2.5 Thioflavin T fluorescence measurements

Suspensions of approximately 25 μg of aggregates formed under acidic (pH 3) and neutral (pH 7) conditions were diluted 20x to 200x (depending on the pre-aggregation concentration) into a ThT stock solution (50 μM ThT, 10 mM sodium phosphate, 150 mM NaCl, pH 7). Samples were excited at 445 nm and fluorescence emission was recorded at 489 nm using a Fluoromax-4 Research Spectrofluorometer (Horiba; Kyoto, Japan). Excitation and emission slits were 2 nm and 4 nm, respectively. In addition, ThT fluorescence measurements were also performed on a systematic series of P23T hγD aggregates formed at pH 3 to pH 8 in 0.5 pH increments. To prepare these aggregates, the pH of a low concentration solution of P23T hγD (0.13 mg/mL) was adjusted using 10% and 1% (w/w) HCl and 0.5 M NaOH. Samples were allowed to aggregate at 37 °C for five days, after which approximately 2.5 μg (20 μL) of the resuspended P23T hγD aggregates were diluted to 400 μL in the same ThT buffer mix, for each ThT fluorescence measurement.

4.2.6 MAS ssNMR spectroscopy

Hydrated [U-^{13}C-^{15}N]-labeled P23T hγD aggregates formed under acidic (pH 3) and neutral (pH 7) conditions were separately packed into regular- and thin-walled 3.2 mm zirconia MAS rotors (Bruker Biospin, Billerica, MA). The sample packing was done by pelleting the hydrated sample suspension directly into the MAS rotor using an ultracentrifugal packing device, as previously described [4]. This was done by centrifugation at ~130,000 g for 2-3 h in a Beckman Coulter Optima L-100 XP ultracentrifuge with a SW-32 Ti rotor. Excess supernatant was removed, the fully hydrated sample was sealed, and subsequently studied by ssNMR in an unfrozen and hydrated state. All MAS ssNMR experiments were performed with a widebore Bruker Avance I NMR spectrometer operating at a ^{1}H Larmor frequency of 600 MHz (14.1 T) using a 3.2 mm MAS ssNMR probe with an HCN “EFree” coil (Bruker Biospin). A constant flow of cooled gas was used to control the sample temperature. Spectra were acquired with Bruker Topspin software, processed in NMRPipe, and analyzed with Sparky software and the CcpNmr Analysis program developed by the Collaborative Computation Project for the NMR community (CCPN) [259, 260]. A 90^o shifted sine bell followed by a Lorentzian-to-Gaussian transformation function was applied in all three dimensions of the NCACX and NCOCX spectra during processing. ^{13}C and ^{15}N chemical shifts were indirectly referenced to 4,4-dimethyl-4-silapentane-1-1 sulfonic acid (DSS) and liquid ammonia, respectively, based on external reference measurements of the ^{13}C signals of adamantane [261]. Cross-polarization (CP) spectra were used to study the immobilized parts of the protein aggregates. Note that soluble proteins would not be visible in the CP experiments, such that the signal must be from the aggregated protein. 1D ^{13}C CP spectra were acquired at a 12.5 kHz MAS rate using ramped ^{1}H-^{13}C CP. Intra-residue through-space ^{13}C-^{13}C correlations were obtained from 2D experiments that combined ramped ^{1}H-^{13}C CP with 8 ms dipolar assisted rotational resonance (DARR) ^{13}C-^{13}C mixing [158]. Partial assignments (Table B.2) were obtained for the aggregates formed under neutral conditions, with the aid of the ^{13}C-^{13}C spectra as well as 2D and 3D ^{15}N-^{13}C correlation experiments (NCO, NCA, NCOCX, NCACX), measured at 10 and 12.5 kHz MAS. These experiments involved 1.5 ms of ^{1}H-^{13}N CP, 5.0 ms of ramped ^{13}C-^{15}N SPECIFIC CP (spectrally induced filtering in combination with cross polarization) [162] and 15.0 ms of ^{13}C-^{13}C DARR mixing, where applicable. The rotor-synchronized refocused INEPT (insensitive nuclei enhanced by polarization transfer) scheme was used to acquire 1D J-coupling-based ^{13}C spectra at a 8.333 kHz MAS rate [156]. These experiments employed τ_1 = 1.2 ms and τ_2 = 1.0 ms for aggregates formed at neutral pH, and τ_1 = 1.5 ms and τ_2 = 1.0 ms for the fibrils formed at pH 3. 2D spectra showing through-bond ^{13}C-^{13}C correlations between highly mobile carbons were obtained using a combination of ^{1}H-^{13}C INEPT and 6.0 ms total through bond correlation spectroscopy (TOBSY) ^{13}C-^{13}C mixing [161]. Two-pulse phase modulation (TPPM) ^{1}H decoupling of typically 83 kHz was applied during acquisition for all 1D experiments, and during acquisition and evolution for 2D experiments [227]. Additional experimental details can be found in Table B.1.

4.2.7 Synthetic MAS ssNMR spectra

Synthetic NMR spectra were generated for both native and amyloid-like reference structures using methods similar to those used in our prior work [262]. Spectra representative of the native protein structure were generated starting with the published solution NMR chemical shifts of soluble P23T hγD [24]. Alternatively, a hypothetical completely parallel β-sheet structure of P23T hγD was built using the PyMOL Molecular Graphics System (Schrödinger, LLC), from which chemical shifts were predicted using the SPARTA+ program [263]. Subsequently, synthetic 2D NMR spectra showing the predicted one-bond cross-peaks were generated using utilities from the NMRPipe software package [259]. The solution NMR peak markers in Figure 4.9 were generated from the solution NMR chemical shifts [24] using the synthetic peak list creation routines of the CcpNmr Analysis program [260].

4.3 RESULTS

4.3.1 Polymorphic aggregation of P23T hγD

First, we compare the aggregation that occurs under physiological and denaturing conditions. At physiological pH (pH 7), the 20.6 kDa P23T hγD is soluble at dilute concentrations but undergoes a temperature-sensitive aggregation at 10-15 mg/mL concentration (Figure 4.2). Negative-stain transmission electron microscopy (TEM) reveals the aggregates to be non-fibrillar, with an amorphous appearance (Figure 4.3, a). Under acidic conditions (pH 3), we observe aggregation of P23T hγD, even at 1.3 mg/mL, which yields 7 nm-wide worm-like unbranched fibrils (Figure 4.3, b). In fluorescence assays using the amyloid-binding dye thioflavin T (ThT), the two aggregates show a dramatically different response (Figure 4.3, c). Consistent with their fibrillar appearance, the acid-induced aggregates lead to high ThT fluorescence, while the amorphous-looking aggregates do not. In a systematic ThT-binding screen across a broad range of pH values (Figure 4.3, d) high ThT fluorescence response is observed at both low and high pH, but not under the neutral pH conditions present in eye lenses [264]. ThT assays are not foolproof measures of amyloid structure [46, 255]. Therefore, we also performed X-ray powder diffraction measurements on the hydrated aggregates. *Bona fide* amyloids exhibit a characteristic cross-β diffraction pattern, dominated by inter-β-sheet and intra-β-sheet repeat distances of ~4.7 and ~8-10 Å, respectively (e.g. see the polyglutamine data in Figure 4.3, g [1]). This amyloid hallmark is observed for the aggregates formed at pH 3 (Figure 4.3, e), but is absent for those that form at pH 7 (Figure 4.3, f). The latter show a different diffraction pattern, featuring 12.6, 16.3 and 28 Å repeat distances. Thus, the fibrils formed upon exposure to acidic denaturing conditions have all amyloid hallmarks, but the amorphous-looking aggregates formed under more physiological conditions feature none of them.

4.3.2 The amorphous-looking aggregates have high internal order

MAS ssNMR was used to probe the molecular structure of both types of deposits prepared from uniformly ^{13}C and ^{15}N (U-^{13}C,^{15}N)-labeled P23T hγD. First, cross-polarization (CP) ssNMR experiments are used to select for the signals of only the immobilized parts of protein deposits. The 1D ^{13}C CP spectrum of the pH 7 aggregates is shown in Figure 4.4, a. The observed peaks have relatively narrow widths, as low as 35 Hz for methyl side chain peaks. More detailed information can be gleaned from 2D spectra, such as the ^{13}C-^{13}C spectrum in Figure 4.4, b. In this spectrum, the off-diagonal cross-peaks reflect pairs of carbons that are connected via one or two bonds. Despite the amorphous appearance of the aggregates, the spectrum shows many well-resolved peaks with remarkably narrow peak widths. The narrow linewidths indicate a high degree of structural homogeneity (i.e. little structural variation among individual molecules throughout the sample) and a lack of line-broadening dynamics. It appears that the crystal-like structural homogeneity applies to much or all of the aggregated protein, given that these narrow peak widths are seen throughout the spectrum. The 2D spectrum is also characterized by a high degree of peak dispersion, which is normally associated with a well-folded state, especially for proteins with multiple types of secondary structure.

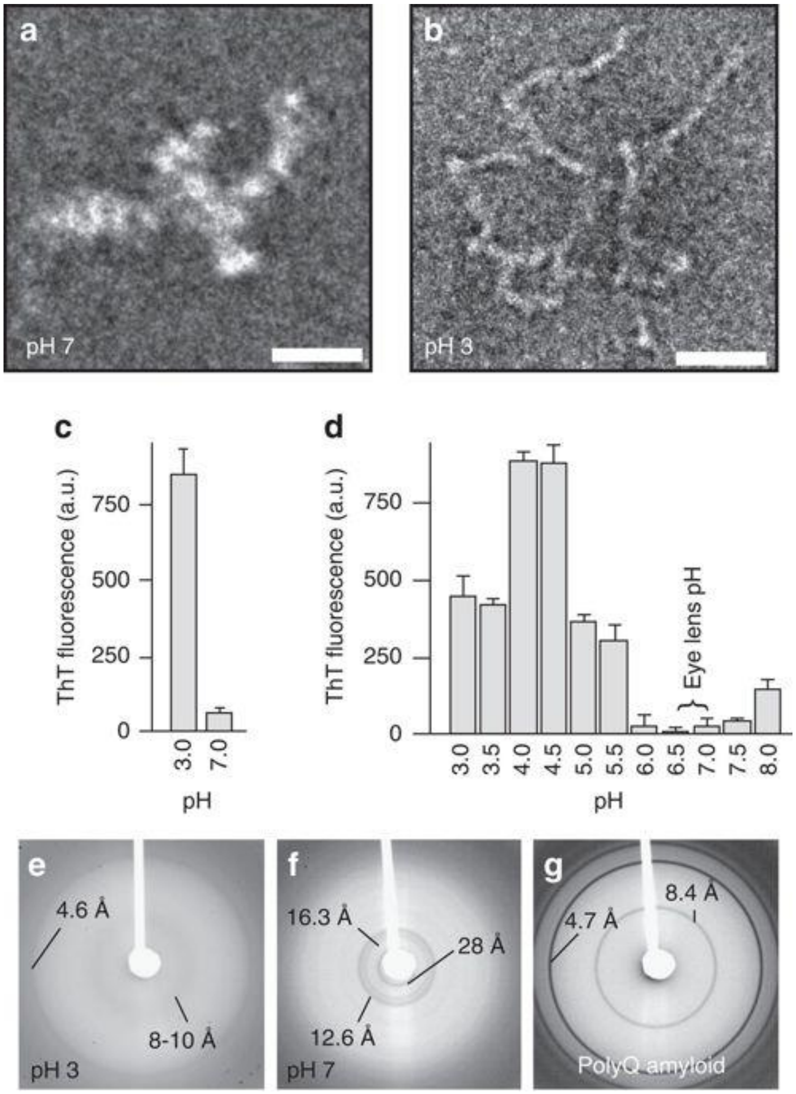


Figure 4.3. Characterization of aggregates formed at 37° C under different pH conditions

(a, b) Negative-stain TEM images of P23T hγD aggregates, formed at pH 7 and pH 3, show amorphous-looking aggregates and worm-like fibrils, respectively. (c) ThT fluorescence assay for both of these P23T hγD aggregates. (d) ThT fluorescence data for 0.13 μM P23T hγD incubated at several pH values. (e, f) X-ray powder diffraction of hydrated P23T hγD aggregates formed at pH 3 and pH 7, respectively. Prominent repeat distances are indicated. (g) X-ray powder diffraction on hydrated $K_2Q_{31}K_2$ *polyglutamine fibrils, showing the characteristic amyloid cross-β dimensions of 4.7 Å and 8.4 Å, as indicated [1].*

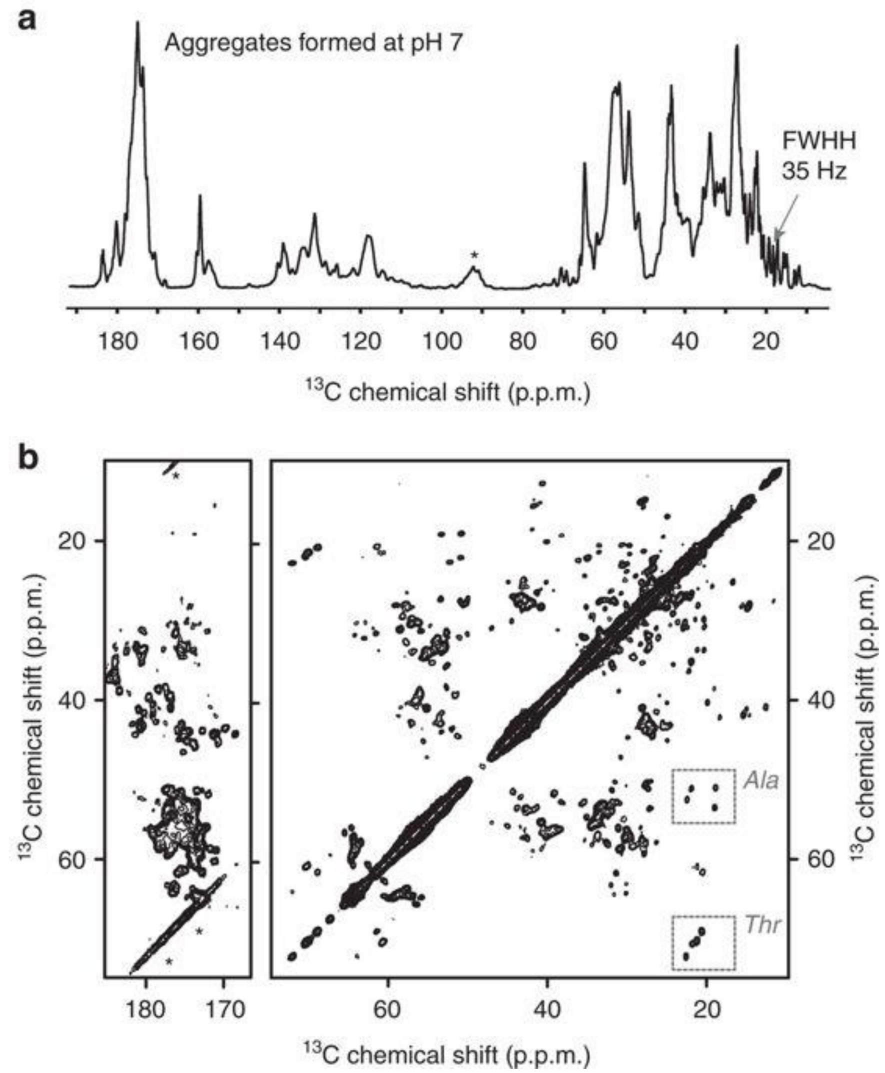


Figure 4.4. MAS ssNMR spectra of uniformly ^{13}C,^{15}N-labeled P23T hγD aggregates formed at pH 7

(a) 1D ^{13}C CP spectrum. The width at half height (FWHH) of one of the methyl peaks is indicated. (b) 2D ^{13}C-^{13}C ssNMR spectrum with an 8 ms DARR mixing time, resulting in one- and two-bond cross-peaks. Spectra were obtained at 600 MHz (^{1}H) and 12.5 kHz MAS. Spinning side-band peaks are marked with asterisks (*). Dashed boxes mark Ala Cα-Cβ and Thr Cβ-Cγ peaks.

4.3.3 Solid-state NMR of the acid-induced fibrils

Analogous ssNMR data were obtained for the acid-induced fibrils. The obtained spectra (Figure 4.5) are dramatically different, and surprisingly indicate a less ordered internal structure than was present in the amorphous-looking deposits. Figure 4.5, a, shows that the 1D spectrum has broader peaks and a limited resolution, which is indicative of a higher degree of structural heterogeneity in these fibrils. The 2D data (Figure 4.5, b) feature notably fewer and broader peaks compared to the aggregates that form at pH 7. This can be appreciated for instance by comparing the Ala peaks in the boxed regions of Figure 4.4, b and Figure 4.5, b P23T hγD has four Ala residues that give four well-resolved peaks in the pH 7 aggregates, but in the fibrils we see a much broader unresolved Ala peak. This observation suggests a lack of chemical shift differences among the observable residues, which points to them being in a similar structure and environment. That said, these CP-based experiments show only the immobilized parts of the protein aggregates; extensive dynamics of parts of the protein could also prevent specific residues from being detected.

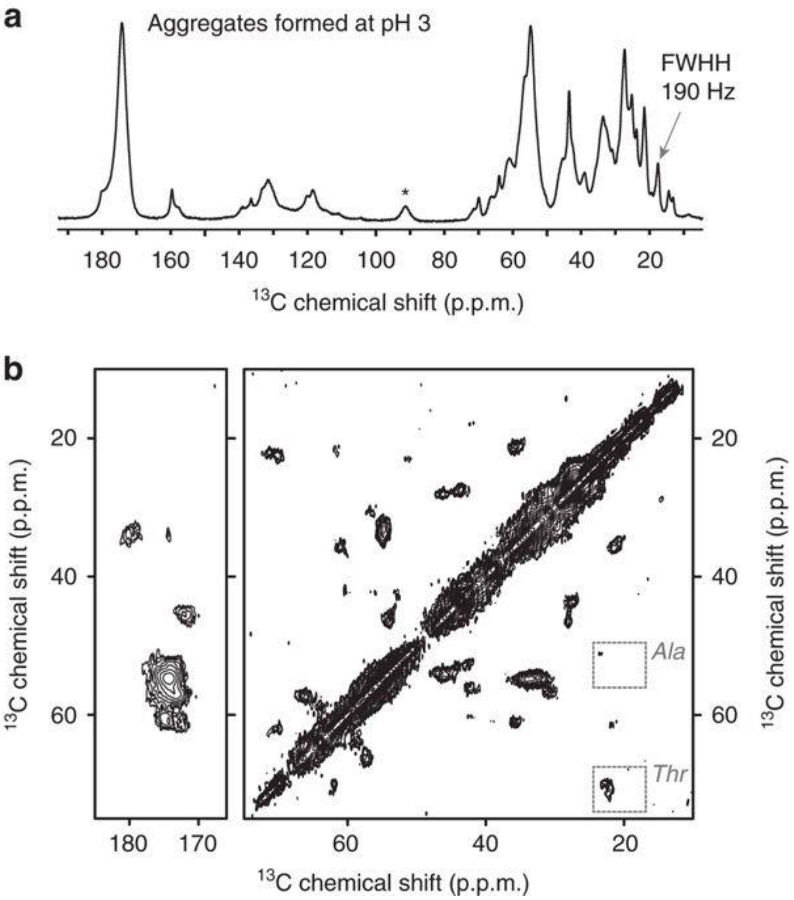


Figure 4.5. MAS ssNMR spectra of U-^{13}C,^{15}N-labeled P23T hγD aggregates formed at pH 3

(a) 1D ^{13}C CP spectrum. (b) 2D ^{13}C-^{13}C ssNMR spectrum with 8 ms DARR mixing. Both spectra were obtained at 600 MHz (^{1}H) and 12.5 kHz MAS. A spinning side-band peak is marked with an asterisk (*). Ala Cα-Cβ and Thr Cβ-Cγ cross peaks are enclosed in dashed boxes.

4.3.4 Detection of dynamic regions by ssNMR

To probe for such dynamic protein segments, MAS ssNMR experiments based on the INEPT (Insensitive nuclei enhanced by polarization transfer) scheme were also performed. In these INEPT-based ssNMR experiments, *only* the most dynamic residues are seen, while signals from rigid and immobilized parts of the protein are eliminated [265]. When the same ^{1}H-^{13}C INEPT experiment is applied to both types of aggregates, we see a striking difference (Figure 4.6, a). The aggregates formed at pH 3 feature a substantial number of mobile residues, as indicated by the presence of intense peaks in the INEPT spectrum. In striking contrast, the INEPT spectrum of the pH 7 aggregates is devoid of peaks (Figure 4.6, a; bottom). A 2D ^{13}C-^{13}C INEPT-based ssNMR spectrum for the acid-induced fibrils allows us to probe the identity of the highly mobile amino acids (Figure 4.6, b). No Ala peaks are detected, indicating a lack of highly mobile Ala residues, which implies that the broad Ala signal in the CP spectrum must contain the overlapping peaks from multiple residues in similar conformations. The spectra contain signals from both backbone and side chain atoms, including those of a Lys residue. Since it is the only Lys in the protein, those mobile Lys signals must stem from Lys-2 near the N-terminus of the NTD. Moreover, other visible peaks include Gly, Lys, Ile, Thr, and Leu residues, which happen to make up the initial five residues of the NTD (see Figure 4.1, a). Thus, these results are consistent with the initial segment of the NTD being unfolded and flexible, outside the mostly immobilized core of the amyloid fibril assemblies. Additional residues, such as Ser and Arg, are not found at the very N-terminus, indicating that other parts of the protein also end up in a similarly dynamic state.

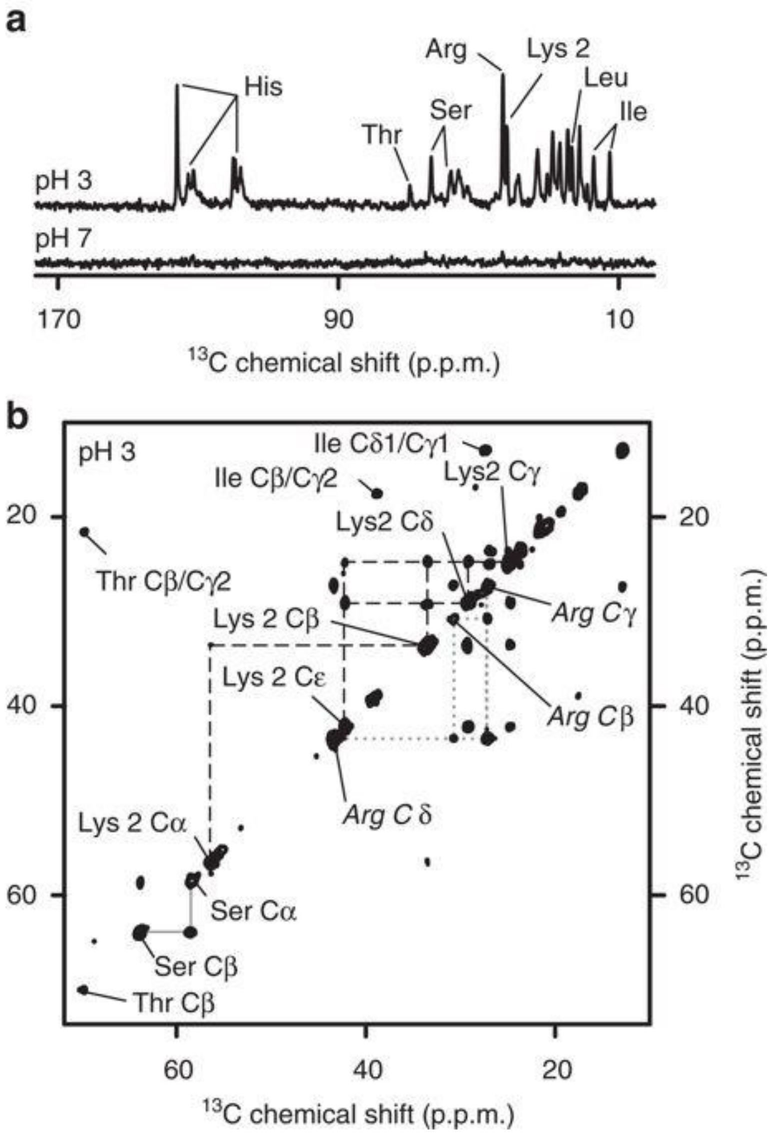


Figure 4.6. MAS ssNMR spectra of mobile residues within aggregated P23T hγD

(a) 1D ^{13}C INEPT spectra of U-^{13}C,^{15}N P23T hγD aggregates formed at pH 3 (top) and at pH 7 (bottom). (b) 2D ^{13}C-^{13}C INEPT/TOBSY spectrum of the U-^{13}C,^{15}N fibrils formed at pH 3. Lys2, Arg, and Ser peaks are connected with dashed, dotted, and continuous (grey) lines, respectively. Spectra were obtained at 600 MHz (^{1}H) and 8.33 kHz MAS.

4.3.5 Secondary structure content of the aggregates

Solid- and liquid-state NMR resonance frequencies are determined largely by the local bond geometry. The dependence of ^{13}Cα and ^{13}Cβ resonance frequencies on the local secondary structure is such that ^{13}Cα-^{13}Cβ peaks occur in distinct spectral regions for residues in random coil, α-helical and β-sheet structures [266]. In Figure 4.7, a-c we visualize this for some well-resolved amino acids by superimposing the canonical secondary-structure-dependent peak positions on our experimental ssNMR data. Note that these CP-based spectra detect specifically the immobilized parts of the protein aggregates. Independent of the aggregation condition, many of the experimental peaks line up with those typical of β-sheets. The dominance of β-sheet structure is most pronounced in the pH 3 fibrils (Figure 4.7, b-d). The fibrils lack peaks in the α-helical region, but signals are detected in the intervening spectral region where one expects the signals for residues that lack a well-defined α-helical or β-sheet structure. Interestingly, both Ser (Figure 4.7, a) and Ala peaks (Figure 4.7, c-d) with α-helical chemical shifts are observed in the spectrum of the pH 7 aggregates. Thus, compared to the fibrils, these amorphous-looking aggregates have a more diverse secondary structure content that includes both β-sheet and α-helical elements.

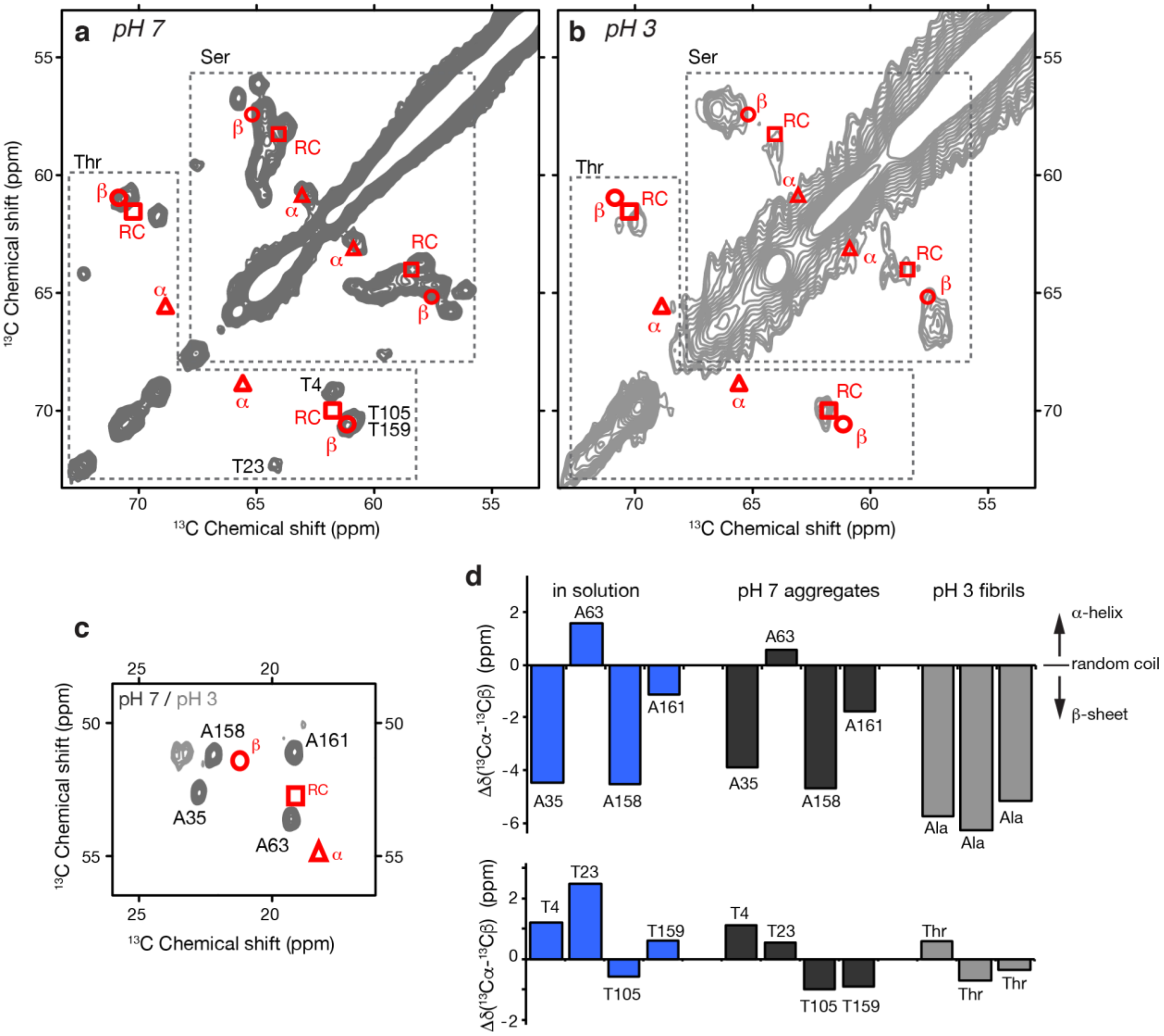


Figure 4.7. NMR chemical shift analysis of residue-specific secondary structure

(a) Thr and Ser Cα/Cβ region of the 2D ^{13}C-^{13}C *ssNMR spectrum for P23T hγD aggregated at neutral pH. The expected peak positions [266] for Thr and Ser in random coil (RC), α-helix (α) and β-sheet (β) structural motifs are indicated with red markers superimposed on the experimental spectrum (grey). Indicated peak assignments are based on 2D and 3D ssNMR spectra. (b) Analogous data overlay for the acid-induced P23T hγD fibrils. (c) Overlay of the Ala Cα/Cβ cross-peaks for the aggregates formed at pH 7 (dark grey) and pH 3 (light grey), along with the expected peak positions for Ala in different secondary structures. Indicated peak assignments are based on 2D and 3D ssNMR spectra. (d) Secondary chemical shifts* ($\Delta\delta(^{13}C\alpha-^{13}C\beta)$) *calculated for the Ala (top) and Thr (bottom) residues of P23T hγD in solution (left), aggregated at pH 7 (middle), or at pH 3 (right). These secondary shifts,*

$\Delta\delta$(^{13}Cα-^{13}Cβ), are normalized relative to the shift of a random coil conformation, such that negative values are indicative of β-sheet structure, whereas positive bars indicate α-helical structure (as indicated on far right).

4.3.6 Spectral modeling of misfolded and native-like aggregates

As noted above, crystallin aggregation may proceed via mechanisms characterized by different extents of misfolding. Proposed aggregation mechanisms involving extensive un- and misfolding are expected to yield amyloid-like fibrils that feature a β-sheet-based core stabilized by intermolecular hydrogen bonds [85, 239, 267]. Runaway domain-swapping [65-67] or surface-mediated ‘condensation’ mechanisms [69-71] would generate aggregates that retain much of the native fold and lack an amyloid core [258, 268]. To evaluate and visualize the expected spectral differences between misfolded amyloid and native-like aggregates, we generated schematic simulated ^{13}C-^{13}C ssNMR spectra. To coarsely model a typical amyloid structure, we applied canonical parallel β-sheet torsion angles (ϕ,φ = -119°, 113°) to the P23T hγD primary sequence. This backbone structure model was then used to predict the ^{13}C’, ^{13}Cα, and ^{13}Cβ chemical shifts, which were used to generate the ^{13}C-^{13}C peak patterns in Figure 4.8, a. To simulate the ssNMR spectrum of native-like aggregates, the NMR shifts of the soluble protein were used as a starting point [24]. These solution NMR shifts by definition correspond to the soluble protein’s native fold, and led to the simulated spectrum in Figure 4.8, b. The two simulated spectra illustrate the marked qualitative differences expected for amyloid-like fibrils and native-like aggregates. The former is characterized by a lack of peak dispersion for residues in the purely β-sheet amyloid core. With their mixed secondary structure, the native-like aggregates have a much larger peak dispersion that resolves many of the individual amino acids.

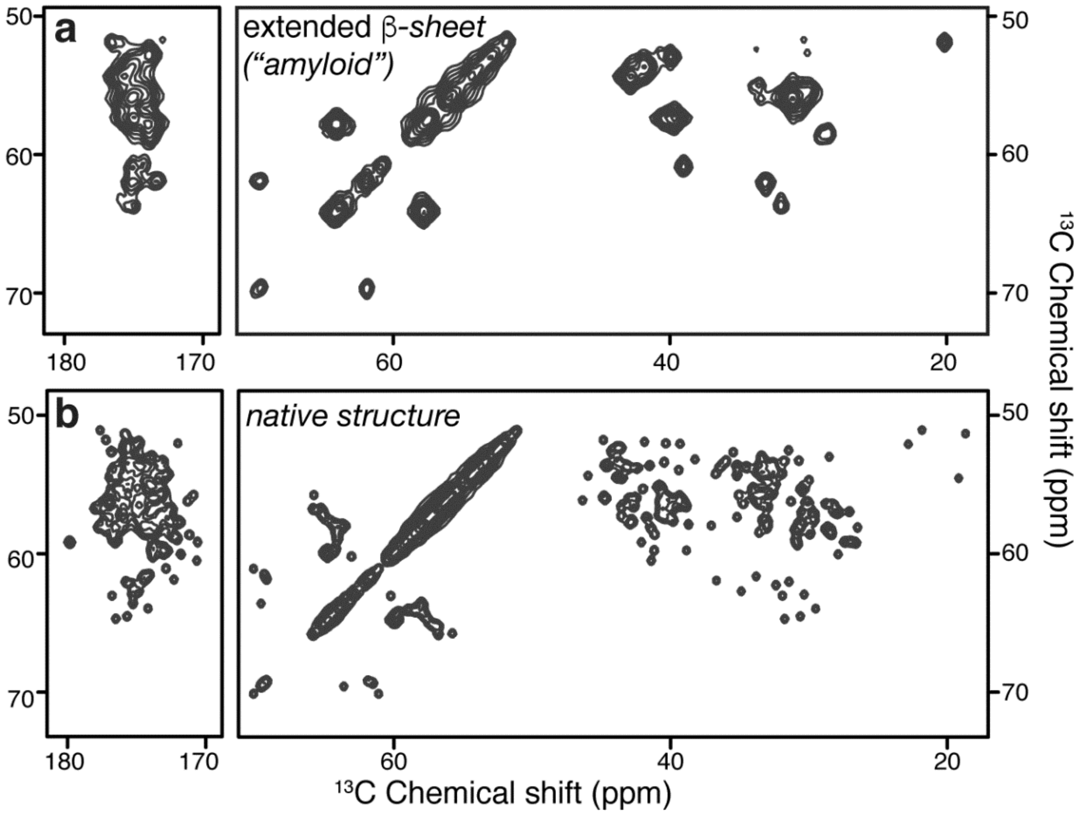


Figure 4.8. Synthetic ^{13}C-^{13}C ssNMR spectra predicted for aggregates containing either typical amyloid-like misfolded proteins or natively folded P23T hγD

(a) Simulated spectrum predicted for P23T hγD in a parallel in-register β-sheet structure, as is typical for amyloid-like fibrils. (b) Simulated spectrum generated from the solution NMR chemical shifts of natively folded P23T hγD [24].

4.3.7 Acid-induced P23T hγD fibrils lack native-like structure

The experimental ssNMR data for the acid-induced fibrils (Figure 4.5) match well to the amyloid-like fibril model, while showing little to no correspondence to the spectra expected for native-like aggregates. As noted above, the experimental spectrum features some peaks from non-β residues that do not appear in the synthetic spectrum. We attribute these peaks to loops and other segments that are outside the β-sheet amyloid core, but are immobilized enough to be seen in the CP-based spectra, similar to the non-amyloid flanking domains of huntingtin exon 1 fibrils [91]. Such non-core residues were not represented in our simplistic amyloid model designed to predict the spectral features of a canonical amyloid core architecture.

4.3.8 The amorphous aggregates have a native-like structure

In contrast to the acid-induced fibrils, the spectrum of amorphous-looking aggregates obtained at pH 7 (Figure 4.4) shows a striking resemblance to the simulated spectrum expected for native-like aggregates. To provide more detailed insights into the structural similarities and differences between the aggregated and soluble state, Figure 4.9 overlays the 2D ^{13}C-^{13}C ssNMR spectrum of P23T hγD aggregated at pH 7 with peak markers representing the solution NMR assignments. Even in absence of site-specific assignments (see below), a close peak-by-peak inspection of this single 2D spectrum reveals that the majority of peak positions are remarkably similar in the solid and solution states. A limited number of well-resolved solution peak markers do not overlap with a matching ssNMR peak, as indicated with circled crosses in Figure 4.9.

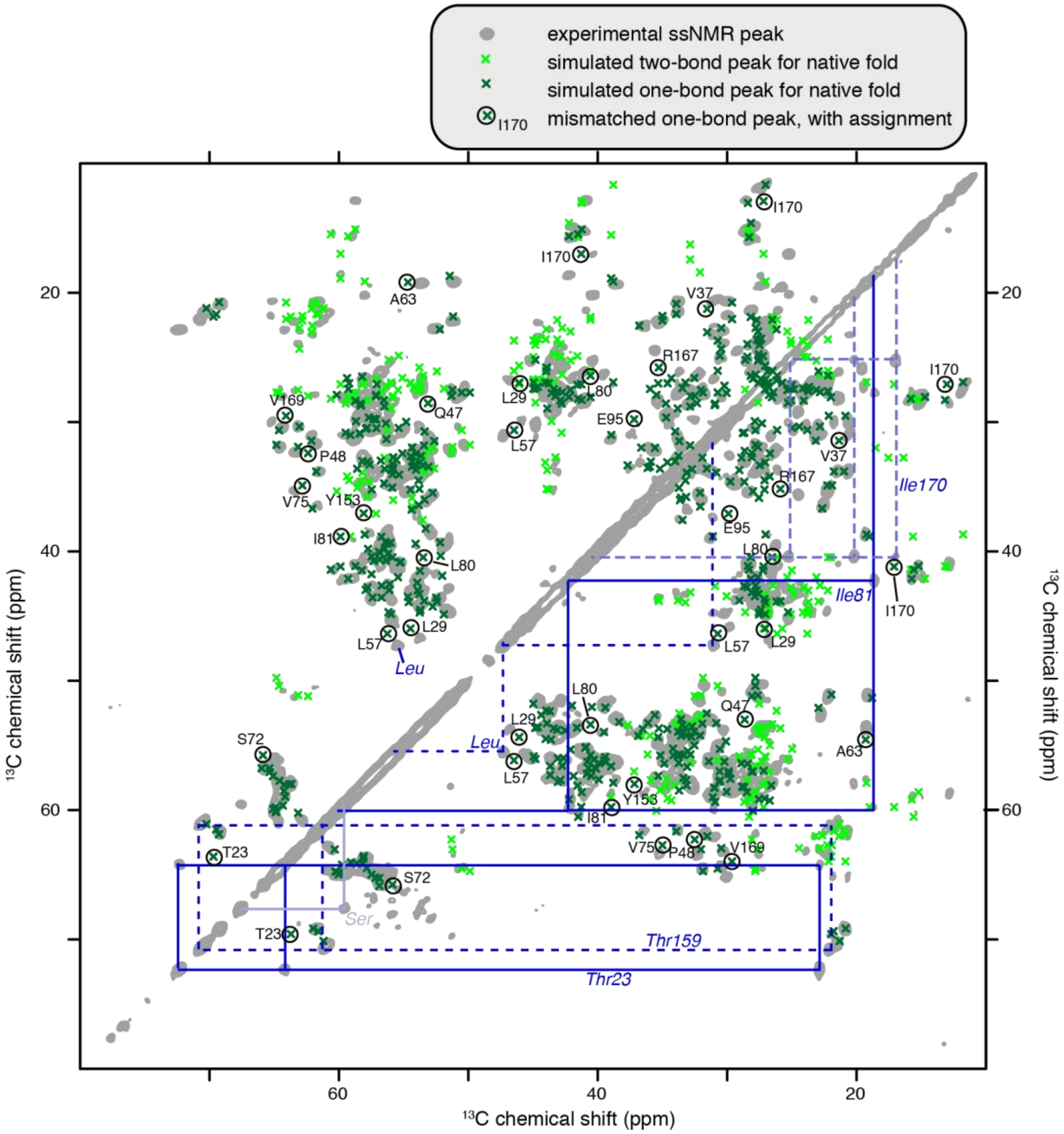


Figure 4.9. Experimental ^{13}C-^{13}C ssNMR spectrum of amorphous P23T hγD aggregates annotated with the liquid-state NMR resonances of P23T hγD in solution

The grey spectrum is a 2D ^{13}C-^{13}C MAS ssNMR spectrum with 8 ms of ^{13}C-^{13}C DARR mixing. The off-diagonal peaks are mostly one-bond ^{13}C-^{13}C contacts, along with a few two-bond cross-peaks. Dark and light 'x' markers indicate one- and two-bond peak positions, respectively, as predicted from the solution NMR assignments of soluble P23T hγD [24]. Circled crosses with their single-letter residue assignments mark well-resolved one-bond solution NMR peaks that do not overlap with peaks in the ssNMR spectrum. Conversely, colored lines connect sets of ssNMR *peaks that do not have overlapping solution NMR peak markers. Different coloring and line styles (dashed/solid) connect together individual residues ("spin systems"). Where available (see Table B.2), color-coded three letter residue assignments are provided; otherwise the residue type is indicated (also with its three-letter code).*

Conversely, a subset of ssNMR peaks do not have a matching solution NMR resonance (marked with colored lines in Figure 4.9). Assignment experiments (below) show that the 'missing' solution NMR peaks and 'unmatched' ssNMR peaks represent the same specific residues. Thus, we are not seeing a disappearance of peaks, but rather a changing of NMR resonance frequencies. This is consistent with the fact that no highly flexible residues (which would be invisible in the CP spectra) were detected in the pH-7 aggregates (Figure 4.6).

Residue-specific assignments of several key segments of the aggregated protein were obtained based on a backbone-walk analysis of 2D and 3D homo- and heteronuclear ssNMR experiments (see Figure B.1, Figure B.2). Comparing the available solid- and solution-state chemical shifts, Figure B.1 shows the residue-specific chemical shift perturbation (CSP) that accompanies the formation of the amorphous aggregates. These data support the visual analysis of the 2D ^{13}C- ^{13}C data above, showing that many residues have a CSP similar to the uncertainty in the chemical shifts (the average ^{13}C CSP is 0.55 ppm). The largest ^{13}C chemical shift deviations are ~2 ppm, but this represents only a few residues among those that were assigned: G1, T23, I81 and I170. The chemical shift differences between the pH 7 aggregates and the hypothetical amyloid-like β-sheet model from Figure 4.8, a, are more than twice as large (Figure B.2).

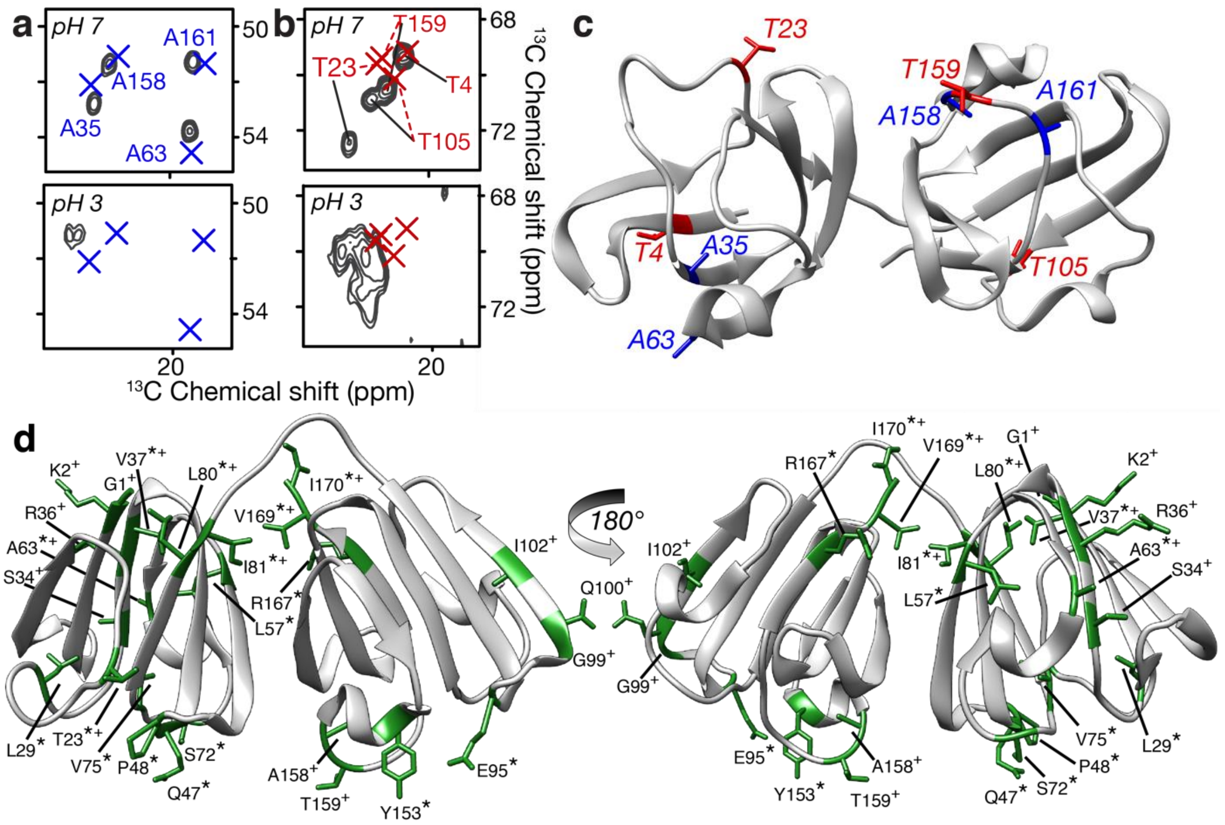


Figure 4.10. Mapping of NMR-detected structural changes and similarities

(a) Ala $C\alpha$-$C\beta$ cross-peaks for the aggregates obtained at pH 7 (top) and pH 3 (bottom). Experimental ssNMR cross peaks are depicted in dark grey and blue crosses mark the solution NMR shifts of the soluble protein [24]. (b) Analogous depiction of the Thr $C\beta$-$C\gamma$ cross peaks. (c) P23T hγD structure [130] highlighting the locations of all eight Ala and Thr residues. (d) Structural mapping of residues whose NMR resonances change upon aggregation at pH 7 (see Figure 4.9 and Figure B.2) due to aggregation-associated changes in the local structure or environment. Residues whose resonances are different between the solution and aggregated (pH 7) states are labeled in green on the ribbon diagram. Residues marked with "+" have unambiguous ssNMR assignments (Table B.2); those marked with "" were identified in the visual analysis of Figure 4.9.*

This overall pattern is visualized in Figure 4.10, a-b, based on the well-resolved Ala and Thr signals. Unlike the acid-induced fibrils (bottom panels), the amorphous-looking aggregates have Ala peaks that closely resemble the solution NMR shifts (marked with X). Given the Ala distribution throughout the protein (Figure 4.10, c), this similarity in itself strongly suggests a retention of much of the native fold. Among the four Thr residues, two (T23 and T159) have significantly changed chemical shifts indicating a localized change in structure. Returning to a more global view of the protein aggregates formed at pH 7, Figure 4.10, d highlights those residues for which we observe significant (> 0.5 ppm) ^{13}C shift changes between the solution and solid-state NMR data. The indicated residues are identified either based on *de novo* ssNMR assignments (Table B.2; marked with + in Figure 4.10, d), the spectral analysis of Figure 4.9 (*), or both. This representation indicates that residues that sense a structural change upon aggregation are mostly outside the core β-strand segments of the Greek key motifs and are found instead on the surface-exposed regions of the protein's folded subdomains.

4.4 DISCUSSION

We observed the aggregation of P23T hγD across a range of pH conditions and found that the nature of the resulting aggregates is different at different pH values. The aggregates formed at non-physiological pH values showed high fluorescence in assays using the amyloid-binding ThT dye. The amyloid-like nature of the acid-induced aggregates was validated by their fibrillar morphology and characteristic cross-β X-ray diffraction pattern (Figure 4.3, e). As is typical for amyloid-like fibrils, the ssNMR peak patterns of the fibrils differed dramatically from those expected for the protein's native fold. The protein had undergone extensive restructuring during aggregation at pH 3. The ssNMR data we have obtained thus far are unable to distinguish the different kinds of β-sheet amyloid core architectures [267]. That said, as previously suggested [85], it seems likely that the fibril core would feature the prototypical parallel in-register β-sheet structure most commonly seen in other amyloid-like fibrils [239]. The ssNMR spectra contained signals from residues lacking β-sheet structure that are either dynamically (Figure 4.6, b) or statically disordered (Figure 4.7). The amino acids are presumably located outside the β-sheet-based amyloid core. Thus, our data support a model featuring fibrils built around a β-sheet amyloid core that is decorated with disordered non-β segments. This structural motif qualitatively resembles the fibril assembly model described in prior IR studies of hγD aggregated under acidic conditions *in vitro* [85].

Consistent with prior studies [242, 243, 269], we found that denaturing conditions were not necessary for P23T hγD to aggregate at concentrations well below the high solubility of WT hγD. The P23T hγD aggregates formed at physiological pH values lack the characteristic cross-β signature of amyloids, have low ThT fluorescence, and have an amorphous appearance by TEM. A similar amorphous morphology was previously noted for P23T hγD aggregated under similar conditions [243], and also for other β- and γ-crystallins [250-252, 270]. As noted in the introduction, the same nomenclature is used to describe non-fibrillar aggregates formed by other aggregation-prone proteins [244-247]. Prior studies offer little insight into the internal structure of these deposits, which are often described as "unstructured" [243, 248, 249] or "disordered" [116, 249, 251, 271]. Our ssNMR studies paint quite a different picture. Both the dispersion and widths of our ssNMR peaks are reminiscent of those typically seen for crystalline protein preparations [272]. This means that the individual protein molecules in the sample all share a single, well-defined conformation. In CP-based spectra of the immobilized parts of the aggregates we observed residue counts that match the solution NMR data, with no sign of missing peaks. This was reaffirmed by INEPT spectra that were devoid of 'mobile' signals (in contrast to the acid-induced fibrils). Thus, the seemingly amorphous deposits are actually highly homogeneous assemblies of well-structured protein monomers on the microscopic level, with no evidence of either dynamic or static disorder. Figure 4.9 shows the striking similarity between the 2D ssNMR spectra of the pH 7 aggregates and the known chemical shifts of the natively folded state present in solution. Residues featuring atoms with a significant change in their chemical shifts were found to be outside the β-strand segments of the native fold's Greek key motifs (Figure 4.10, d; Figure B.2). These sites of structural change likely co-localize at least in part with the intermolecular contacts that arise upon aggregation, but are absent in the monomeric soluble protein. The limited spectral changes detected in the aggregates' ssNMR data are incompatible with large-scale un- or misfolding.

The fact that the ssNMR spectra of the pH 7 aggregates indicate a native-like structure is most consistent with two types of aggregation mechanisms proposed in the literature. 'Condensation' mechanisms suggest that minor changes in surface charges or surface hydrophobicity would be sufficient to increase the molecule's aggregation propensity, without substantial mis- or re-folding (Figure 4.11, b) [69-71, 273]. Even relatively small changes in surface characteristics may disrupt the careful balance of inter-protein interactions that normally facilitate the remarkable protein concentrations in the eye lens. The proteins could then aggregate while preserving a largely native fold, with the mutations affecting the surface patches that mediate the protein-protein contacts in the aggregated state. Figure 4.10, d summarizes the identities and locations of residues for which we observe above-average chemical shift deviations between the soluble and aggregated protein (ref. Figure 4.9 and Figure B.2). These are the residues that experience a change in local structure as a result of the aggregation process. They are found to reside outside the core β-strands of the NTD and CTD, and are instead primarily on the surfaces of the domains, indicating an apparent preservation of the Greek key motifs.

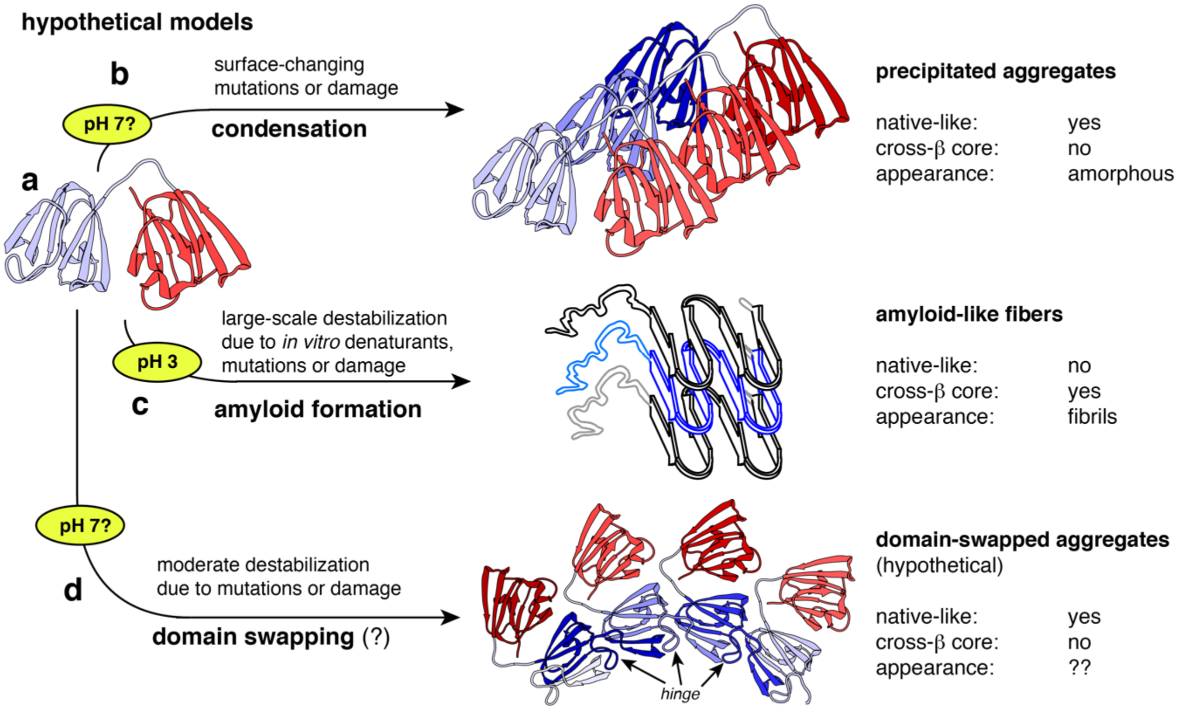


Figure 4.11. Proposed potential aggregation mechanisms

(a) The native hγD fold, with the NTD and CTD in blue and red, respectively. (b) Condensation into aggregates within which individual proteins retain a native-like fold. Mutations or chemical damage may lead to changes in surface charges or hydrophobicity, thus disrupting the native protein-protein interactions essential for high solubility. Individual protein monomers are depicted in different color shades. (c) Denaturing conditions, including low pH, induce a loss of the native fold and formation of fibrils with a β-sheet-based amyloid core. Unfolding may also result from disruptive mutations or chemical damage to buried residues. (d) Hypothesized domain-swapping-based mechanism that generates aggregates in which much of the native fold is preserved, except for loops that act as hinges and undergo substantial rearrangements.

Deposits assembled from intertwined or 3D domain-swapped proteins, as previously proposed for cataract-associated aggregates [65-67], would also be predicted to have native-like spectra

[258]. On the basis of known domain swapped multimers as large as tetramers, it has been proposed that "runaway" domain swapping in a daisy-chain fashion could be a mechanism of protein aggregation. A hypothetical domain-swapped P23T hγD assembly is shown in Figure 4.11, d. We stress that this is a hypothetical graphic, purely meant to offer a schematic visual representation of the structural implications of run-away domain swapping. The swapped multimers reproduce much of the native fold, with structural changes centered on loop regions that act as hinges between the swapped domains. A characteristic feature of domain swapped assemblies would be that notable NMR chemical shift changes would primarily cluster in such a hinge loop region. The observed ssNMR shift differences (Figure 4.10, d) fail to match this particular pattern, leaving us unable to identify a specific hinge loop region, or build a particularly convincing runaway domain-swapped model. As a result, we interpret the available data to be most consistent with the kind of aggregates expected for a condensation-type aggregation process. That said, there are many different ways in which a domain-swapped assembly could be created, such that it is difficult to exclude a role for domain swapping until more structural data are available.

Various γ-crystallins form amyloid-like fibrils *in vitro*, at acidic pH [71, 85, 252, 274] or when exposed to other denaturing conditions [65, 71, 250, 253, 254, 275-277]. A similar behavior has been reported for α-and β-crystallins [71, 255, 277]. However, numerous globular proteins can be induced to form amyloid-like fibrils when one destabilizes their native fold, independent of their propensity to form amyloid under physiological conditions [236, 258]. In the case of P23T hγD, the mutation itself does not cause much change in the structure, as mentioned in the Introduction and discussed in detail in prior studies [24, 241]. Additional destabilization of the native fold, for instance by exposure to acidic or basic pH values, appears to be necessary for amyloid formation to become the dominant aggregation pathway (Figure 4.11, c). Thus, it is tempting to conclude that P23T hγD aggregation under *in vivo* conditions is likely to occur via a non-amyloid mechanism. Of course, extrapolation to the complex *in vivo* conditions is always difficult, given the potential impacts of other β/γ-crystallins, α-crystallin chaperones, and the build-up of aging-induced chemical modifications [71, 251]. That said, we do want to point out one way in which cataract-related protein aggregation would appear to be fundamentally different from amyloid formation in neurodegenerative disorders. In the latter case, the amyloid-forming proteins or peptides are often present at very low concentrations [278, 279]. The concentration of protein in the eye lens is orders of magnitude higher, requiring the remarkable eye lens protein solubility that is based on a careful balancing of repulsive and attractive protein-protein interactions [280]. Amyloidogenic aggregation at low concentrations tends to involve exposure of hydrophobic sites through misfolding, and proceed via specific nucleation events [1, 281]. At very high protein concentrations, aggregation can result from surface-mediated changes in protein-protein interactions, well before large-scale structural changes trigger the formation of an amyloidogenic state.

Protein aggregation processes involved in cataract formation may be qualitatively different from the misfolding-based amyloid formation associated with neurodegenerative disease. Recent efforts to develop and screen anti-cataract drugs have focused on their inhibition of amyloid formation [276, 277]. Our results argue that it may be important to also screen against disease-related crystallins that form non-amyloid aggregates. We showed that the amorphous-looking aggregates are much more ordered and structured that one might have expected based on their morphology. As such, at least in this case, the “amorphous” label appears to be a misnomer when it comes to the atomic level structure of the aggregated protein. One implication of this remarkable structural order and homogeneity is that the deposits presumably form via a reproducible and specific aggregation mechanism, rather than a multiplicity of parallel or non-specific pathways. A better molecular understanding of this aggregation process may pave the way for the development and design of small-molecule aggregation inhibitors [282, 283]. Age-dependent changes in the operation of nature's own aggregation inhibitors, protein chaperones, play a key role in aging-dependent aggregation diseases. In the eye, the α-crystallins perform this protective role. Intriguingly, α-crystallins act on amorphous-looking and amyloid-like substrates through different mechanisms [118, 237]. Further structural studies of the crystallins' polymorphic aggregates will enhance our understanding of these protective processes and may enhance ongoing efforts to develop new preventative or curative treatments.

4.5 ACKNOWLEDGMENTS

We thank Drs. Cody Hoop, Jinwon Jung and Fangling Ji for helpful discussions and for performing early experiments related to the current work, and Dr. Abhishek Mandal for help with experiments and helpful discussions. We acknowledge Michael Delk, Doowon Lee, and Dr. Alexander Makhov for assistance with NMR, X-ray, and EM instrumentation. We thank Dr. James Conway for use of the departmental EM facility. This work was enabled by funding from the University of Pittsburgh, the National Institutes of Health (R01 EY021193 to A.M.G., R01 GM112678 to P.v.d.W., T32 GM088119 to J.C.B.), the Achievement Rewards for College Scientists (ARCS) Foundation (J.C.B.), and grant UL1 RR024153 from the National Center for Research Resources (NCRR). Molecular graphics were prepared with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

4.6 AUTHOR CONTRIBUTIONS

J.C.B. and P.v.d.W. conducted the ssNMR experiments. J.C.B, M.L., and P.v.d.W. analyzed ssNMR data. J.C.B. and M.J.W. prepared samples. J.C.B. performed the TEM, X-ray, ThT, and aggregation experiments. J.C.B., P.v.d.W., M.J.W., and A.M.G. designed the experiments; J.C.B. and P.v.d.W. wrote the paper.

5.0 MAGIC ANGLE SPINNING SOLID STATE NMR STUDIES OF METHIONINE OXIDIZED APOLIPOPROTEIN A-I AGGREGATES

Adapted and reprinted with permission from: Witkowski, A., Chan, G. K. L, **Boatz, J. C.**, Li, N. J., Wong, J. C., van der Wel, P. C. A., Cavigiolio, G. *Methionine oxidized apolipoprotein A-I at the crossroads of HDL biogenesis and amyloid formation*, The FASEB Journal 32(6):3149-3165. Copyright (2018) Witkowski, A., et. al.

I contributed data towards one figure in the main text of the published manuscript, and three figures in the supplemental information in the published manuscript, reprinted with permission as Figure 5.2, Figure 5.5, Figure 5.6, and Figure C.1. I built the cartoon of the amyloid structure reprinted with permission in Figure 5.7. I obtained images from samples using transmission electron microscopy, depicted in Figure 5.2 and Figure 5.3. I also performed several magic angle spinning solid state NMR experiments, depicted in Figure 5.5, Figure 5.6, Figure C.1, and Figure C.2.

Atherosclerosis is a state of chronic inflammation in the arteries, initiated by deposition of oxidized lipids, and is one of the primary causes of cardiovascular disease. Incidence of amyloid deposits in atherosclerotic lesions is very high; however, a causal relationship between arterial amyloid deposition and atherosclerosis development has not yet been unequivocally established. Surprisingly, the known anti-atherogenic factor apolipoprotein A-I (apoA-I) has been identified as the protein contributing to over 90% of amyloid deposits found in atherosclerotic plaques. Previously, it was demonstrated that methionine oxidation promotes apoA-I aggregation. In my doctoral work, I applied magic angle spinning solid state NMR to probe the structure of aggregated apoA-I. This technique is exquisitely sensitive to structural changes and readily allows direct comparison of secondary structure changes between samples. We find that methionine oxidized apoA-I forms amyloid-like fibrils that can propagate through seeding. Notably, a large portion of the aggregated protein retains a native-like structure. These results suggest that methionine oxidation can promote vascular amyloidosis, and that this phenomenon impacts the anti-atherosclerotic properties of apoA-I.

5.1 INTRODUCTION

Cardiovascular disease accounts for 1 in 4 deaths in the United States [284]. One of the leading causes of cardiovascular diseases is atherosclerosis, which is defined as plaque buildup within arteries [285]. More specifically, oxidized lipids infiltrate the sub-endothelial space of arteries and initiate an inflammatory response. Within this inflammatory and oxidative milieu, foam cells and fatty deposits accumulate, forming atherosclerotic plaques [285]. Vascular amyloidosis (VA) and atherosclerosis are highly correlated, as high incidence of amyloid fibrils occurs in atherosclerotic plaques [89]. However, it is uncertain whether there is a causal relationship between VA and atherosclerosis and the role of protein aggregation in atherogenesis remains unclear.

High density lipoproteins (HDL) are known anti-atherosclerotic factors [286]. The major protein component of HDL particles is apolipoprotein A-I (apoA-I). ApoA-I is a large (45 kDa) protein with high structural plasticity that readily binds to phospholipids and cholesterol and is important for lipid metabolism [287]. By extracting cholesterol from immune cells strongly implicated in the process of atherogenesis, such as macrophages, apoA-I reduces foam cell formation and atherosclerosis progression [287]. By binding cell membrane cholesterol and forming HDL particles, apoA-I facilitates transport of this excess of cholesterol out of the sub-endothelial space of the arteries and through the bloodstream, delivery to the liver for excretion, a process often called “reverse cholesterol transport” [287].

The cholesterol release function of apoA-I is hindered in atherosclerosis [288]. The macrophage secreted enzyme myeloperoxidase (MPO), strongly implicated in atherosclerosis and found at high concentrations in the atherosclerotic plaques, has been shown to inhibit apoA-I function [289-291]. Under physiological conditions, MPO oxidizes methionine residues within apoA-I and targets tryptophan, tyrosine, and lysine residues for additional modifications including chlorination, nitration, and hydroxylation [14, 15, 292-299]. It has been unclear which MPO induced modification is responsible for the loss of function of apoA-I, but the cholesterol release function of apoA-I is not altered when only methionine (Met) residues are oxidized [6, 293, 295, 300, 301]. Conversely, when the MPO reaction is performed in the presence of 10-fold molar excess of H_2O_2, a condition that promotes Trp and Tyr modifications in addition to full Met oxidation, cholesterol release by oxidized apoA-I is significantly hindered [6]. Together, this indicates a possible causal relationship between oxidization at Met and amyloid formation.

Although the cholesterol binding function of apoA-I is not altered by oxidation of methionines, recent results from the Cavigiolio lab indicate that Met oxidation by MPO is sufficient to promote amyloid formation in oxidized lipid-free apoA-I; amyloid formation is abolished when the three Met residues of apoA-I are substituted with Leu [302]. Interestingly, amyloid formation can also be induced by *in vitro* oxidation with up to 1000-fold molar excess of H_2O_2 (H_2O_2-wt-ApoA-I) [6], and even in such large molar excess, H_2O_2 exclusively oxidizes Met residues [6]. Consistent with the lack of detrimental consequences of the oxidation state of methionine residues on the cholesterol binding function of apoA-I, interaction with macrophages and cholesterol release not only is not hindered, but it is enhanced in H_2O_2-wt-ApoA-I [6, 303]. Intriguingly, the same phenomenon is observed after thermal treatment at 90 °C for 1h [6]. Both thermal treatment and treatment with H_2O_2 disrupt self-association of apoA-I, the latter irreversibly [6].

Thus, oxidation of Met residues promotes structural changes that disrupt self-association and facilitate lipid binding, while also facilitating protein aggregation in amyloid-like assemblies [6, 303]. As the formation of amyloid fibrils processed by a monomer-dependent mechanism, and lipid free apoA-I amyloid fibrils can be propagated through seeding [6], a more thorough investigation of the structural changes that occur in apoA-I upon Met oxidation is warranted.

ApoA-I primarily adopts an α-helical structure that self-associates in the lipid-free state, and this self-association affects protein function [304]. Fluorescence studies indicate that Met oxidation and thermal treatment induces aromatic residues to become more solvent exposed, likely by disrupting an N-terminal 4-helix bundle [6, 305-309]. In this chapter, the structure and dynamics of H_2O_2-wt-ApoA-I were probed by MAS ssNMR in hopes of gaining atomic-level insight to answer this question.

5.2 METHODS

5.2.1 Preparation of apoA-I samples

I thank G. K. L. Chan, who expressed, purified, and prepared all apoA-I samples. ApoA-I (wt-^{13}C,^{15}N-ApoA-I) was expressed in BL21(DE3)pLysS *Escherichia coli* (Agilent Technologies) containing a pET-20b bacterial expression vector (Novagen) and encoding for an N-terminal His tag [302, 304]. Uniformly ^{13}C, ^{15}N labeled protein was overexpressed in minimal media using U-^{13}C-D-glucose and ^{15}N-ammonium chloride as the only carbon and nitrogen sources [210, 310]. The protein was purified using nickel affinity chromatography and the His tag was then removed with TEV protease [302, 311]. Protein purity was confirmed by SDS-PAGE. Samples were frozen were stored in 6M guanidine chloride buffer and then refolded before use. Samples were oxidized overnight at 37 °C in oxidation buffer (OB, 10 mM sodium phosphate, 100 μM DTPA, 100 mM NaCl, pH 7.5) with a 1000:1 molar excess of H_2O_2 (H_2O_2-wt-^{13}C,^{15}N-ApoA-I), and then were dialyzed into fibrillation buffer (10 mM sodium phosphate, pH 6.0) [293, 302]. Samples were then diluted to 1.0 mg/mL and aggregated at 37 °C with continuous vortexing [302]. Aggregation of non-oxidized wt-^{13}C,^{15}N-ApoA-I was initiated by seeding with unlabeled pre-formed H_2O_2-wt-ApoA-I seeds.

5.2.2 Thioflavin T fluorescence assays

I thank G. K. L. Chan and G. Cavigiolio for performing the Thioflavin T fluorescence studies. ThT fluorescence was recorded during aggregation in samples with and without oxidized apoA-I seeds at the timepoints indicated in Figure 5.1. Samples were diluted into ThT stock solution (500 μM ThT, 50 mM sodium phosphate, pH 7.4) for a final concentration of 2.84 μM apoA-I. Fluorescence spectra were recorded on a Horiba Jobin-Yvon FluoroMax-4 spectrofluorometer. Samples were excited at 450 nm and the emission was recorded between 460 and 500 nm. Slit widths were 2.5 and 5.0 nm, respectively.

5.2.3 Fourier transform infrared (FTIR) spectroscopy

I thank G. K. L. Chan and G. Cavigiolio for obtaining the FTIR measurements. FTIR absorbance spectra were recorded between 4000 and 800 cm^{-1} on a Direct Detect™ spectrometer (EMD Millipore Co.) [302].

5.2.4 Transmission electron microscopy

Samples were adsorbed onto freshly glow discharged carbon-coated copper grids for 1 minute and stained with 1% (w/v) uranyl acetate for 30 s. Images were taken on a Tecnai T12 transmission electron microscope (FEI; Hillsboro, OR) operating at 120 kV and equipped with an UltraScan 1000 CCD camera (Gatan; Pleasanton, CA) [3, 5].

5.2.5 MAS solid state NMR spectroscopy

Hydrated samples were packed into thin-wall 3.2 mm zirconia MAS rotors (Bruker Biospin, Billerica, MA) at 100,500 g for 60 min using a home-built swinging-bucket ultracentrifugal packing device, and then rotors were sealed with epoxy [4]. All MAS ssNMR experiments were performed on a widebore Bruker Avance I NMR spectrometer operating at a ^{1}H Larmor frequency of 600 MHz (14.1 T) and equipped with a 3.2 mm MAS ssNMR probe with a triple channel (HCN) “EFree” coil (Bruker Biospin). All spectra were obtained at 13 kHz MAS unless otherwise indicated. A constant flow of cooled gas was used to regulate sample temperature. Spectra were acquired with Bruker Topspin, processed with NMRPipe, and analyzed with the CcpNmr Analysis program [259, 260]. Chemical shifts were referenced to 4,4-dimethyl-4-silapentane-1-1 sulfonic acid by indirectly referencing to the ^{13}C signals of adamantane [261]. Secondary structure analysis was derived from empirical reference data [266, 312]. Fully α-helical and β-sheet models of apoA-I were created in Chimera [23]. Synthetic NMR spectra were generated using NMRPipe and Sparta+ as previously described from the secondary structure models and also from a crystal structure of C-terminally truncated apoA-I (PDB 3R2P, ref. [313]) [5, 259, 262, 263].

5.3 RESULTS

5.3.1 Amyloid formation occurs in non-oxidized apolipoprotein A-I upon seeding with pre-formed amyloid fibrils

Amyloid fibrils often can act as nuclei to induce fibril formation in natively-folded protein. To test whether H_2O_2-wt-ApoA-I amyloid fibrils can induce amyloid formation in intact wt-ApoA-I, non-oxidized wt-ApoA-I was incubated in the presence of 0, 1, and 10% (w/w) pre-formed H_2O_2-ApoA-I amyloid seeds and ThT fluorescence was monitored over time. The results are summarized in Figure 5.1. As expected, no significant increase in ThT fluorescence was observed in the presence of 0% seeds even after 23 days of incubation. In the presence of 1% seeds, ThT fluorescence is observed to increase after a lag phase of about 14 days. The amyloid elongation phase lasts another 9 days, as indicated by a steady increase in ThT fluorescence before it plateaus. In the presence of 10% seeds fibrils form much more quickly; the lag phase ends after 8 days and the elongation phase begins to conclude at after another 6 days of incubation.

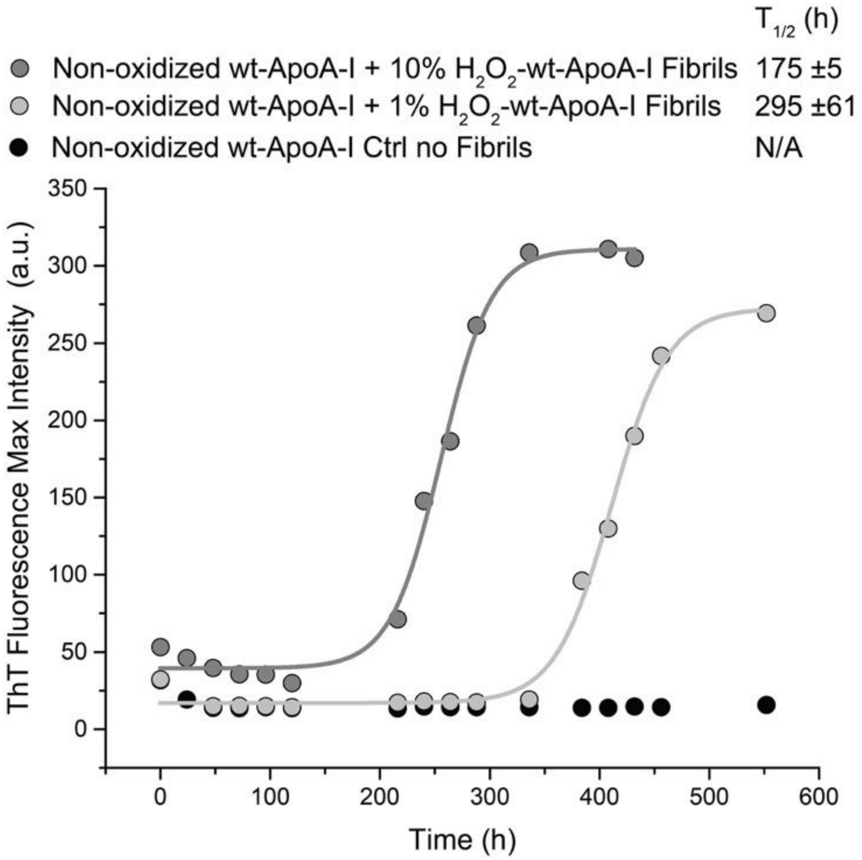


Figure 5.1. Seeding of non-oxidized wt-ApoA-I with molar 10% and 1% of pre-formed H_2O_2-wt-ApoA-I aggregates

As a control, non-oxidized wt-ApoA-I was incubated in the absence of seeds (black dots). Total protein concentration in each sample was 1.0 mg/ml. The samples were incubated in fibrillation buffer (pH 6.0) at 37 °C with continuous vortexing at 800 rpm, and then ThT fluorescence was measured at the indicated time points. Solid lines are fitting of the experimental values by sigmoidal curves. Mean $T_{1/2}$ *and SEM values from ≥3 independent experiments are reported in figure. N/A = not applicable.*

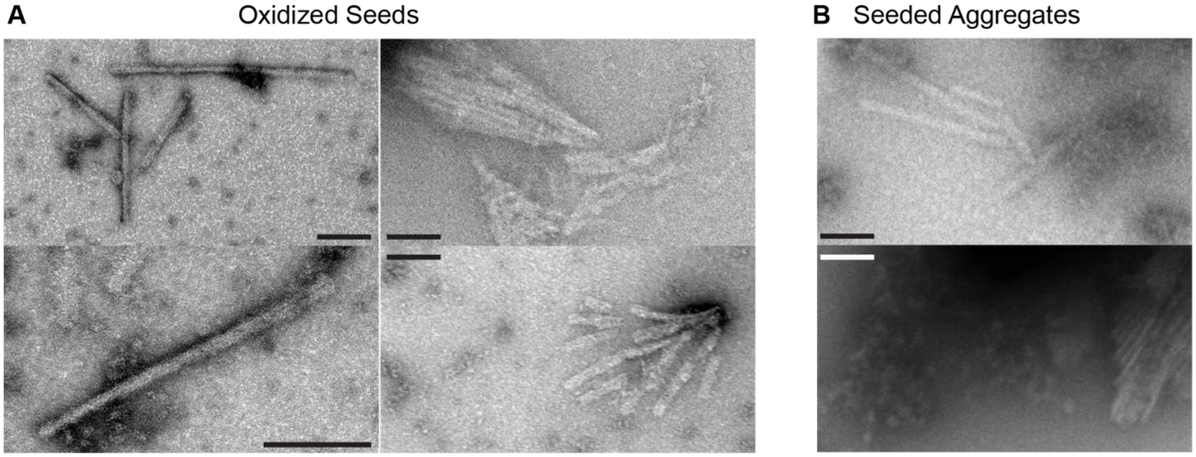


Figure 5.2. TEM analysis of H_2O_2-ApoA-I seeds and seeded non-oxidized apoA-I aggregates

Negative staining EM analysis was performed on H_2O_2-ApoA-I seeds after 4 days under fibrillation conditions (panel A) and on non-oxidized apoA-I seeded with 10% H_2O_2-ApoA-I seeds and incubated for 20 days under fibrillation conditions (panel B). For each sample, two micrographs of different grid areas are reported at the top and bottom of each column. Two independent samples are reported in panel A, left and right columns. These images illustrate well the variability in morphology typical of amyloid fibrils of apoA-I that are oxidized at Met. Bars represent 100 nm.

Both the seeds and the seeded samples form aggregates that are visible by eye. The H_2O_2-wt-ApoA-I seeds display morphological heterogeneity by TEM (Figure 5.2, a), visible as short single fibrils and large clusters of fibrils. Curly structures were also interspersed among the fibrils in the micrographs [302, 314]. The non-oxidized wt-ApoA-I fibrils formed in the presence of 10% H_2O_2-wt-ApoA-I seeds are not distinguishable from H_2O_2-wt-ApoA-I fibrils (Figure 5.2, b). Aggregates were not visible by eye in control samples of non-oxidized wt-ApoA-I following 23 days of incubation, however TEM images revealed several small clusters of protein reminiscent of the amorphous-appearing aggregates of P23T hγD (Figure 5.3) [5].

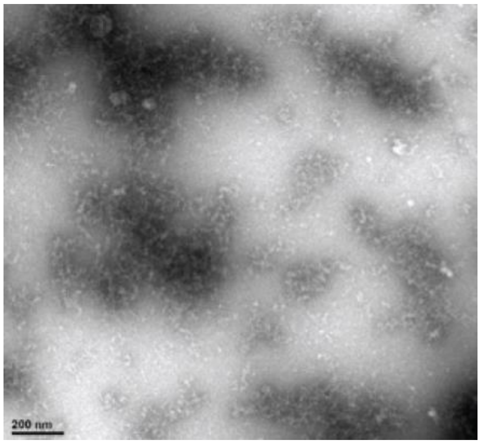


Figure 5.3. TEM analysis of non-oxidized apoA-I in the absence of H_2O_2-ApoA-I seeds

TEM images were obtained following 7 days under fibrillation conditions and 7 days at 4° C. Scale bar is 200 nm.

5.3.2 Non-oxidized apolipoprotein A-I adopts a partial β-sheet structure upon seeding with pre-formed oxidized amyloid seeds

FTIR was employed to monitor the secondary structure of non-oxidized wt-ApoA-I in the presence and absence of 10% (w/w) H_2O_2-wt-ApoA-I seeds over an 8 day incubation period. Non-oxidized and natively folded wt-ApoA-I was found to be predominantly α-helical and random coil in agreement with previous work, manifesting in the FTIR spectrum as single peak at about 1655 cm^{-1} in the amide I region (Figure 5.4, a) [302, 315]. The intensity of the peak does not appear to change significantly over 48 hours or 8 days of incubation, suggesting that the native fold of wt-ApoA-I remains intact. A peak at about 1622 cm^{-1} becomes visible in the seeded sample after 8 days of incubation (Figure 5.4, a). This band indicates the formation of a β-sheet secondary structure that is characteristic of amyloid structures. The intensity of the 1655 cm^{-1} band remains unchanged, in agreement with previous observations that substantial amounts of the native structure persists in the wt-ApoA-I fibrils [302, 315].

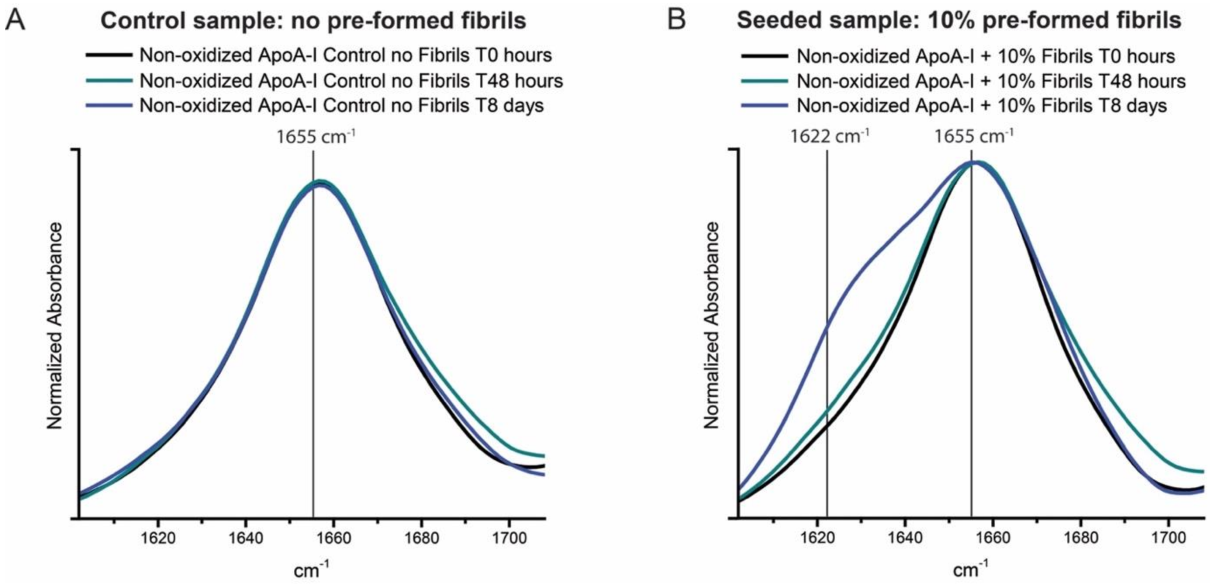


Figure 5.4. FTIR spectra in the amide I region (1602-1708 cm^{-1}) of non-oxidized wt-ApoA-I after seeding with pre-formed H_2O_2-wt-ApoA-I aggregates

Control (panel A) and 10% (molar) seeded (panel B) mixtures before incubation under fibrillation conditions (T0) and after 48 h and 8 day incubations.

5.3.3 Seeded non-oxidized apolipoprotein A-I aggregates retain a partial native structure

Oxidized H_2O_2-wt-ApoA-I seeds and non-oxidized wt-ApoA-I (10% seeded) were compared using MAS ssNMR in order to identify changes in the local secondary structure of individual residues and residue types [5, 258, 316]. Samples were uniformly labeled with ^{13}C and ^{15}N. We note that the seeds present in the non-oxidized seeded sample were prepared at isotopic natural abundance and attribute less than 1% towards the ^{13}C signals in that sample. The one dimensional direct excitation ^{13}C MAS ssNMR spectra of the seeds and seeded sample (Figure 5.5, a, full size: Figure C.1, a,b) are very similar. The most obvious difference is observed as a decrease in the intensity of a resonance frequency at 33 ppm upon oxidation, as indicated by a gray arrow (Figure 5.5, a). This resonance frequency is consistent with the chemical shifts expected for unoxidized Met Cγ and Cβ, which change dramatically upon oxidation of Met side chains [317, 318]. There is an additional decrease in the intensity of a resonance frequency at 17 ppm (gray arrow), which we assign as Met Cε. We also note the appearance of a small peak near 39 ppm (black arrow) that is observed in oxidized Met but is not present in unoxidized Met [317, 318].

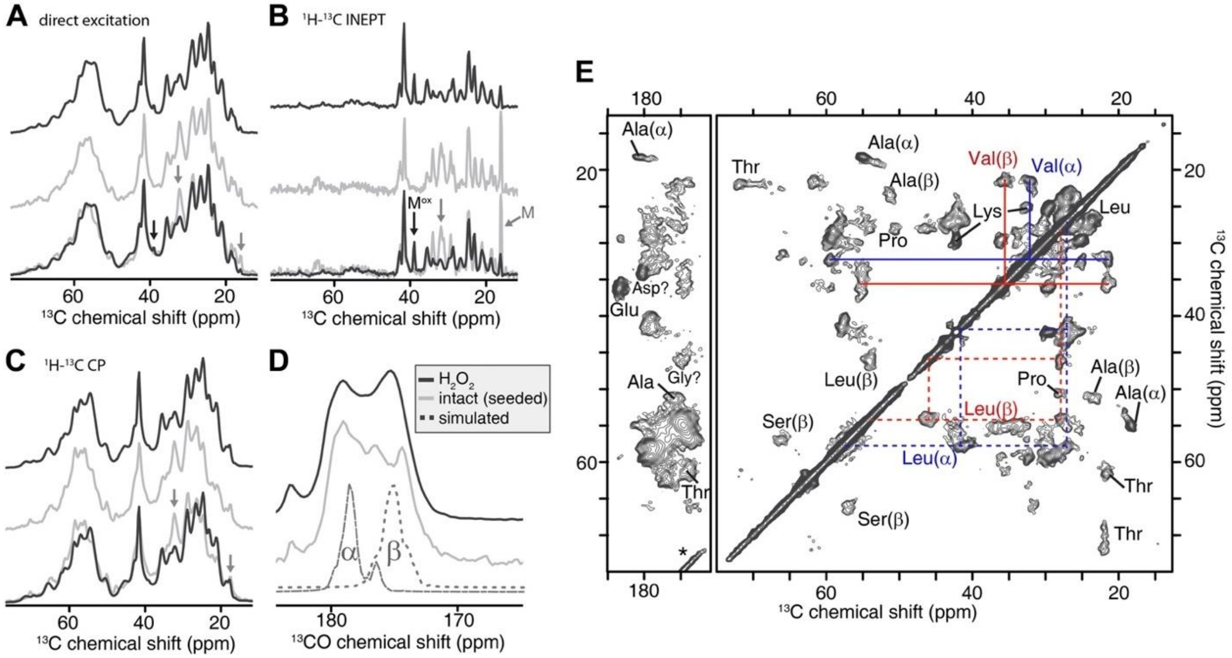


Figure 5.5. MAS ssNMR spectral analysis of pelleted [U-^{13}C,^{15}N] wt-ApoA-I aggregates

(A) Direct excitation ^{13}C spectra of H_2O_2-wt-^{13}C,^{15}N-ApoA-I aggregates (black; top) and aggregates formed by non-oxidized wt-^{13}C,^{15}N-ApoA-I incubated for 8 days with 10% (molar) of unlabeled H_2O_2-wt-ApoA-I seeds (light gray; middle). An overlay of the two spectra is reported at the bottom. The black and gray arrows mark the ^{13}C signals of oxidized and unoxidized Met side chains, respectively. (B) ^{13}C spectra obtained via ^{1}H-^{13}C refocused INEPT, which only shows highly dynamic residues. (C) ^{13}C spectra of the same samples using ^{1}H-^{13}C cross polarization, which is selective for the more rigid parts of the pelleted protein samples. In panels (A-C) only aliphatic spectral regions are shown; full spectra are reported in Figure C.1. (D) ^{1}H-^{13}C CP spectra in the carbonyl region of H_2O_2-wt-^{13}C,^{15}N-ApoA-I (black; top) and non-oxidized wt-^{13}C,^{15}N-ApoA-I after seeding (light gray;

middle). Dotted lines (bottom) show simulated spectra for hypothetical fully α-helical and β-sheet structural models of ApoA-I, to illustrate the secondary structure dependence. (E) 2D ^{13}C*-*^{13}C *spectrum of the same* H_2O_2*-wt-*^{13}C*,*^{15}N*-ApoA-I as in panels (A-D), using 20 ms DARR mixing. The amino acid type assignments of select cross peaks are indicated. Color-coded lines connect sets of peaks from co-existing α -helical and β-sheet Val and Leu. All spectra were acquired at 600 MHz (*^{1}H *frequency).*

Dynamic differences between the oxidized seeds and the non-oxidized seeded sample were also investigated by MAS ssNMR. Direct excitation spectra capture ^{13}C signals from both rigid and mobile protein domains. The most dynamic parts of the aggregated samples are detected in spectra obtained from experiments modified from solution NMR (INEPT) (Figure 5.5, b, full size: Figure C.1, c,d). As before, there is a striking similarity observed between the INEPT spectra of the seeds and the seeded samples. Met Cε, Cγ, and Cβ frequencies are assigned in the non-oxidized seeded sample at 16.7, 32.5, and 34.6 ppm, respectively and are marked with gray arrows in Figure 5.5, b. The intensity of the Met peaks is dramatically decreased in the oxidized seeds. A single peak at 39 ppm is observed in the oxidized seeds that is not present in the non-oxidized seeded sample. The remaining peaks do not differ significantly between samples.

In addition to Met and some signals from aromatic residues (Phe and/or Tyr), Lys is also detected in the INEPT spectra. Interestingly, there are no observable signals that can readily be assigned to other residue types. The few signals observed indicate that the number of highly dynamic and unstructured residues is relatively small. This contrasts from the highly unstructured C-terminal domain of monomeric non-oxidized wt-ApoA-I observed by solution NMR [309]. However, it is feasible the lack of signals is a result of surrounding ordered domains interacting with the unstructured regions, resulting in partial immobilization [3, 37].

Rigid regions of the aggregates were probed using ramped cross-polarization (CP) based experiments (Figure 5.5, c, full size: Figure C.1, e,f). The spectra of the oxidized seeds and non-oxidized seeded samples are again very similar, however there are several differences which we interpret as revealing subtle changes in the structure and dynamics of the immobilized core and assemblies within the aggregates. These differences are apparent not only in the Met side chain but also in the Met backbone (Cα and CO). The chemical shifts of backbone carbons are exquisitely sensitive to secondary structure changes, allowing for a direct comparison between the local secondary structure of individual residues in the seeds and the seeded samples. A mixture of α-helical and β-sheet structure are observed in both samples, consistent with the above FTIR data and earlier studies [302, 316]. There is a greater β-sheet content in the core of the H_2O_2-wt-ApoA-I seeds than in the non-oxidized (but seeded) wt-ApoA-I aggregates (Figure 5.5, d). Several amino acid types (including Ala, Val, Ser, and Leu) exhibit both α-helical and β-sheet structure, as shown in the CP-based 2D ^{13}C-^{13}C DARR spectrum of the ^{13}C,^{15}N-labeled H_2O_2-wt-ApoA-I seeds (Figure 5.5, e), consistent with prior analysis and results [302, 315, 316].

The native fold of wt-ApoA-I is predominantly α-helical without any β-sheet content. This is demonstrated by a synthetic NMR spectrum simulated from predicted NMR chemical shifts for the native X-ray structure of C-terminal truncated Δ(185-243)-ApoA-I (Figure 5.6, red). An overlay of the synthetic spectrum of the native protein and the experimental spectrum of H_2O_2 wt-^{13}C,^{15}N-ApoA-I aggregates clearly shows that there is extensive α-helical content in the latter. There is greater heterogeneity in the non-native β-sheet content than the native-like α-helical content in the H_2O_2 wt-^{13}C,^{15}N-ApoA-I aggregates, evident by broader peaks observed in the former. Taken together with an observed similar peak pattern between the synthetic and experimental spectra of the both the H_2O_2 wt-^{13}C,^{15}N-ApoA-I seeds and the non-oxidized but seeded wt-^{13}C,^{15}N-ApopA-I aggregates (Figure C.2), it is likely that there is extensive native-like structure present in both samples, with more native-like structure found in the non-oxidized aggregates.

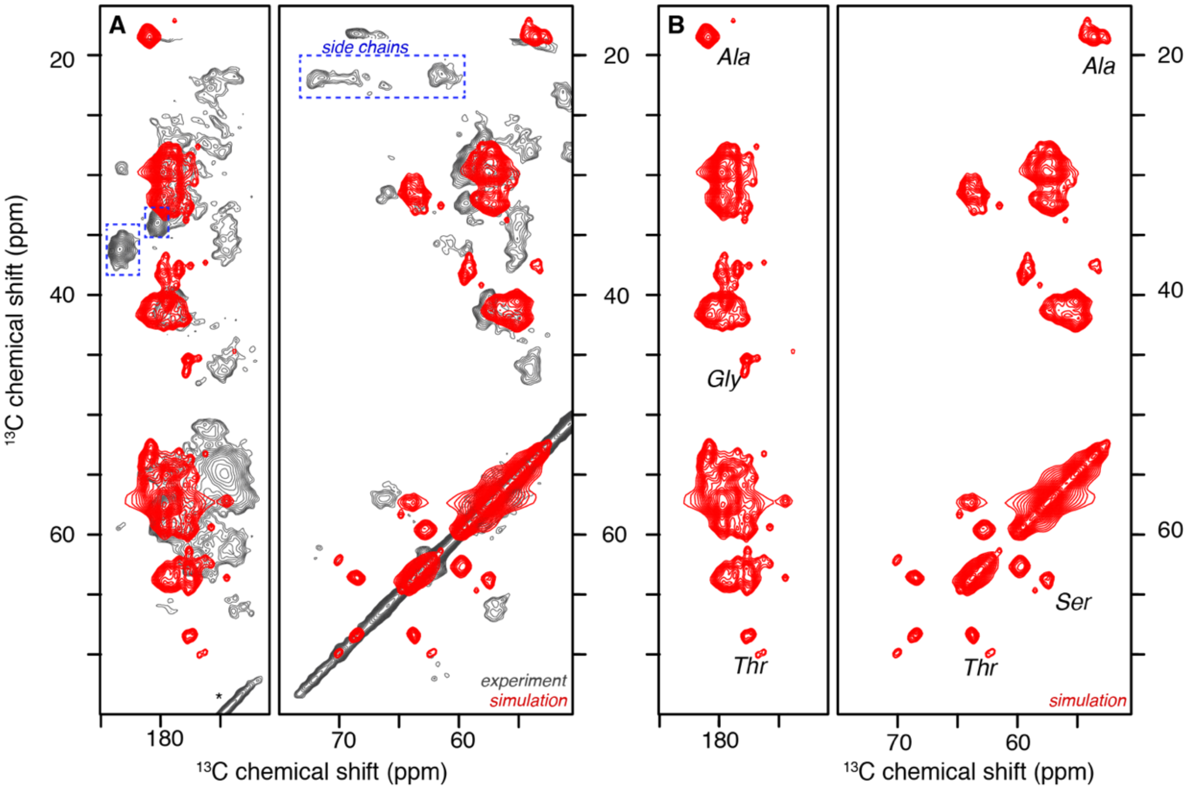


Figure 5.6. Comparison of experimental MAS NMR 2D ^{13}C-^{13}C spectrum of H_2O_2 wt-^{13}C,^{15}N-ApoA-I aggregates to a simulated spectrum based on the X-ray determined α-helical native structure of C-terminal truncated Δ(185-243)-ApoA-I

(A) Overlay of experimental and simulated data [313]. (B) Simulation by itself. Since the ^{13}C chemical shift prediction with the SPARTA+ program [263] is limited to backbone and Cβ carbons, only part of the spectrum is shown. Colored boxes mark side chain peaks visible in the experimental data that were not simulated.

5.4 DISCUSSION

These results in context with the recent results from the Cavigiolio lab that were discussed in the Introduction are presented in a schematic summary in Figure 5.7. Briefly, exchangeable apolipoproteins display structural flexibility that aids in their transition between lipid-bound and lipid-free states. Lipid-free apoA-I self-associates and this phenomenon provides stabilization to the native state structure. Both apoA-I oxidized at Met and thermally treated apoA-I associate with macrophages to extract cellular cholesterol. Interestingly, both treatments disrupt self-association; however, the thermal treatment disrupts self-association only temporarily, while disruption from oxidation at Met is non-reversible. This irreversibility lends apoA-I oxidized at Met vulnerable to aggregation. The self-associated native structure of the lipid-free apoA-I functions to protect apoA-I monomers from aggregating into amyloid-like fibrils. This is reminiscent of transthyretin, which is an amyloid forming protein in the monomeric state but remains soluble as a tetramer [319]. Of note, drugs that stabilize transthyretin tetramers are successful at protecting against cardiac amyloidosis [320].

ApoA-I oxidized at Met forms amyloid fibrils that fluoresce in the presence of ThT, have a significant amount of β-sheet structure, and can propagate through seeding. Interestingly, the FTIR and MAS ssNMR results indicate that the oxidized apoA-I fibrils contain a higher amount of β-sheet structure than the seeded (but not oxidized) apoA-I aggregates. Furthermore, although apoA-I fibrils adopt a β-sheet structure in the amyloid core, the native structure of apoA-I is not completely abolished and much of the α-helical content is retained in the aggregates. The N-terminal 4-helix bundle motif present in lipid-free self-associated apoA-I likely becomes destabilized upon Met oxidation and monomerization, resulting in surface changes that expose hydrophobic residues that readily bind to lipids including cholesterol. Although the specific residues within the amyloid core could not be assigned in the primary sequence, various algorithms exist that aid in predicting amyloidogenic sequences [321-324]. Four separate algorithms (3D Profile (via ZipperDB), Zyggregator, CamSol, and Tango) were used to identify residues potentially found within the amyloid core; the results are highlighted in Figure C.3 [321-324]. Briefly, residues 13-22 and 50-59 at the N-terminal end and residues 224-332 at the C-terminal end were reproducibly identified as candidates for existing within the amyloid core.

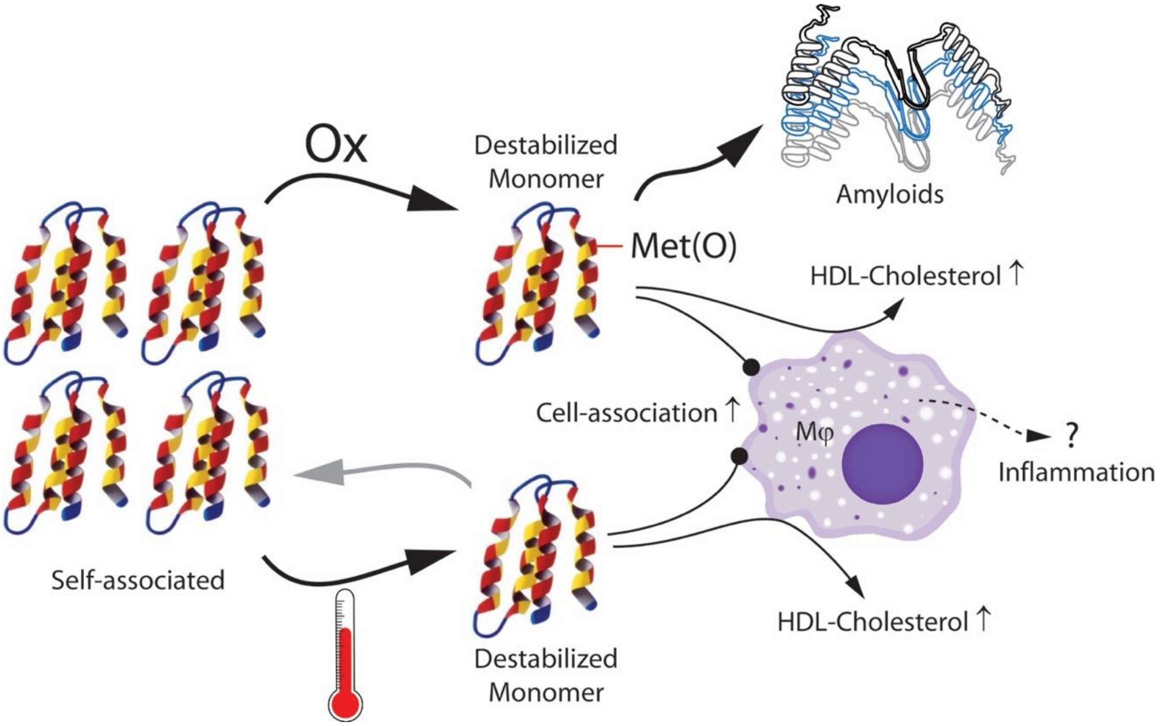


Figure 5.7. Schematic of the consequences of the destabilization of self-associated apolipoprotein A-I

ApoA-I native self-association is disrupted by thermal treatment or Met-oxidation. Destabilized monomeric apoA-I associates with macrophages more strongly and extracts cholesterol more efficiently than native apoA-I. Destabilized monomeric Met(O)-ApoA-I is also amyloidogenic.

These results give support to the idea that apoA-I can aggregate through both amyloid-like and native-like aggregation pathways. The native-like aggregation pathway is most likely driven by interactions between the α-helices found in the native structure of lipid-free apoA-I, while the amyloid-like pathway is driven by β-sheet elongation. It is unclear whether native-like intermolecular interactions between lipid-free apoA-I protects from or promotes further aggregation through the amyloid-like pathway following secondary nucleation events. If the former case is true, therapeutics that promote native-like apoA-I aggregation may be beneficial for fighting VA and atherosclerosis.

5.5 ACKNOWLEDGEMENTS

I thank Dr. Giorgio Cavigiolio for his assistance in interpreting the background literature, for his enthusiasm for science, and for a fruitful collaboration.

This work was supported in whole or part by National Institutes of Health Grants R01 HL113059 to G.C., R01 GM112678 to P.C.A.v.d.W, and T32 GM088119 to J.C.B. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

5.6 AUTHOR CONTRIBUTIONS

GC conceived and coordinated the study. JCB and PCAvdW designed, performed and analyzed all TEM and MAS ssNMR experiments. All other experiments were designed and analyzed by GC and AW. GKLC produced the protein samples. GKLC and GC performed the Thioflavin T fluorescent studies. GKLC and GC performed the FTIR experiments. GKLC, API, JCW, NJL and GC performed additional experiments discussed in the published manuscript. GC, AW, and PCAvdW wrote the published manuscript.

6.0 DISSERTATION SUMMARY AND THE BIG PICTURE

The aggregation of normally soluble protein in tissues is rampant and directly correlated to tissue dysfunction in protein deposition diseases (PDDs). Several classes of protein aggregates are found in PDDs, including amyloid-like fibrils, native-like aggregates, and amorphous aggregates. Aggregate polymorphism is a common theme of PDDs that has implications in disease severity. In this dissertation, I explored aggregate polymorphism in proteins that have implications in neurodegenerative disease, cataract, and vascular amyloidosis. I specifically targeted my investigations towards httEx1, P23T hγD, and apoA-I. HttEx1 forms amyloid fibrils that have implications in Huntington's disease and P23T hγD is an extremely aggregation prone crystallin that is associated with eye lens cataract formation in very young children. ApoA-I is an important anti-atherosclerotic factor that has been identified in vascular amyloidosis and may be involved in atherogenesis upon oxidation of Met residues.

I show in Chapter 3.0 that httEx1 amyloid fibrils can form several polymorphs as a direct consequence of monomer concentration and the rate of monomer release into the aggregating reaction mixture. Furthermore, my results lend further support to the idea that the proline rich domain that flanks the polyQ amyloid core is an important factor in httEx1 fibril polymorphism. In Chapter 4.0, I show that cataract-associated P23T hγD can be induced to form amyloid-like fibrils or native-like aggregates in a pH dependent manner. Of note, MAS ssNMR revealed that homogenous, well-structured native-like aggregates form at a pH that matches the pH of the eye lens. Surprisingly, the native-like aggregates of P23T hγD appear strikingly amorphous by TEM. In Chapter 5.0, I show that apoA-I amyloid fibrils retain a significant amount of native structure and apoA-I oxidized at Met residues can aggregate through native-like and amyloid-like aggregation pathways.

These results illustrate several aspects of PDDs that extend beyond Huntington's disease, cataract, and vascular amyloidosis. In amyloid forming proteins specifically, flanking domains have been repeatedly shown to be important for influencing fibril polymorphism. Of note, many proteins can be induced to adopt multiple classes of protein aggregates and some aggregates can be amyloid-like while retaining a surprising amount of native-like structure. Furthermore, it is important to study the amorphous-looking aggregates that are found in PDDs as they can be surprisingly well structured.

APPENDIX A

A.1. SUPPORTING INFORMATION FOR CHAPTER 3

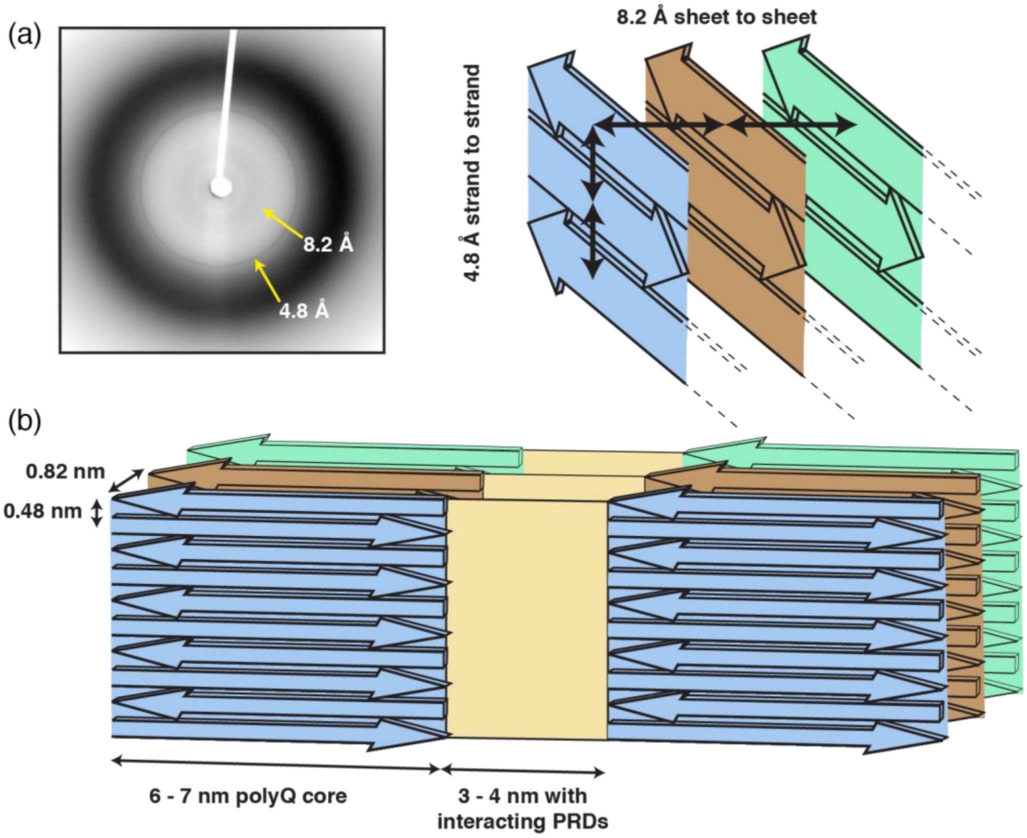


Figure A.1. Q44-httEx1 fibril dimensions.

(a) X-ray powder diffraction pattern of Q44-httEx1 reveals that the distance between β-strands along the fibers axis is 4.8 Å and the distance between β-sheets (depicted in different colors in cartoon) is 8.2 Å, in agreement with previously work [3, 325]. (b) Individual filaments are proposed to interact through interactions of the PRDs in Q44-httEx1 polymorphs [3].

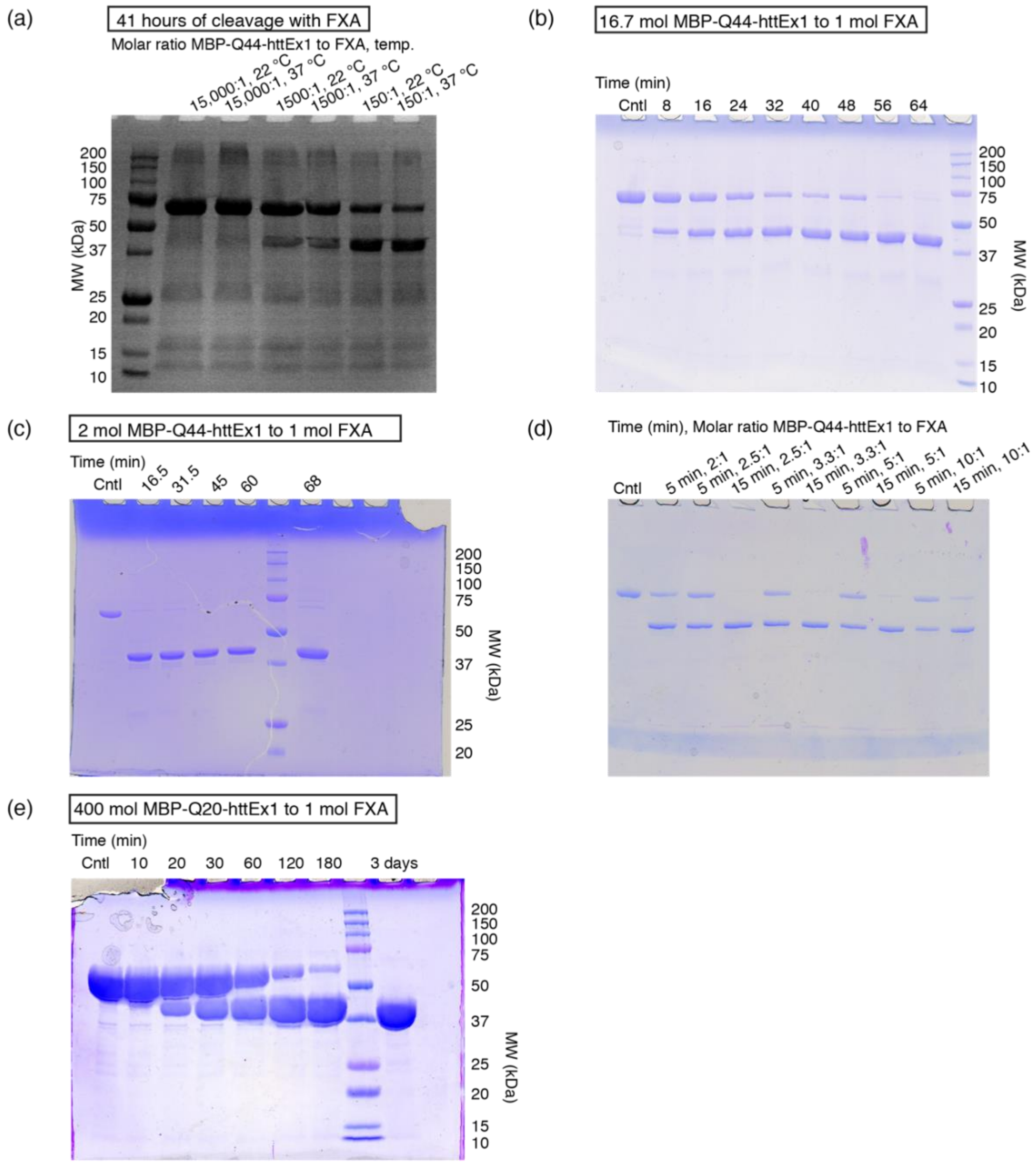


Figure A.2. Cleavage kinetics of MBP-httEx1 fusion proteins by FXA protease.

(a) Cleavage of MBP-Q44-httEx1 after 41 hours at fusion protein to protease molar ratios (FP:P ratios) of: 15,000:1, 1500:1, and 150:1, and temperatures of 22 and 37 °C. Cleavage did not reach completion under all circumstances. (b) Cleavage of MBP-Q44-httEx1 over 64 minutes at 37 °C with an FP:P molar ratio of 16.7:1. Cleavage is nearly complete after 64 minutes. (c) Cleavage of MBP-Q44-httEx1 over 68 minutes at 37 °C at an FP:P molar ratio of 2:1. Cleavage appears complete at 16.5 minutes. (d) Cleavage of MBP-Q44-httEx1 at 37 °C after 5 and 15 minutes for various FP:P ratios: 2:1, 2.5:1, 3.3:1, 5:1, and 10:1. e) Cleavage of MBP-Q20-httEx1at 22 °C and an FP:P molar ratio of 400:1. Most of the fusion protein has been cleaved after 180 minutes and is no longer observed after 3 days.

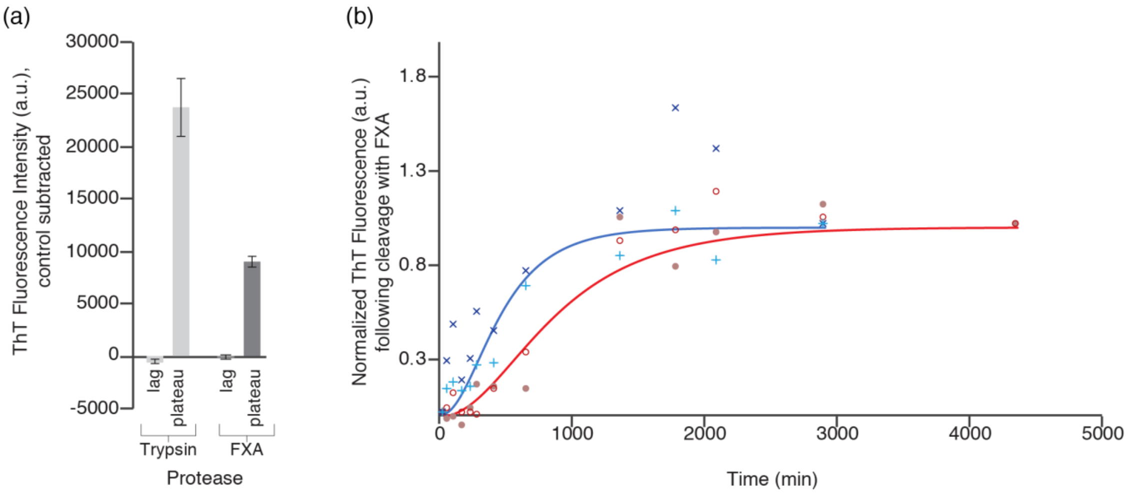


Figure A.3. ThT Fluorescence following cleavage with trypsin and FXA.

(a) Average ThT fluorescence in the lag and plateau phase for MBP-Q44-httEx1 cleaved with trypsin (light gray) and FXA (dark gray). The average fluorescence of a control sample containing MBP-Q44-httEx1 only was subtracted for each. Lag phase was measured at 50 and 60 minutes for trypsin (n = 3) and FXA (n = 2), respectively. Plateau phase was measured at 4 and 3 days for trypsin (n = 3) and FXA (n = 2) respectively. (b) Red: ThT fluorescence of Q44-httEx1 in the absence of seeds (n = 2). The values recorded for the individual measurements are marked in empty and filled red circles. Blue: ThT fluorescence of Q44-httEx1 in the presence of 20% (molar) Q44-httEx1 preformed seeds (n = 2). The values recorded for the individual measurements are marked as blue X's and +'s. The data was fit using AmyloFit [225].

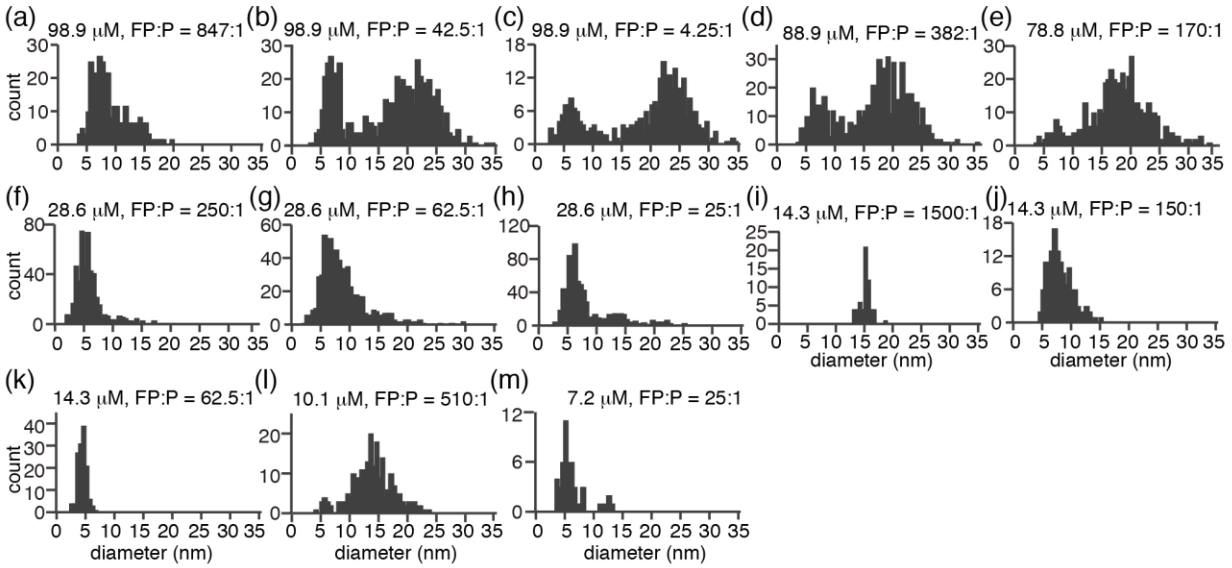


Figure A.4. Distribution of fibril widths visible by TEM for Q44-httEx1 fibrils formed at 37 °C, varying concentrations of MBP-Q44-httEx1, and varying molar ratios of MBP-Q44-httEx1 to FXA protease. *(a) 98.9 μM MBP-Q44-httEx1, 847 moles per 1 mole FXA. (b) 98.9 μM MBP-Q44-httEx1, 42.5 moles per 1 mole FXA. (c) 98.9 μM MBP-Q44-httEx1, 4.25 moles per 1 mole FXA. (d) 88.9 μM MBP-Q44-httEx1, 382 moles per 1 mole FXA. (e) 78.8 μM MBP-Q44-httEx1, 170 moles per 1 mole FXA. (f) 28.6 μM MBP-Q44-httEx1, 250 moles per 1 mole FXA. (g) 28.6 μM MBP-Q44-httEx1, 62.5 moles per 1 mole FXA. (h) 28.6 μM MBP-Q44-httEx1, 25 moles per 1 mole FXA. (i) 14.3 μM MBP-Q44-httEx1, 1500 moles per 1 mole FXA. (j) 14.3 μM MBP-Q44-httEx1, 150 moles per 1 mole FXA. (k) 14.3 μM MBP-Q44-httEx1, 62.5 moles per 1 mole FXA. (l) 7.2 μM MBP-Q44-httEx1, 510 moles per 1 mole FXA. (m) 7.2 μM MBP-Q44-httEx1, 25 moles per 1 mole FXA.*

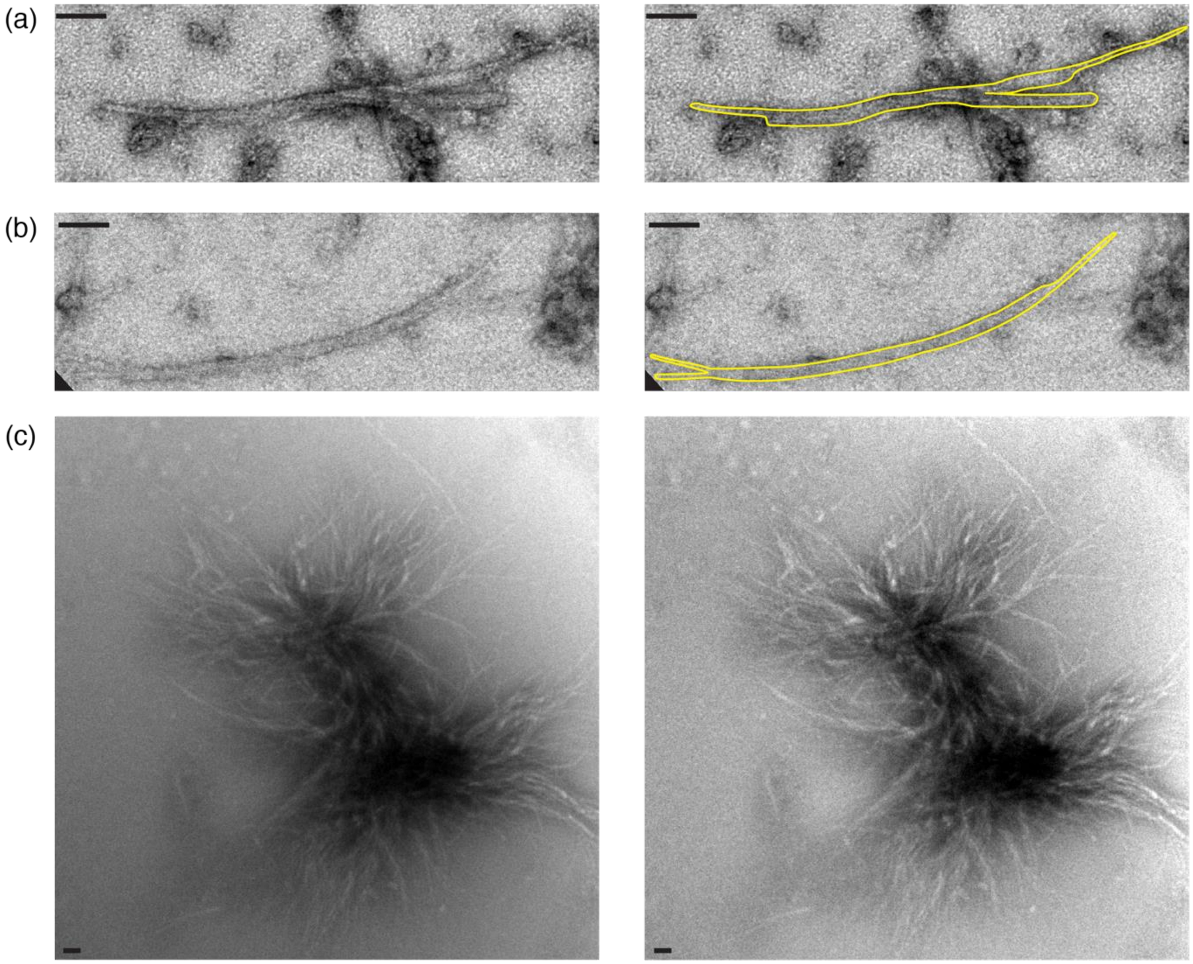


Figure A.5. Branching and bundling of Q44-httEx1 fibrils.

(a-b, left) Examples of fibrils which appear to branch, observed at high concentrations (88.9, and 98.9 μM). Fibrils are outlined (right) to highlight differences in width across the fibril axis and branching. (c) Left: Bundle of fibrils, original unaltered image. Right: Image was subjected to a band-pass filter to increase contrast and highlight the density of the bundled fibrils. All scale bars are 100 nm.

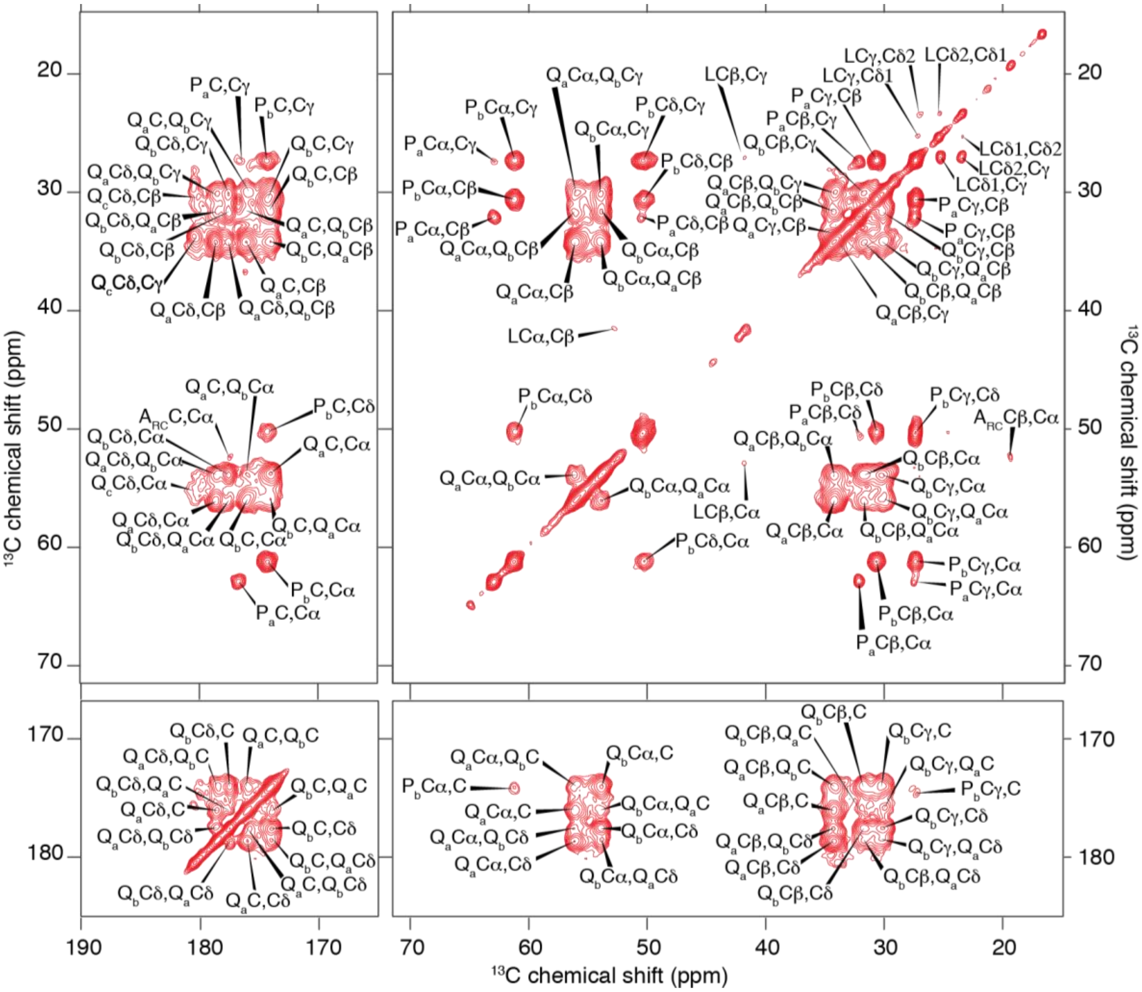


Figure A.6. PDSD spectrum of ΔN15-Q44-httEx1.

Fully assigned 250 ms PDSD spectrum. Pro$_{RC}$ is labeled as Pro$_A$ and Pro$_{PPII}$ is labeled as Pro$_B$.

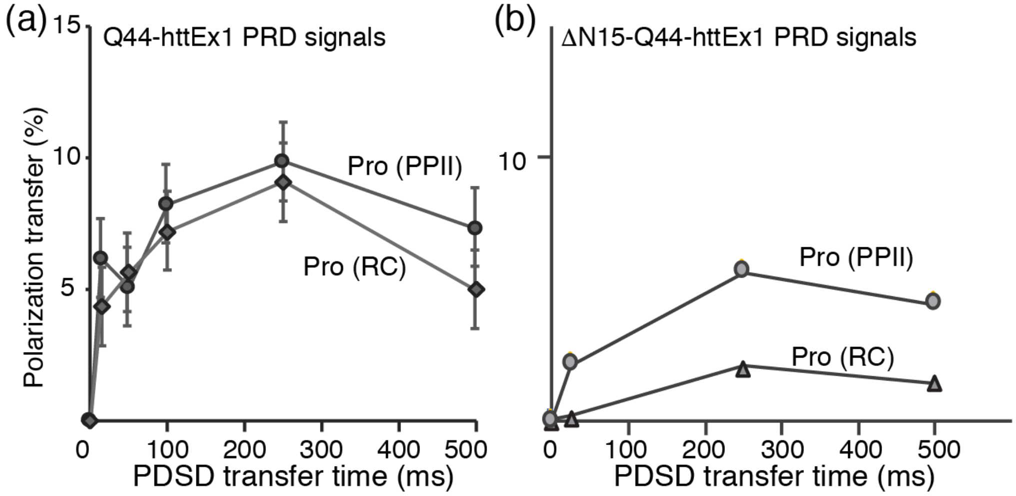


Figure A.7. Polarization transfer curves.

(a) PDSD buildup profiles for proline residues with random coil (RC), circles, and PPII-structured, diamonds, secondary structure in Q44-httEx1 [3], and (b) for RC, circles, and PPII-structured, triangles, Pro residues in ΔN15-Q44-httEx1. Both Q44-httEx1 and ΔN15-Q44-httEx1 were formed at 22 °C. Panel (a) reprinted with permission from Lin, H. K., ***Boatz, J. C.****, Krabbendam, I. E., Kodali, R., Hou, Z., Wetzel, R., Dolga, A. M., Poirier, M. A., van der Wel, P. C. A. Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core, Nature Communications 8:15462. Copyright (2017) Lin, H. K., et. al. [3]*

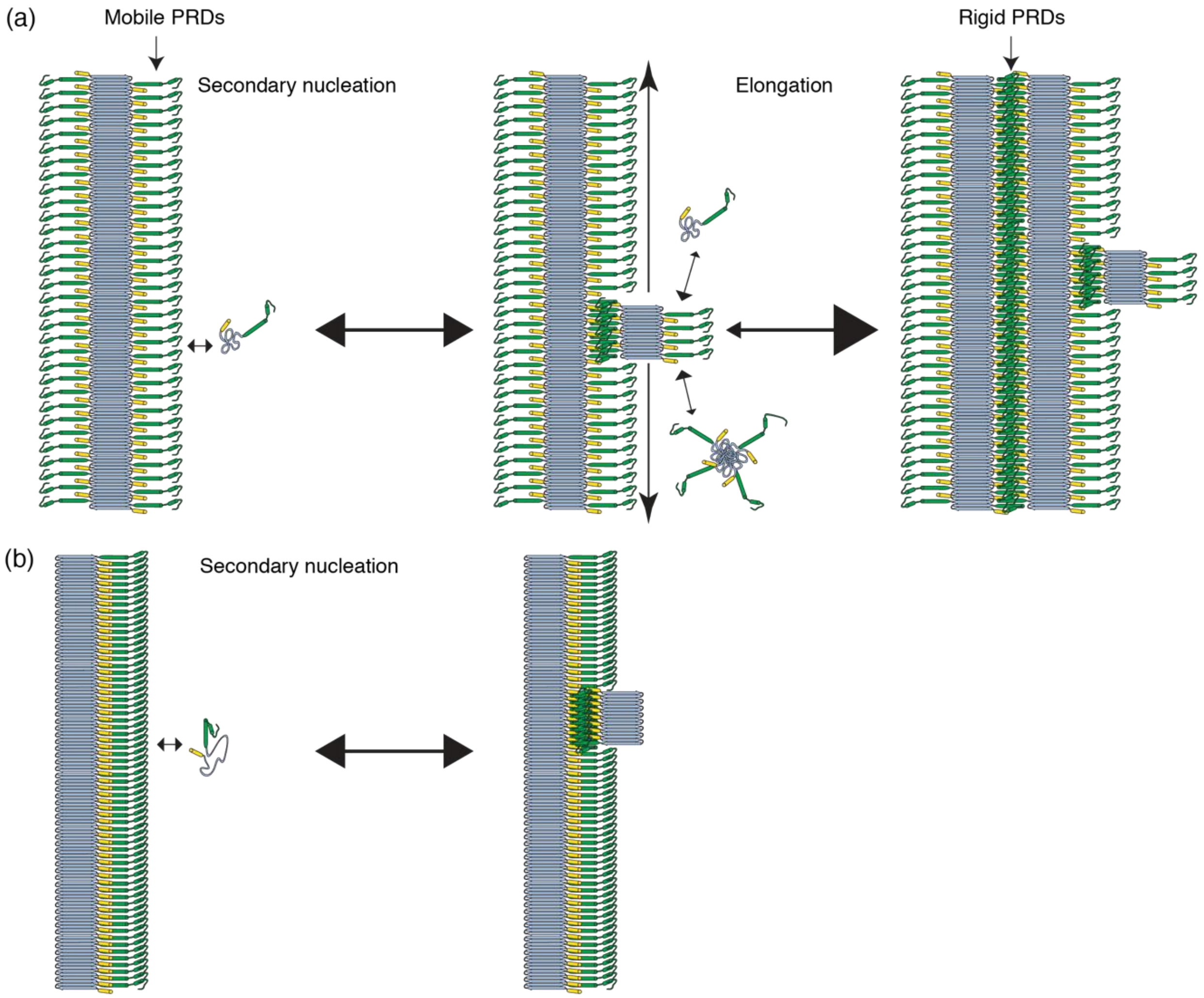


Figure A.8. Hypothetical flanking domain distribution and implications for secondary nucleation

(a) An extremely symmetric distribution of flanking domains is expected to promote secondary nucleation and elongation. (b) An extremely asymmetric distribution of flanking domains may experience some secondary nucleation but is not expected to experience any elongation of the nucleated fiber.

A.2. TRYPSIN CLEAVAGE PRODUCTS OF MBP-Q44-HTTEX1 AND FREE MBP

In Figure 3.6, d, the area under the curve of the MBP peak at 6 hours is 47% the area at 5 minutes. Peak area in deconvoluted ESI-TOF MS spectra cannot be quantitatively compared to protein concentration in the injected sample, however the large difference in peak area between these two time points suggests that free MBP is degrading within the sample in the presence of trypsin. In contrast, free MBP appears stable by SDS-PAGE after 3 days of FXA cleavage [3]. The ESI-TOF MS spectrum at 6 hours was screened against the MW of 759 theoretical cleavage products of MBP-Q44-httEx1 by trypsin protease. A small peak was observed at 42623.5 Da that was not present at 30 seconds, 1 minute, or 5 minutes. This peak was identified as MBP(I[-391]-R[-1]) (theoretical MW 42624.7 Da). To confirm the presence of MBP cleavage products, an additional seven trials were screened within the first six hours of the cleavage reaction. Three MBP cleavage products were observed in at least two of the seven trials: MBP(I[-391]-R[-1]), MBP(D[-363]-K[-256]) (theoretical MW 12115.9 Da), and MBP(D[-363]-R[-295]) (theoretical MW 7721.7 Da). Figure A.1, b shows the ESI-TOF MS spectrum of one of the seven trials, where all three products are observed. Two additional cleavage products are visible by SDS-PAGE (Figure A.1, e), obtained from a reaction mixture with 10 μM MBP-Q44-httEx1 and an FP:P ratio of 1:1. These products are MBP(Y[-294]-R[-1]) and MBP(E[-255]-R[-1]). MBP(E[-255]-R[-1]) appears unstable, disappearing from the gel within the first five minutes of the reaction.

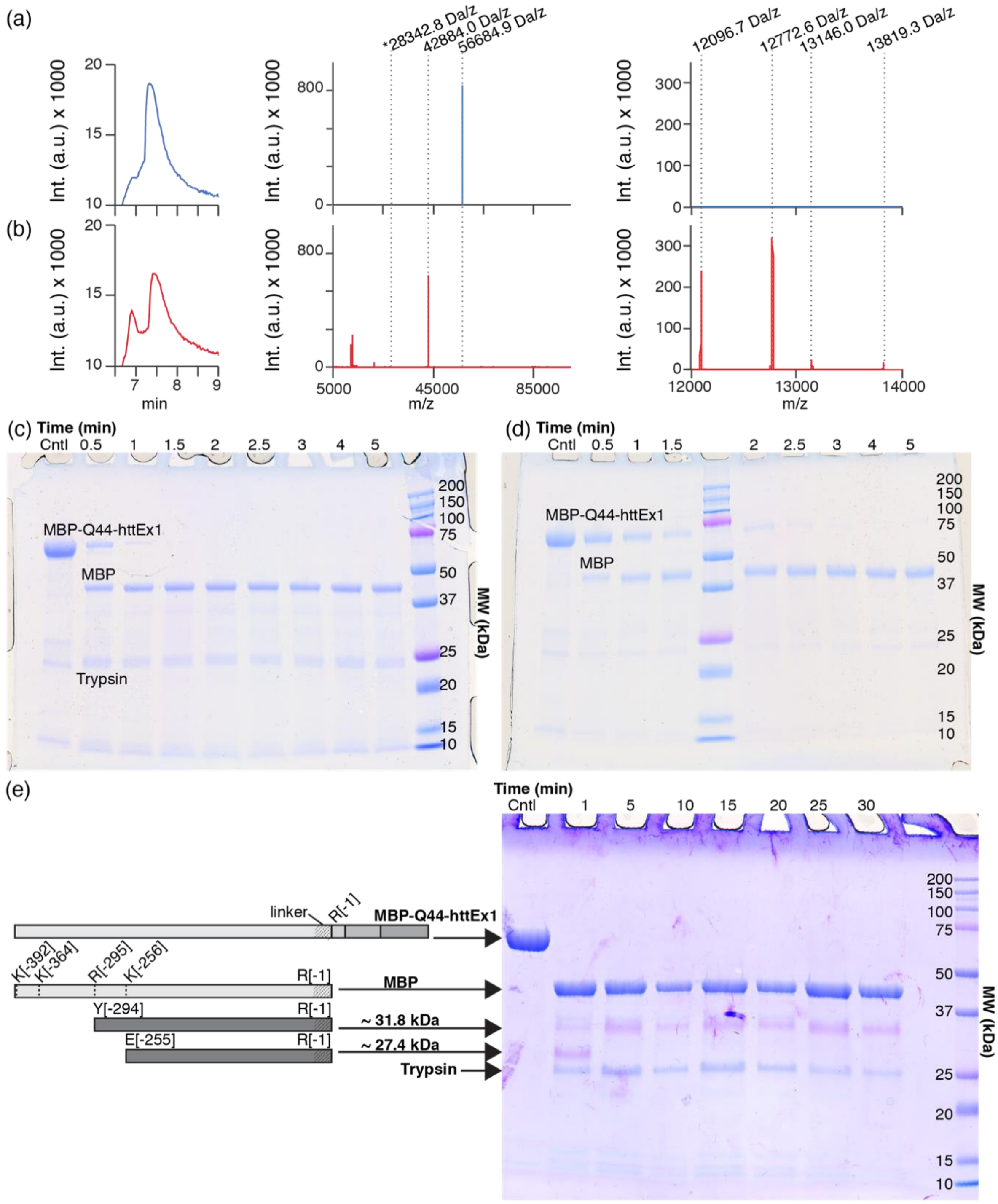


Figure A.1. Cleavage products of Q44-httEx1 by trypsin protease.

(a, left) Mass spectra chromatogram from LC elution profile of MBPhttExQ44 and (b, left) its trypsin cleavage products. (a,b center) Mass spectra were deconvoluted by maximum entropy from 6.6-8.2 minutes. MBP-Q44-httEx1 (56684.9 Da with z = 1 and 28342.8 Da with z = 2) is observed prior to trypsin cleavage, but not after. Free MBP

(42884.0 Da) is observed following trypsin cleavage but not prior to cleavage. (a,b, right) Zoomed in region of (b) from 12000-14000 m/z. ΔN15-Q44-httEx1 (12096.7 Da), ΔN9 htt exon1 (12772.6 Da), ΔN6 htt exon1 (13146.0 Da), and htt exon1 (13819.3 Da) are observed prior to ΔN15-Q44-httEx1 aggregation. (c,d) Cleavage of 10 μM MBP-Q44-httEx1 via trypsin protease with FP:P molar ratios of 1:1 (c) and 625:10 (d). (e) Cleavage with FP:P molar ratio of 1:2. Cleavage products with molecular weights of approximately 31.9 and 27.4 Da are observed. These bands likely represent MBP cleavage products MBP(Y100-R393) and MBP(E139-R393), respectively. MBP(E139-R393) appears unstable, disappearing from the reaction mixture within 5 minutes.

All MBP-Q44-httEx1 cleavage products observed, their MW, and their sequence are available in Table A.1. The cleavage products observed in each individual trial are recorded in Table A.2. Four additional trials were analyzed after 1 day of aggregation (Table A.3). ΔN15-Q44-httEx1 was not observed in these trials, suggesting complete aggregation into mature fibrils. Together these results suggest 8 total cleavage products of MBP-Q44-httEx1 resulting from cleavages at K2, K30, R[-295], and K[-256]: MBP(M1-K2), MBP(I[-391]-R[-1]), MBP(M1-K30), MBP(D[-363]-K[-256]), MBP(E[-255]-R[-1]), MBP(D[-363]-R[-295]), and MBP(Y[-294]-R[-1]). These cleavage sites were mapped onto the predicted structure of free MBP, derived from the X-ray crystal structure of MBP-L30 (1NMU in the protein data bank) as shown in Figure A.2 [226]. Notably, K30, R[-295], and K[-256] cleavage sites are surface residues exposed to water and located at the edges between α-helices and random coil regions of the protein. K2 is not mapped due to the flexibility of the MBP N-termini that prevented crystallization of this region. Residue K2 is the most frequent cleavage site observed in free MBP by ESI-TOF MS. The dynamics of the MBP N termini likely increase the accessibility of the K2 residue to trypsin cleavage compared to the accessibility of the K30, R[-295], and K[-256] cleavage sites.

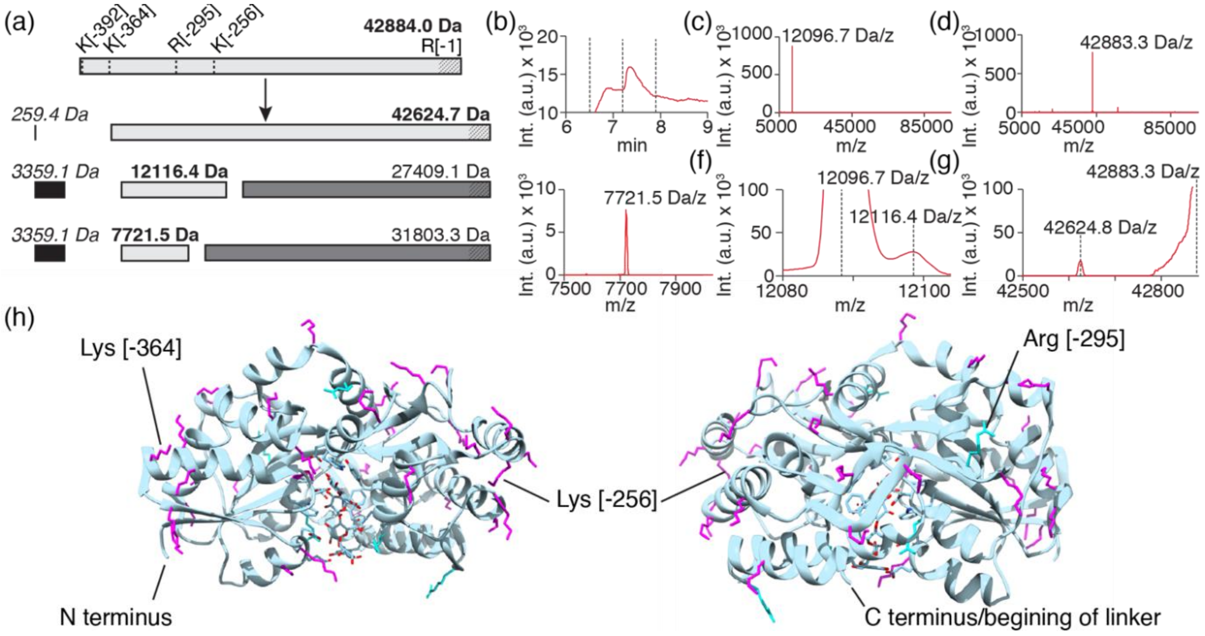


Figure A.2. Trypsin cleavage of free MBP.

(a) Free MBP (top, light gray, theoretical MW 42884.0, (bold) is cleaved by trypsin into several cleavage products. Cleavage sites are marked. Three cleavage products (light gray, I[-391]-R[-1], D[-363]-K[-256], and D[-363]-R[-295]) with theoretical MW of 42624.7, 12116.4, and 7721.5 Da (bold) were reproducibly observed by ESI-TOF MS. Two cleavage products (dark gray, E[-255]-R[-1] and Y[-294]-R[-1]) with theoretical MW of 27409.1 and 31803.3 Da were observed by SDS-PAGE but not by MS, likely due to low ionization or low stability of these products. Peptides (black, M1-K2, M1-K30) with theoretical MW of 259.4 and 3359.1 Da (oblique) are expected based on the presence of the other cleavage products observed, but below the ESI-TOF MS detection limit of 5000 Da/z. (b) Mass spectra chromatogram from LC elution profile of trypsin cleaved MBP-Q44-httEx1, with individual peaks observed between 6.5-7.2 and 7.2-7.9 minutes. (c) Deconvoluted mass spectrum from 6.5-7.2 minutes. ΔN15-Q44-httEx1 is observed as 12096.7 Da. (d) Deconvoluted mass spectrum from 7.2-7.9 minutes. Free MBP is observed as 42883.3 Da (e) Zoomed in region of deconvoluted mass spectrum from (c). MBP cleavage product D[-363]-R[-295] is observed at 7721.5 Da. (f) Zoomed in region of deconvoluted mass spectrum from (c). MBP cleavage product from D[-363]-K[-256] is observed at 12116.4 Da. (g) Zoomed in region of deconvoluted mass spectrum from (d). MBP cleavage product from I[-391]-R[-1] is observed at 42624.8 Da. (h) Predicted structure of free MBP, adapted from MBP fusion protein MBP-L30, 1NMU in the Protein Data Bank [226]. The C-terminal end of each Lys (pink) and

Arg (cyan) residue is a potential trypsin cleavage site, however cleavage is only observed at K2 (not shown), K30, R[-295], and K[-256].

Table A.1. Observed trypsin cleavage products of MBP-Q44-httEx1.

Name: Residues numbered relative to Q44-httEx1	MW (Da)	Sequence
MBP-Q44-httEx1	56685.3	MKIEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ VAATGDGPDI IFWAHDRFGG YAQSGLLAEI TPDKAFQDKL YPFTWDAVRY NGKLIAYPIA VEALSLIYNK DLLPNPPKTW EEIPALDKEL KAKGKSALMF NLQEPYFTWP LIAADGGYAF KYENGKYDIK DVGVDNAGAK AGLTFLVDLI KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK VNYGVTVLPT FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN AASGRQTVDA ALAAAQTNAA AASEFSSNNN NNNNNNNLGI EGRMATLEKL MKAFESLKSF QQQQQQQQQQ QQQQQQQQQQ QQQQQQQQQQ QQQQQQQQQQ QQQQPPPPPP PPPPPQLPQP PPQAQPLLPQ PQPPPPPPPP PPGPAVAEEP LHRPSGSHHH HHH
Free MBP	42884.0	MKIEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ VAATGDGPDI IFWAHDRFGG YAQSGLLAEI TPDKAFQDKL YPFTWDAVRY NGKLIAYPIA VEALSLIYNK DLLPNPPKTW EEIPALDKEL KAKGKSALMF NLQEPYFTWP LIAADGGYAF KYENGKYDIK DVGVDNAGAK AGLTFLVDLI KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK VNYGVTVLPT FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN AASGRQTVDA ALAAAQTNAA AASEFSSNNN NNNNNNNLGI EGR
MBP(I[-391]-R[-1])	42624.7	IEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ VAATGDGPDI IFWAHDRFGG YAQSGLLAEI TPDKAFQDKL YPFTWDAVRY NGKLIAYPIA VEALSLIYNK DLLPNPPKTW EEIPALDKEL KAKGKSALMF NLQEPYFTWP LIAADGGYAF KYENGKYDIK DVGVDNAGAK AGLTFLVDLI KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK VNYGVTVLPT FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN AASGRQTVDA ALAAAQTNAA AASEFSSNNN NNNNNNNLGI EGR
MBP(Y[-294]-[-1])*	31803.3	Y NGKLIAYPIA VEALSLIYNK DLLPNPPKTW EEIPALDKEL KAKGKSALMF NLQEPYFTWP LIAADGGYAF KYENGKYDIK DVGVDNAGAK AGLTFLVDLI KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK VNYGVTVLPT FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN AASGRQTVDA ALAAAQTNAA AASEFSSNNN NNNNNNNLGI EGR
MBP(E[-255]-R[-1])*	27409.1	EL KAKGKSALMF NLQEPYFTWP LIAADGGYAF KYENGKYDIK DVGVDNAGAK AGLTFLVDLI KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK VNYGVTVLPT FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN AASGRQTVDA ALAAAQTNAA AASEFSSNNN NNNNNNNLGI EGR
Q44-httEx1	13819.3	MATLEKL MKAFESLKSF QQQQQQQQQQ QQQQQQQQQQ QQQQQQQQQQ QQQQQQQQQQ QQQQPPPPPP PPPPPQLPQP PPQAQPLLPQ PQPPPPPPPP PPGPAVAEEP LHRPSGSHHH HHH
ΔN6-Q44-httEx1	13145.4	L MKAFESLKSF QQQQQQQQQQ QQQQQQQQQQ QQQQQQQQQQ QQQQQQQQQQ QQQQPPPPPP PPPPPQLPQP PPQAQPLLPQ PQPPPPPPPP PPGPAVAEEP LHRPSGSHHH HHH
ΔN9-Q44-httEx1	12772.9	AFESLKSF QQQQQQQQQQ QQQQQQQQQQ QQQQQQQQQQ QQQQQQQQQQ QQQQPPPPPP PPPPPQLPQP PPQAQPLLPQ PQPPPPPPPP PPGPAVAEEP LHRPSGSHHH HHH

Name: Residues numbered relative to Q44-httEx1	**MW (Da)**	**Sequence**
MBP(D[-363]-K[-256])	12115.9	`DTGIKVTVEH PDKLEEKFPQ VAATGDGPDI IFWAHDRFGG YAQSGLLAEI TPDKAFQDKL` `YPFTWDAVRY NGKLIAYPIA VEALSLIYNK DLLPNPPKTW EEIPALDK`
ΔN15-Q44-httEx1	12097.1	`SF QQQQQQQQQQ QQQQQQQQQQ QQQQQQQQQQ QQQQQQQQQQ QQQQPPPPPP` `PPPPPQLPQP PPQAQPLLPQ PQPPPPPPPP PPGPAVAEEP LHRPSGSHHH HHH`
MBP(D[-363]-R[-295])	7721.7	`DTGIKVTVEH PDKLEEKFPQ VAATGDGPDI IFWAHDRFGG YAQSGLLAEI TPDKAFQDKL` `YPFTWDAVR`

**MBP(Y[-294]-[-1]) and MBP(E[-255]-R[-1]) were observed by SDS-PAGE only.*

Table A.2. Observed trypsin cleavage products of MBP-Q44-httEx1 by ESI-TOF MS prior to aggregation of the ΔN15-Q44-httEx1 monomer.

Cleavage product	**Frequency observed (n=7)**	**Trial 1**	**Trial 2**	**Trial 3**	**Trial 4**	**Trial 5**	**Trial 6**	**Trial 7**
Free MBP	1.00	X	X	X	X	X	X	X
MBP(I[-391]-R[-1])	0.29	X	X				X	
Q44-httEx1	0.29					X	X	
ΔN6-Q44-httEx1	0.29					X	X	
ΔN9-Q44-httEx1	0.43					X	X	X
MBP(D[-363]-K[-256])	0.57	X	X	X				X
ΔN15-Q44-httEx1	1.00	X	X	X	X	X	X	X
MBP(D[-363]-R[-295])	0.57	X	X	X				X

Table A.3. Observed trypsin cleavage products of MBP-Q44-httEx1 by ESI-TOF MS following aggregation of ΔN15-Q44-httEx1 into mature fibrils.

Cleavage product	**Frequency observed (n=4)**	**Trial 1**	**Trial 2**	**Trial 3**	**Trial 4**
Free MBP	1.00	X	X	X	X
MBP(I[-391]-R[-1])	0.75	X	X	X	
Q44-httEx1	0				
ΔN6-Q44-httEx1	0				
ΔN9-Q44-httEx1	0				
MBP(D[-363]-K[-256])	0				
ΔN15-Q44-httEx1	0				
MBP(D[-363]-R[-295])	0.25	X			

**Separate trials than indicated in Table A.2.*

Table A.4. Experimental conditions for MAS ssNMR experiments.

Abbreviations: NS, number of scans per t_1 point; MAS, magic angle spinning rate; RD, recycle delay; TPPM, ^{1}H decoupling power during evolution and acquisition using the two-pulse phase modulation scheme; t_1 evol., t_1 evolution time = (real and imaginary t_1 points) X t_1 increment time; Mixing, ^{13}C-^{13}C or ^{1}H-^{1}H mixing time (ms).

Fig.	**Sample**	**Experiment**	**NS**	**Temp. (K)**	**MAS (kHz)**	**RD (s)**	**TPPM (kHz)**	**t_1 evol. (μs)**	**Mixing (ms)**
Figure 3.8	Q44-httEx1, 22 °C [3]	^{1}H-^{13}C CP	1024	275	13	2.8	83	NA	NA
Figure 3.8	Q44-httEx1, 22 °C [3]	^{1}H-^{13}C INEPT	1024	275	10	2.8	83	NA	NA
Figure 3.8	ΔN15-Q44-httEx1, 22 °C	^{1}H-^{13}C CP	1024	277	12.5	2.8	83	NA	NA
Figure 3.8	ΔN15-Q44-httEx1, 22 °C	^{1}H-^{13}C INEPT	4096	277	14	3.0	83	NA	NA
Figure 3.8	ΔN15-Q44-httEx1, 22 °C	2D ^{13}C-^{13}C INEPT-TOBSY	64	277	8.333	2.8	83	256 * 35.99	25
Figure 3.8	Q44-httEx1, 22 °C [3]	2D ^{13}C-^{13}C INEPT-TOBSY	96	275	8.333	2.6	83	388 * 83.33	6
Figure 3.8	ΔN15-Q44-httEx1, 22 °C	2D ^{13}C-^{13}C DARR	64	277	14	2.8	83	256 * 31.10	8
Figure 3.8	Q44-httEx1, 22 °C [3]	2D ^{13}C-^{13}C DARR	325	275	13	2.8	83	724 * 27.6	15
Figure A.7	Q44-httEx1, 22 °C [3]	2D ^{13}C-^{13}C PDSD	96	275	13	2.8	83	526 * 33.11	1, 15, 50, 100, 250, 500
Figure A.6	ΔN15-Q44-httEx1, 22 °C	2D ^{13}C-^{13}C PDSD	32	277	13	2.8	83	256 * 35.99	250
Figure A.7	ΔN15-Q44-httEx1, 22 °C	2D ^{13}C-^{13}C PDSD	128	277	13	2.8	83	256 * 35.99	0, 25, 250, 500

APPENDIX B

B.1. SUPPORTING INFORMATION FOR CHAPTER 4

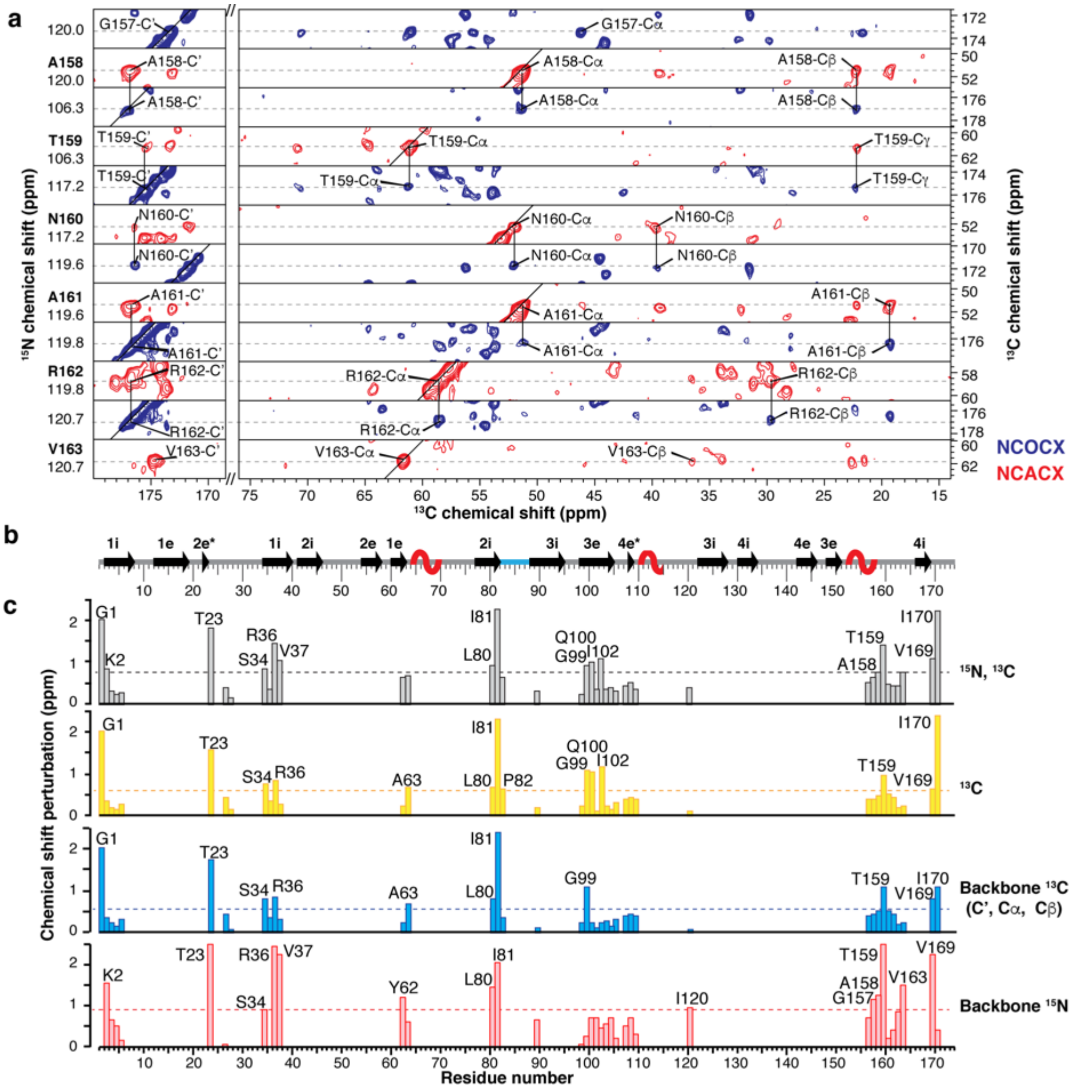


Figure B.1. MAS ssNMR chemical shift assignments of amorphous P23T hγD aggregates formed at physiological pH

(a) Sequential backbone walk for residues A158-V163 based on 3D NCOCX (navy) and 3D NCACX (red) spectra. (b) Secondary structure content of the solution-state structure. The β-strands (arrows) and single-residue β-bridges (*) are labeled as internal (i) or external (e) within Greek key motifs 1-4, while α-helices are marked as red helices. The linker between the NTD and CTD is shown in blue. (c) Chemical shift perturbation (CSP) of P23T hγD upon formation of amorphous aggregates, based on the difference between the assigned ssNMR shifts (Table B.2) and the P23T hγD solution NMR chemical shifts. Bars indicate the combined RMSD of ^{13}*C and* ^{15}*N atoms,* ^{13}*C, backbone* ^{13}*C (C', Cα and Cβ) and backbone* ^{15}*N, respectively. The average RMSDs of 0.75, 0.55, 0.62, and 0.91 are plotted as dashed lines. No* ^{15}*N solution NMR shifts were available for G1, P27, and P82. Reprinted from:* ***Boatz, J.C.****, Whitley M.J., Li M, Gronenborn A.M., and van der Wel, P.C.A. Cataract-associated P23T γD-crystallin retains a native-like fold in amorphous-looking aggregates formed at physiological pH, Nature Communications 8, 15137. Copyright (2017) Boatz, J. C., et. al. under Creative Commons Attribution 4.0 International License.*

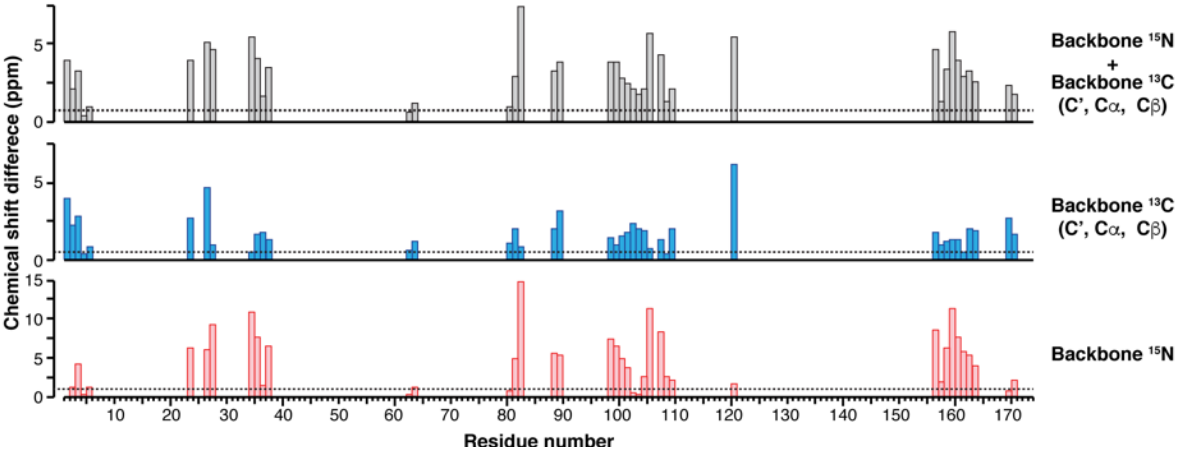


Figure B.2. MAS ssNMR chemical shift differences between amorphous P23T hγD aggregates and the simulated amyloid-like model

Bars show the chemical shift differences between the experimental chemical shifts of the amorphous P23T hγD aggregates Table B.2) and the chemical shifts predicted for our schematic amyloid-like model (Figure 4.8, a). Shown are the combined RMSD of the backbone ^{13}C and ^{15}N atoms, the backbone ^{13}C (C', Cα and Cβ) and backbone ^{15}N, respectively. The employed prediction algorithm does not produce the ^{13}C shifts of the entire side chain, leading us to omit the full-residue ^{13}C chemical shift deviations present in Figure B.1. Note the different scale of the vertical axes compared to Figure B.1, reflecting the much larger chemical shift deviation between the schematic amyloid-like model and the observed ssNMR spectra. For comparison, the average RMSDs from Figure B.1 are included as dashed lines. Reprinted from: ***Boatz, J.C.****, Whitley M.J., Li M, Gronenborn A.M., and van der Wel, P.C.A. Cataract-associated P23T γD-crystallin retains a native-like fold in amorphous-looking aggregates formed at physiological pH, Nature Communications 8, 15137. Copyright (2017) Boatz, J. C., et. al. under Creative Commons Attribution 4.0 International License.*

Table B.1. Detailed experimental conditions of the MAS ssNMR experiments on P23T hγD crystallin

Abbreviations: NS, number of scans per t_1 point; Temp., temperature of cooling gas; MAS, magic angle spinning rate; RD, recycle delay. TPPM, ^{1}H decoupling power during evolution and acquisition using the two-pulse phase modulation scheme was 83 kHz for all experiments.

Fig./ Table	**pH**	**Experiment**	**NS**	**Temp (K)**	**MAS (kHz)**	**RD (s)**	**t_1 evol. (μs)**	**t_2 evol. (μs)**	**Mix. (ms)**	**H-C/N, C-N CP time (ms)**	**τ_1, τ_2 (ms)**	**Exp. time (h)**
Figure 4.4, a	7	^{1}H-^{13}C CP	102	298	12.5	3.0	NA	NA	NA	2.0, NA	NA	0.8
Figure 4.4, b Figure 4.10, a,b Figure 4.7 Figure 4.9	7	2D ^{13}C-^{13}C CP-DARR	40	283	12.5	2.8	34.85* 860	NA	8.0	2.0, NA	NA	54
Figure 4.5, a	3	^{1}H-^{13}C CP	1024	298	12.5	3.0	NA	NA	NA	1.5, NA	NA	0.8
Figure 4.5, a Figure 4.10, a,b	3	2D ^{13}C-^{13}C CP-DARR	64	298	12.5	2.6	37.84* 792	NA	8.0	1.5, NA	NA	70
Figure 4.6, a	7	^{1}H-^{13}C INEPT	256	298	8.333	2.8	NA	NA	NA	NA, NA	1.5, 1.0	0.2
Figure 4.6, b	3	2D ^{13}C-^{13}C TOBSY	64	298	8.333	2.8	50.94* 400	NA	6.0	NA, NA	1.2, 1.0	20
Figure 4.9	7	3D NCACX	80	298	12.5	2.5	410.76 * 16	275.89 *24	15	2.0, 6.0	NA	87.0
Figure 4.9	7	3D NCOCX	80	298	12.5	2.5	410.76 * 16	275.89 *24	30	2.5, 5.5	NA	87.5
Table B.1	7	2D NCACX	112	298	10.0	3.0	100.00 * 100	NA	25	1.5, 6.0	NA	19.0

Table B.2. Chemical shifts of residues assigned in amorphous P23T hγD aggregates at physiological pH determined by MAS ssNMR

The reported ^{15}N and ^{13}C chemical shifts are referenced relative to liquid ammonia and dilute aqueous DSS, respectively (see methods), and have an experimental error of 0.4 ppm for ^{13}C and 0.5 ppm for ^{15}N.

Res. Num.	Type	N	C'	Cα	Cβ	Cγ1	Cγ2	Cδ	Cε
1	Gly		169.5	43.6					
2	Lys	124.4	172.9	57.8	35.7	24.6		26.6	42.1
3	Ile	128.6	171.5	58.9	41.3	28.5	15.3	12.9	
4	Thr	122.5	172.5	61.8	69.2	20.9			
5	Leu	128.2	175.1	53.7	44.1	27.4			
23[a]	Thr	113.2	176.4	64.3	72.3	22.9			
26	Gln	119.5	174.0	59.5	28.2				
27	Pro	130.6	175.7	64.6	30.4	28.0		50.7	
34	Ser	107.9	173.6	58.8	64.4				
35	Ala	118.9	175.4	52.7	22.8				
36	Arg	123.8	175.5	54.9	33.9				
37	Val	131.8	175.5	62.2	31.0	23.2			
62[b]	Tyr	122.6	174.4	56.3*	39.8*			132.2	
63	Ala	126.5	175.7	53.8	19.3				
80	Leu	126.0	175.7	53.6	41.8	26.8			
81	Ile	130.3	173.8	60.0	42.2	27.7	18.6		
82	Pro	138.1	177.1	63.6	32.3	26.4		51.3	
89	Ile	130.0	170.7	59.1	42.1	28.1	15.7*	14.9	
98	Arg	117.7	175.2	54.4	33.1	27.5			
99	Gly	105.2	172.5	44.0					
100[b]	Gln	119.6	174.2	57.7*	29.8*	32.6			
101	Met	120.0	175.3	53.4	36.1				
102	Ile	124.0	171.0	60.4	41.4	28.2	15.5	12.9	
103	Glu	127.4	175.3	53.7	33.8	37.5		183.6	
104	Phe	120.8	176.1	57.6	43.5				
105[b]	Thr	108.8	173.3	61.0*	70.5*	21.5			
107	Asp	115.8	176.4	55.5	41.4				
108	Cys	121.3	174.2	57.8	28.7				
109[b]	Ser	124.5	173.9	60.9*	62.8*				
120[c]	Ile	122.7	173.9	57.9	38.7	26.9*	19.4	11.5*	
156	Trp	114.9	176.6	57.1	28.2				
157	Gly	108.0	173.4	46.1					
158	Ala	120.0	177.0	51.4	22.2				
159	Thr	106.3	175.4	61.3	70.7	22.2			
160	Asn	117.2	171.7	52.1	39.5	176.5			

Res. Num.	**Type**	**N**	**C'**	**Cα**	**Cβ**	**Cγ1**	**Cγ2**	**Cδ**	**Cε**
161	Ala	119.6	176.4	51.4	19.3				
162	Arg	119.8	176.9	58.6	29.6				
163	Val	120.7	174.4	61.7	36.4				
169	Val	124.3	174.7	65.2	30.0	21.8	20.7		
170	Ile	130.1	173.9	58.1	40.4	25.3	20.2	17.0	

[a] *The mutant residue T23 was assigned to these chemical shifts on the basis that the three other (out of four) Thr residues were unambiguously assigned in the 2D and 3D ssNMR spectra.*

[b] *Atoms labeled with asterisk (*) are tentatively assigned based on close correspondence to solution NMR shifts.*

[c] *Residue I120 was assigned to these chemical shifts on the basis that the other five (out of six) Ile residues were unambiguously assigned in the 2D and 3D ssNMR spectra. This tentative assignment finds further support in the observation that the Cδ and Cγ1 chemical shifts of this spin system match the distinct and characteristic pattern that sets I120 apart from the other Ile residues in the solution NMR data.*

APPENDIX C

C.1. SUPPORTING INFORMATION FOR CHAPTER 5

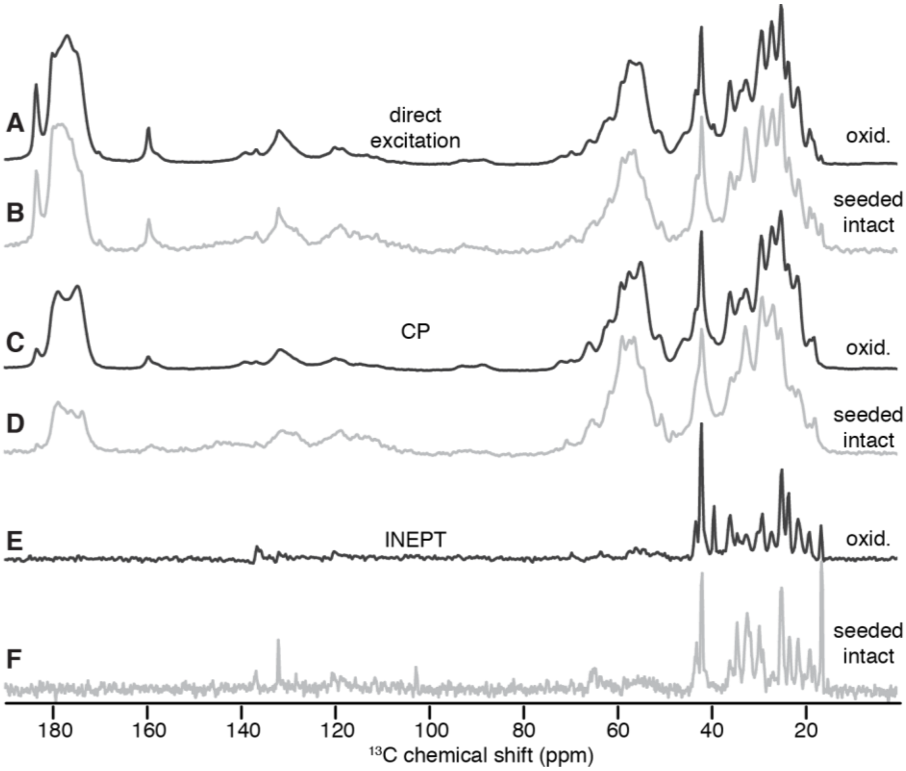


Figure C.1. Full 1D MAS ssNMR spectra of [U-^{13}C,^{15}N] apoA-I

*(A-B) direct excitation 13C, (C-D) ^{1}H-^{13}C CP, and (E-F) ^{1}H-^{13}C refocused INEPT spectra. Aggregates of H_2O_2-wt-^{13}C,^{15}N-ApoA-I and aggregates of non-oxidized wt-^{13}C,^{15}N-ApoA-I seeded with unlabeled pre-aggregated H_2O_2-wt-ApoA-I are reported in black (A, C, E) and gray (B, D, F), respectively. Reprinted with permission from Witkowski, A., Chan, G. K. L, **Boatz, J. C.**, Li, N. J., Wong, J. C., van der Wel, P. C. A., Cavigiolio, G. Methionine oxidized apolipoprotein A-I at the crossroads of HDL biogenesis and amyloid formation, The FASEB Journal 32(6):3149-3165. Copyright (2018) Witkowski, A., et. al.*

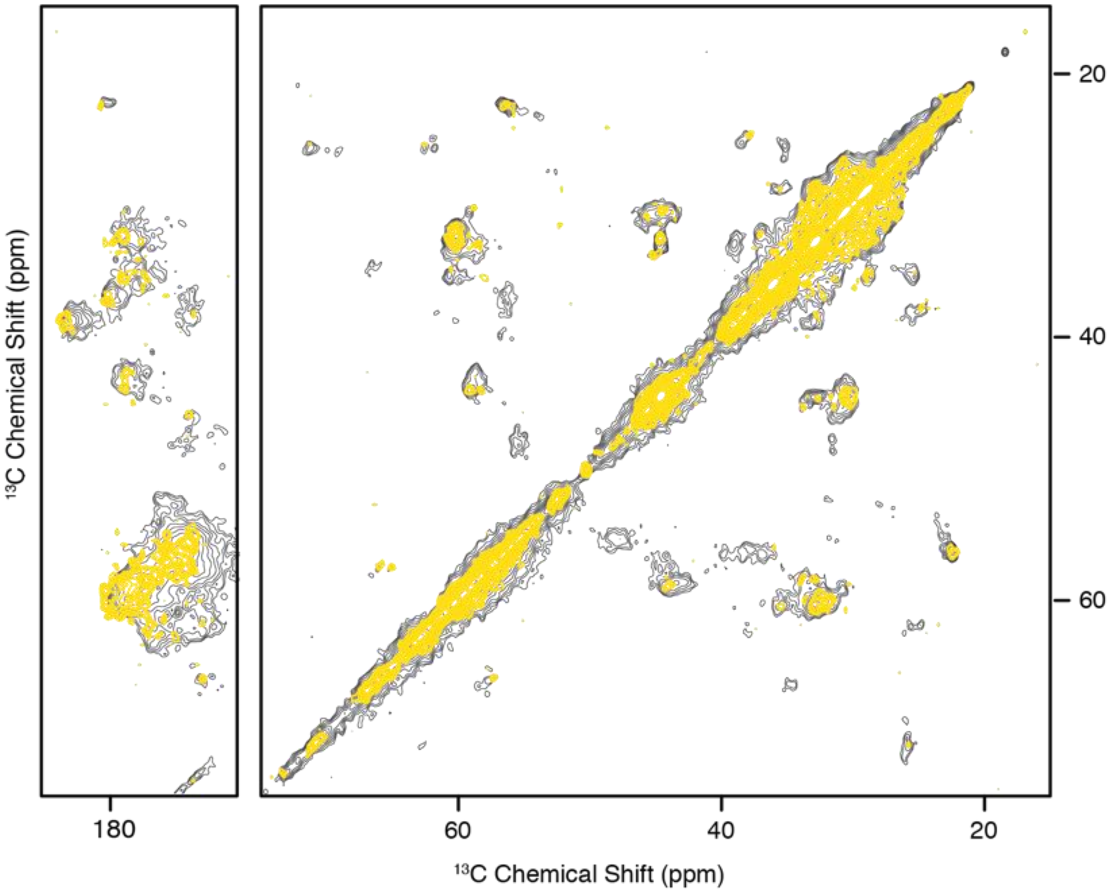


Figure C.2. MAS ssNMR spectral comparison between a replicate H_2O_2 wt-^{13}C,^{15}N-ApoA-I sample and the seeded intact wt-^{13}C,^{15}N-ApoA-I aggregates

2D ^{13}C-^{13}C CP-based DARR MAS ssNMR spectra of H_2O_2 wt-^{13}C,^{15}N-ApoA-I (gray) and the seeded intact wt-^{13}C,^{15}N-ApoA-I aggregates (yellow). The spectrum of H_2O_2 wt-^{13}C,^{15}N-ApoA-I was obtained from a separate sample than was used to acquire the 2D spectrum shown in Figure 5.5, e and Figure 5.6.

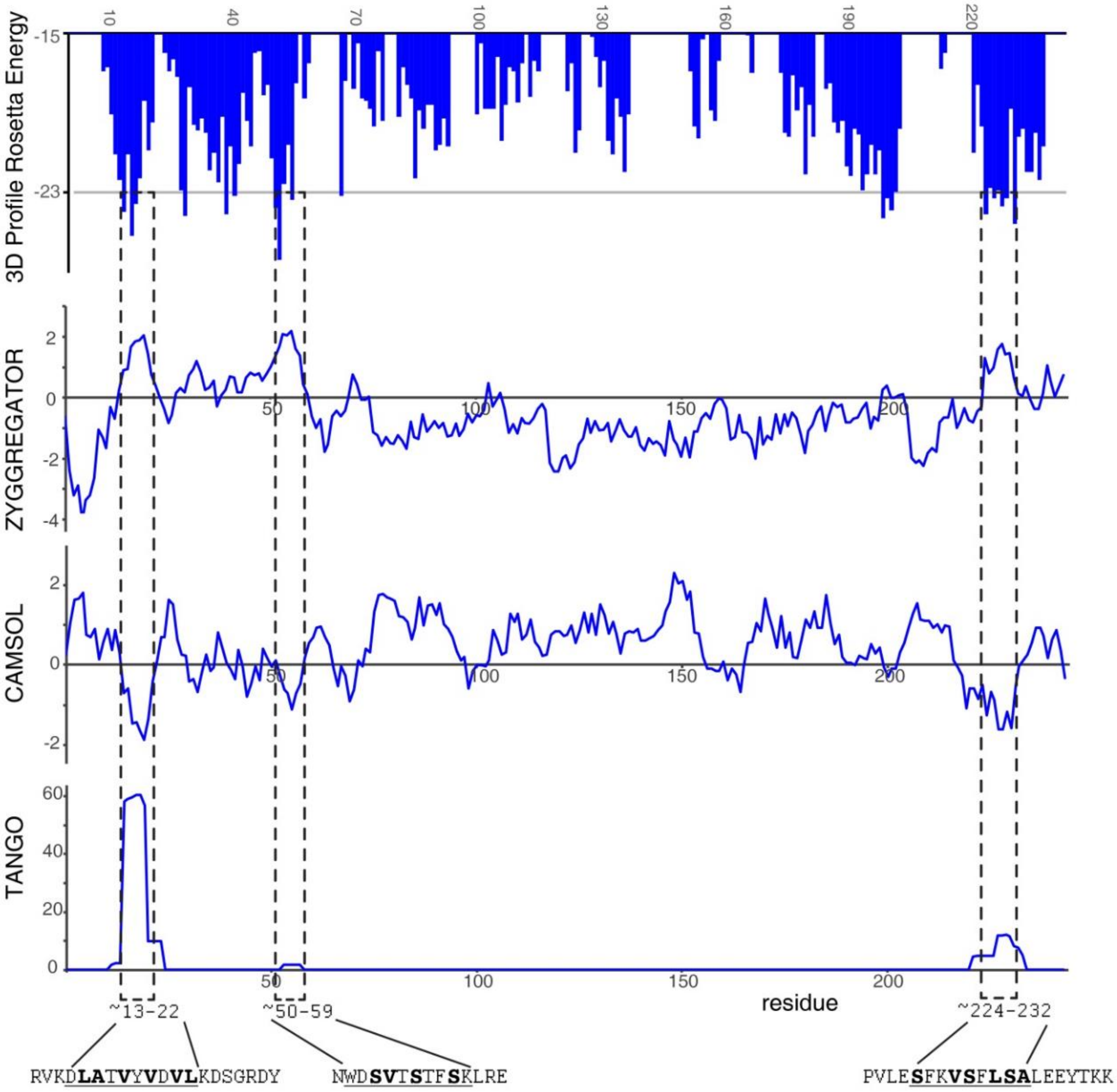


Figure C.3. Amyloid-propensity prediction results.

The primary sequence of apoA-I was analyzed for amyloid and aggregation propensity using four algorithms: 3D Profile (via ZipperDB), Zyggregator, CamSol, and Tango [321-324]. The sequence areas that appear to be consensus elements with high predicted amyloidogenicity are indicated with dashed boxes. The approximate sequence positions (underlined residues), with their primary sequence context, are indicated at the bottom. In the sequence, b-sheet residues detected by MAS ssNMR are shown in bold. Reprinted with permission from Witkowski, A., Chan, G. K. L, ***Boatz, J. C.****, Li, N. J., Wong, J. C., van der Wel, P. C. A., Cavigiolio, G. Methionine oxidized apolipoprotein A-I at the crossroads of HDL biogenesis and amyloid formation, The FASEB Journal 32(6):3149-3165. Copyright (2018) Witkowski, A., et. al.*

BIBLIOGRAPHY

1. Hoop, C.L., et al., *Huntingtin exon 1 fibrils feature an interdigitated β-hairpin-based polyglutamine core.* Proceedings of the National Academy of Sciences, 2016. **113**(6): p. 1546-1551.
2. Merg, A.D., et al., *Peptide-Directed Assembly of Single-Helical Gold Nanoparticle Superstructures Exhibiting Intense Chiroptical Activity.* J Am Chem Soc, 2016. **138**(41): p. 13655-13663.
3. Lin, H.K., et al., *Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core.* Nat Commun, 2017. **8**: p. 15462.
4. Mandal, A., et al., *On the use of ultracentrifugal devices for routine sample preparation in biomolecular magic-angle-spinning NMR.* Journal of Biomolecular NMR, 2017. **67**(3): p. 165-178.
5. Boatz, J.C., et al., *Cataract-associated P23T gammaD-crystallin retains a native-like fold in amorphous-looking aggregates formed at physiological pH.* Nat Commun, 2017. **8**: p. 15137.
6. Witkowski, A., et al., *Methionine oxidized apolipoprotein A-I at the crossroads of HDL biogenesis and amyloid formation.* FASEB J, 2018. **32**(6): p. 3149-3165.
7. Smith, A.N., et al., *Structural Fingerprinting of Protein Aggregates by Dynamic Nuclear Polarization-Enhanced Solid-State NMR at Natural Isotopic Abundance.* J Am Chem Soc, 2018. **140**(44): p. 14576-14580.
8. Chiti, F. and C.M. Dobson, *Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade.* Annu Rev Biochem, 2017. **86**: p. 27-68.
9. Glenner, G.G., et al., *Amyloid fibril proteins: proof of homology with immunoglobulin light chains by sequence analyses.* Science, 1971. **172**(3988): p. 1150-1.

10. Eulitz, M., D.T. Weiss, and A. Solomon, *Immunoglobulin heavy-chain-associated amyloidosis.* Proc Natl Acad Sci U S A, 1990. **87**(17): p. 6542-6.
11. Bhowmick, D.C., et al., *The Molecular Physiopathogenesis of Islet Amyloidosis.* Handb Exp Pharmacol, 2018. **245**: p. 271-312.
12. Das, M. and O. Gursky, *Amyloid-Forming Properties of Human Apolipoproteins: Sequence Analyses and Structural Insights.* Adv Exp Med Biol, 2015. **855**: p. 175-211.
13. Samillan-Sosa Kdel, R., et al., *Hereditary apolipoprotein AI-associated renal amyloidosis: A diagnostic challenge.* Nefrologia, 2015. **35**(3): p. 322-7.
14. Zheng, L., et al., *Apolipoprotein A-I is a selective target for myeloperoxidase-catalyzed oxidation and functional impairment in subjects with cardiovascular disease.* J Clin Invest, 2004. **114**(4): p. 529-41.
15. Didonato, J.A., et al., *Site-specific Nitration of Apolipoprotein A-I at Tyrosine 166 Is Both Abundant within Human Atherosclerotic Plaque and Dysfunctional.* J Biol Chem, 2014. **289**(15): p. 10276-92.
16. Morton, J., et al., *Strikingly Different Atheroprotective Effects of Apolipoprotein A-I in Early- Versus Late-Stage Atherosclerosis.* JACC Basic Transl Sci, 2018. **3**(2): p. 187-199.
17. Magalingam, K.B., et al., *Current Concepts of Neurodegenerative Mechanisms in Alzheimer's Disease.* Biomed Res Int, 2018. **2018**: p. 3740461.
18. Saiki, S., S. Sato, and N. Hattori, *Molecular pathogenesis of Parkinson's disease: update.* J Neurol Neurosurg Psychiatry, 2012. **83**(4): p. 430-6.
19. Bates, G.P., et al., *Huntington disease.* Nat Rev Dis Primers, 2015. **1**: p. 15005.
20. Gajdusek, D.C., C.J. Gibbs, and M. Alpers, *Experimental transmission of a Kuru-like syndrome to chimpanzees.* Nature, 1966. **209**(5025): p. 794-6.
21. Hope, J., et al., *Fibrils from brains of cows with new cattle disease contain scrapie-associated protein.* Nature, 1988. **336**(6197): p. 390-2.
22. Prusiner, S.B., *Novel proteinaceous infectious particles cause scrapie.* Science, 1982. **216**(4542): p. 136-44.
23. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis.* Journal of computational chemistry, 2004. **25**(13): p. 1605-1612.

24. Jung, J., et al., *The structure of the cataract-causing P23T mutant of human gammaD-crystallin exhibits distinctive local conformational and dynamic changes.* Biochemistry, 2009. **48**(12): p. 2597-2609.
25. Uversky, V.N., *Natively unfolded proteins: a point where biology waits for physics.* Protein Sci, 2002. **11**(4): p. 739-56.
26. Nilsson, B. and S. Anderson, *Proper and improper folding of proteins in the cellular environment.* Annu Rev Microbiol, 1991. **45**: p. 607-35.
27. Kanagal-Shamanna, R., et al., *Crystal-storing histiocytosis: a clinicopathological study of 13 cases.* Histopathology, 2016. **68**(4): p. 482-91.
28. Yu, X.J., et al., *Monoclonal light chain crystalline podocytopathy and tubulopathy associated with monoclonal gammopathy of renal significance: a case report and literature review.* BMC Nephrol, 2018. **19**(1): p. 322.
29. Mulay, S.R., et al., *Novel Insights into Crystal-Induced Kidney Injury.* Kidney Dis (Basel), 2018. **4**(2): p. 49-57.
30. Abe, N., et al., *Crystalglobulinemia manifesting as chronic arthralgia and acute limb ischemia: A clinical case report.* Medicine (Baltimore), 2017. **96**(16): p. e6643.
31. Berger, J. and N. Hinglais, *Intercapillary deposits of IgA-IgG (Article in French).* J Urol Nephrol (Paris), 1968. **74**(9): p. 694-5.
32. Jimenez-Zepeda, V.H., *Light chain deposition disease: novel biological insights and treatment advances.* Int J Lab Hematol, 2012. **34**(4): p. 347-55.
33. Stuart, M.J. and R.L. Nagel, *Sickle-cell disease.* Lancet, 2004. **364**(9442): p. 1343-60.
34. Gronenborn, A.M., *Protein acrobatics in pairs--dimerization via domain swapping.* Curr Opin Struct Biol, 2009. **19**(1): p. 39-49.
35. Bennett, M.J., M.P. Schlunegger, and D. Eisenberg, *3D domain swapping: a mechanism for oligomer assembly.* Protein Sci, 1995. **4**(12): p. 2455-68.
36. Greenwald, J. and R. Riek, *Biology of amyloid: structure, function, and regulation.* Structure, 2010. **18**(10): p. 1244-60.
37. van der Wel, P.C.A., *Insights into protein misfolding and aggregation enabled by solid-state NMR spectroscopy.* Solid State Nuclear Magnetic Resonance, 2017. **88**: p. 1-14.
38. Tycko, R., *Amyloid polymorphism: structural basis and neurobiological relevance.* Neuron, 2015. **86**(3): p. 632-45.

39. Korsak, M. and T. Kozyreva, *Beta Amyloid Hallmarks: From Intrinsically Disordered Proteins to Alzheimer's Disease.* Adv Exp Med Biol, 2015. **870**: p. 401-21.
40. Morris, G.P., I.A. Clark, and B. Vissel, *Inconsistencies and controversies surrounding the amyloid hypothesis of Alzheimer's disease.* Acta Neuropathol Commun, 2014. **2**: p. 135.
41. Paravastu, A.K., et al., *Molecular structural basis for polymorphism in Alzheimer's beta-amyloid fibrils.* Proc Natl Acad Sci U S A, 2008. **105**(47): p. 18349-54.
42. Petkova, A.T., W.M. Yau, and R. Tycko, *Experimental constraints on quaternary structure in Alzheimer's beta-amyloid fibrils.* Biochemistry, 2006. **45**(2): p. 498-512.
43. Bertini, I., et al., *A new structural model of Abeta40 fibrils.* J Am Chem Soc, 2011. **133**(40): p. 16013-22.
44. Lu, J.X., et al., *Molecular structure of beta-amyloid fibrils in Alzheimer's disease brain tissue.* Cell, 2013. **154**(6): p. 1257-68.
45. Colvin, M.T., et al., *Atomic Resolution Structure of Monomorphic Abeta42 Amyloid Fibrils.* J Am Chem Soc, 2016. **138**(30): p. 9663-74.
46. Kodali, R., et al., *Abeta(1-40) forms five distinct amyloid structures whose beta-sheet contents and fibril stabilities are correlated.* J Mol Biol, 2010. **401**(3): p. 503-17.
47. Nekooki-Machida, Y., et al., *Distinct conformations of in vitro and in vivo amyloids of huntingtin-exon1 show different cytotoxicity.* Proc Natl Acad Sci U S A, 2009. **106**(24): p. 9679-84.
48. Morris, A.M., M.A. Watzky, and R.G. Finke, *Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature.* Biochim Biophys Acta, 2009. **1794**(3): p. 375-97.
49. Arosio, P., T.P.J. Knowles, and S. Linse, *On the lag phase in amyloid fibril formation.* Phys Chem Chem Phys, 2015. **17**(12): p. 7606-7618.
50. Jarrett, J.T., E.P. Berger, and P.T. Lansbury, Jr., *The carboxy terminus of the beta amyloid protein is critical for the seeding of amyloid formation: implications for the pathogenesis of Alzheimer's disease.* Biochemistry, 1993. **32**(18): p. 4693-7.
51. Bhattacharyya, A.M., A.K. Thakur, and R. Wetzel, *polyglutamine aggregation nucleation: thermodynamics of a highly unfavorable protein folding reaction.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15400-5.

52. Bellesia, G. and J.E. Shea, *Diversity of kinetic pathways in amyloid fibril formation.* J Chem Phys, 2009. **131**(11): p. 111102.

53. Fu, Z., et al., *Mechanism of Nucleated Conformational Conversion of Abeta42.* Biochemistry, 2015. **54**(27): p. 4197-207.

54. Thakur, A.K., et al., *Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism.* Nat Struct Mol Biol, 2009. **16**(4): p. 380-9.

55. Woodard, D., et al., *Gel formation in protein amyloid aggregation: a physical mechanism for cytotoxicity.* PLoS One, 2014. **9**(4): p. e94789.

56. Shen, K., et al., *Control of the structural landscape and neuronal proteotoxicity of mutant Huntingtin by domains flanking the polyQ tract.* Elife, 2016. **5**: p. e18065.

57. Yamaguchi, K., et al., *Seeding-dependent propagation and maturation of amyloid fibril conformation.* J Mol Biol, 2005. **352**(4): p. 952-60.

58. Gillam, J.E. and C.E. MacPhee, *Modelling amyloid fibril formation kinetics: mechanisms of nucleation and growth.* J Phys Condens Matter, 2013. **25**(37): p. 373101.

59. Watzky, M.A., et al., *Fitting yeast and mammalian prion aggregation kinetic data with the Finke-Watzky two-step model of nucleation and autocatalytic growth.* Biochemistry, 2008. **47**(40): p. 10790-800.

60. Serio, T.R., et al., *Nucleated conformational conversion and the replication of conformational information by a prion determinant.* Science, 2000. **289**(5483): p. 1317-21.

61. Wolfe, L.S., et al., *Protein-induced photophysical changes to the amyloid indicator dye thioflavin T.* Proc Natl Acad Sci U S A, 2010. **107**(39): p. 16863-8.

62. Sahoo, B., et al., *Aggregation Behavior of Chemically Synthesized, Full-Length Huntingtin Exon1.* Biochemistry, 2014. **53**(24): p. 3897-3907.

63. 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 U S A, 1999. **96**(8): p. 4604-9.

64. Smith, M.A., et al., *Mutation of interfaces in domain-swapped human betaB2-crystallin.* Protein Sci, 2007. **16**(4): p. 615-25.

65. Das, P., J.A. King, and R. Zhou, *Aggregation of γ-crystallins associated with human cataracts via domain swapping at the C-terminal β-strands.* Proc. Natl. Acad. Sci. U.S.A., 2011. **108**(26): p. 10514-10519.

66. Garcia-Manyes, S., et al., *Single molecule force spectroscopy predicts a misfolded, domain-swapped conformation in human γD-crystallin.* J Biol Chem, 2016. **291**(8): p. 4226-4235.

67. Serebryany, E., et al., *Aggregation of Trp>Glu Point Mutants of Human Gamma-D Crystallin Provides a Model for Hereditary or UV-Induced Cataract.* Protein Science, 2016. **25**(6): p. 1115-1128.

68. Xi, Z., M.J. Whitley, and A.M. Gronenborn, *Human betaB2-Crystallin Forms a Face-en-Face Dimer in Solution: An Integrated NMR and SAXS Study.* Structure, 2017. **25**(3): p. 496-505.

69. Pande, A., et al., *Increase in surface hydrophobicity of the cataract-associated P23T mutant of human gammaD-crystallin is responsible for its dramatically lower, retrograde solubility.* Biochemistry, 2010. **49**(29): p. 6122-6129.

70. Benedek, G.B., *Cataract as a protein condensation disease: the Proctor Lecture.* Investigative Ophthalmology & Visual Science, 1997. **38**(10): p. 1911-1921.

71. Meehan, S., et al., *Amyloid fibril formation by lens crystallin proteins and its implications for cataract formation.* J Biol Chem, 2004. **279**(5): p. 3413-3419.

72. Loewenstein, M.A. and F.A. Bettelheim, *Cold cataract formation in fish lenses.* Exp Eye Res, 1979. **28**(6): p. 651-63.

73. Siezen, R.J., et al., *Opacification of gamma-crystallin solutions from calf lens in relation to cold cataract formation.* Proc Natl Acad Sci U S A, 1985. **82**(6): p. 1701-5.

74. Lo, W.K., *Visualization of crystallin droplets associated with cold cataract formation in young intact rat lens.* Proc Natl Acad Sci U S A, 1989. **86**(24): p. 9926-30.

75. Bermudez, M.A., et al., *Time course of cold cataract development in anesthetized mice.* Curr Eye Res, 2011. **36**(3): p. 278-84.

76. Adachi, M., et al., *Aggregation-phase diagrams of beta2-microglobulin reveal temperature and salt effects on competitive formation of amyloids versus amorphous aggregates.* J Biol Chem, 2018. **293**(38): p. 14775-14785.

77. Verma, M., A. Vats, and V. Taneja, *Toxic species in amyloid disorders: Oligomers or mature fibrils.* Ann Indian Acad Neurol, 2015. **18**(2): p. 138-45.

78. Kumar, J., R. Namsechi, and V.L. Sim, *Structure-Based Peptide Design to Modulate Amyloid Beta Aggregation and Reduce Cytotoxicity.* PLoS One, 2015. **10**(6): p. e0129087.

79. Oskarsson, M.E., et al., *BRICHOS domain of Bri2 inhibits islet amyloid polypeptide (IAPP) fibril formation and toxicity in human beta cells.* Proc Natl Acad Sci U S A, 2018. **115**(12): p. E2752-E2761.

80. Bissig, C., L. Rochin, and G. van Niel, *PMEL Amyloid Fibril Formation: The Bright Steps of Pigmentation.* Int J Mol Sci, 2016. **17**(9): p. 1438.

81. Fowler, D.M., et al., *Functional amyloid--from bacteria to humans.* Trends Biochem Sci, 2007. **32**(5): p. 217-24.

82. McGlinchey, R.P. and J.C. Lee, *Reversing the amyloid trend: Mechanism of fibril assembly and dissolution of the repeat domain from a human functional amyloid.* Isr J Chem, 2017. **57**(7-8): p. 613-621.

83. Maji, S.K., et al., *Functional amyloids as natural storage of peptide hormones in pituitary secretory granules.* Science, 2009. **325**(5938): p. 328-32.

84. Li, J., et al., *The RIP1/RIP3 necrosome forms a functional amyloid signaling complex required for programmed necrosis.* Cell, 2012. **150**(2): p. 339-50.

85. Moran, S.D., et al., *Two-dimensional IR spectroscopy and segmental ^{13}C labeling reveals the domain structure of human γD-crystallin amyloid fibrils.* Proc. Natl. Acad. Sci. U.S.A., 2012. **109**(9): p. 3329-3334.

86. Serebryany, E. and J.A. King, *The betagamma-crystallins: native state stability and pathways to aggregation.* Prog Biophys Mol Biol, 2014. **115**(1): p. 32-41.

87. Stradner, A., et al., *New insight into cataract formation: enhanced stability through mutual attraction.* Phys Rev Lett, 2007. **99**(19): p. 198103.

88. Annunziata, O., et al., *Oligomerization and phase transitions in aqueous solutions of native and truncated human beta B1-crystallin.* Biochemistry, 2005. **44**(4): p. 1316-28.

89. Wang, Y., et al., *Is Vascular Amyloidosis Intertwined with Arterial Aging, Hypertension and Atherosclerosis?* Front Genet, 2017. **8**: p. 126.

90. Schulte, J. and J.T. Littleton, *The biological function of the Huntingtin protein and its relevance to Huntington's Disease pathology.* Curr Trends Neurol, 2011. **5**: p. 65-78.

91. Sivanandam, V.N., et al., *The aggregation-enhancing huntingtin N-terminus is helical in amyloid fibrils.* J Am Chem Soc, 2011. **133**(12): p. 4558-4566.

92. Hoop, C.L., et al., *Polyglutamine amyloid core boundaries and flanking domain dynamics in huntingtin fragment fibrils determined by solid-state nuclear magnetic resonance.* Biochemistry, 2014. **53**(42): p. 6653-66.

93. Mishra, R., et al., *Serine phosphorylation suppresses huntingtin amyloid accumulation by altering protein aggregation properties.* J Mol Biol, 2012. **424**(1–2): p. 1-14.

94. Isas, J.M., R. Langen, and A.B. Siemer, *Solid-State Nuclear Magnetic Resonance on the Static and Dynamic Domains of Huntingtin Exon-1 Fibrils.* Biochemistry, 2015. **54**(25): p. 3942-9.

95. Myers, R.H., *Huntington's disease genetics.* NeuroRx, 2004. **1**(2): p. 255-62.

96. Chen, S., et al., *Amyloid-like features of polyglutamine aggregates and their assembly kinetics.* Biochemistry, 2002. **41**(23): p. 7391-9.

97. Chen, S., F.A. Ferrone, and R. Wetzel, *Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation.* Proc Natl Acad Sci U S A, 2002. **99**(18): p. 11884-9.

98. Lee, J.M., et al., *CAG repeat expansion in Huntington disease determines age at onset in a fully dominant fashion.* Neurology, 2012. **78**(10): p. 690-5.

99. Stricker-Shaver, J., et al., *Genetic Rodent Models of Huntington Disease.* Adv Exp Med Biol, 2018. **1049**: p. 29-57.

100. 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**(3): p. 493-506.

101. Pieri, L., et al., *Fibrillar α-Synuclein and Huntingtin Exon 1 Assemblies Are Toxic to the Cells.* Biophys. J., 2012. **102**(12): p. 2894-2905.

102. Scherzinger, E., et al., *Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo.* Cell, 1997. **90**(3): p. 549-558.

103. Sahl, S.J., et al., *Delayed emergence of subdiffraction-sized mutant huntingtin fibrils following inclusion body formation.* Quarterly reviews of biophysics, 2015. **49**(e2): p. 1-13.

104. Duim, W.C., et al., *Super-resolution fluorescence of huntingtin reveals growth of globular species into short fibers and coexistence of distinct aggregates.* ACS Chem Biol, 2014. **9**(12): p. 2767-78.

105. Sahl, S.J., et al., *Cellular Inclusion Bodies of Mutant Huntingtin Exon 1 Obscure Small Fibrillar Aggregate Species.* Scientific Reports, 2012. **2**: p. 895.

106. Jayaraman, M., et al., *Kinetically competing huntingtin aggregation pathways control amyloid polymorphism and properties.* Biochemistry, 2012. **51**(13): p. 2706-16.

107. Moreau, K.L. and J.A. King, *Protein misfolding and aggregation in cataract disease and prospects for prevention.* Trends Mol Med, 2012. **18**(5): p. 273-282.

108. Pascolini, D. and S.P. Mariotti, *Global estimates of visual impairment: 2010.* Br J Ophthalmol, 2012. **96**(5): p. 614-618.

109. Foster, A. and S. Resnikoff, *The impact of Vision 2020 on global blindness.* Eye (Lond), 2005. **19**(10): p. 1133-5.

110. Bloemendal, H., *The vertebrate eye lens.* Science, 1977. **197**(4299): p. 127-38.

111. Jaenicke, R. and C. Slingsby, *Lens crystallins and their microbial homologs: structure, stability, and function.* Crit Rev Biochem Mol Biol, 2001. **36**(5): p. 435-99.

112. Tikhomirova, T.S., O.M. Selivanova, and O.V. Galzitskaya, *alpha-Crystallins Are Small Heat Shock Proteins: Functional and Structural Properties.* Biochemistry (Mosc), 2017. **82**(2): p. 106-121.

113. Liu, Z., et al., *A Structural View of alphaB-crystallin Assembly and Amyloid Aggregation.* Protein Pept Lett, 2017. **24**(4): p. 315-321.

114. Tardieu, A., et al., *Protein interactions in the calf eye lens: interactions between beta-crystallins are repulsive whereas in gamma-crystallins they are attractive.* Eur Biophys J, 1992. **21**(1): p. 1-12.

115. Magid, A.D., A.K. Kenworthy, and T.J. McIntosh, *Colloid osmotic pressure of steer crystallins: implications for the origin of the refractive index gradient and transparency of the lens.* Exp Eye Res, 1992. **55**(4): p. 615-27.

116. Graw, J., *Genetics of crystallins: cataract and beyond.* Exp Eye Res, 2009. **88**(2): p. 173-189.

117. Xia, Z., et al., *UV-radiation induced disruption of dry-cavities in human gammaD-crystallin results in decreased stability and faster unfolding.* Sci Rep, 2013. **3**: p. 1560.

118. Jehle, S., et al., *Solid-state NMR and SAXS studies provide a structural basis for the activation of alphaB-crystallin oligomers.* Nat Struct Mol Biol, 2010. **17**(9): p. 1037-1042.

119. Jehle, S., et al., *N-terminal domain of alphaB-crystallin provides a conformational switch for multimerization and structural heterogeneity.* Proc Natl Acad Sci U S A, 2011. **108**(16): p. 6409-14.

120. Van Montfort, R.L., et al., *Crystal structure of truncated human betaB1-crystallin.* Protein Sci, 2003. **12**(11): p. 2606-12.

121. Basak, A., et al., *High-resolution X-ray crystal structures of human gammaD crystallin (1.25 A) and the R58H mutant (1.15 A) associated with aculeiform cataract.* J Mol Biol, 2003. **328**(5): p. 1137-1147.

122. Santana, A. and M. Waiswo, *The genetic and molecular basis of congenital cataract.* Arq Bras Oftalmol, 2011. **74**(2): p. 136-42.

123. Nandrot, E., et al., *Gamma-D crystallin gene (CRYGD) mutation causes autosomal dominant congenital cerulean cataracts.* J Med Genet, 2003. **40**(4): p. 262-7.

124. Vanita, V. and D. Singh, *A missense mutation in CRYGD linked with autosomal dominant congenital cataract of aculeiform type.* Mol Cell Biochem, 2012. **368**(1-2): p. 167-72.

125. Mackay, D.S., U.P. Andley, and A. Shiels, *A missense mutation in the gammaD crystallin gene (CRYGD) associated with autosomal dominant "coral-like" cataract linked to chromosome 2q.* Mol Vis, 2004. **10**: p. 155-62.

126. Xu, W.Z., et al., *[Localization and screening of autosomal dominant coralliform cataract associated gene]*. Zhonghua Yi Xue Yi Chuan Xue Za Zhi, 2004. **21**(1): p. 19-22.

127. Shentu, X., et al., *Special fasciculiform cataract caused by a mutation in the gammaD-crystallin gene.* Mol Vis, 2004. **10**: p. 233-9.

128. Santhiya, S.T., et al., *Novel mutations in the gamma-crystallin genes cause autosomal dominant congenital cataracts.* J Med Genet, 2002. **39**(5): p. 352-8.

129. Burdon, K.P., et al., *Investigation of crystallin genes in familial cataract, and report of two disease associated mutations.* Br J Ophthalmol, 2004. **88**(1): p. 79-83.

130. Ji, F., et al., *Crystal structure of the cataract-causing P23T γD-crystallin mutant.* Proteins, 2013. **81**(9): p. 1493-1498.

131. Schrodinger, LLC, *The AxPyMOL Molecular Graphics Plugin for Microsoft PowerPoint, Version 1.8*, 2015.

132. Mucchiano, G., G.G. Cornwell, 3rd, and P. Westermark, *Senile aortic amyloid. Evidence for two distinct forms of localized deposits.* Am J Pathol, 1992. **140**(4): p. 871-7.

133. Yamagishi, M., et al., *Coronary reactivity to nitroglycerin: intravascular ultrasound evidence for the importance of plaque distribution.* J Am Coll Cardiol, 1995. **25**(1): p. 224-30.

134. Sekijima, Y., *Transthyretin-type cerebral amyloid angiopathy: a serious complication in post-transplant patients with familial amyloid polyneuropathy.* J Neurol Neurosurg Psychiatry, 2015. **86**(2): p. 124.

135. Stewart, C.R., et al., *Serum amyloid P colocalizes with apolipoproteins in human atheroma: functional implications.* J Lipid Res, 2007. **48**(10): p. 2162-71.

136. Audemard, A., et al., *AL amyloidosis with temporal artery involvement simulates giant-cell arteritis.* Joint Bone Spine, 2012. **79**(2): p. 195-7.

137. Davies, H.A., M.M. Phelan, and J. Madine, *1H, 15N and 13C assignment of the amyloidogenic protein medin using fast-pulsing NMR techniques.* Biomol NMR Assign, 2016. **10**(1): p. 75-7.

138. Peng, S., J. Glennert, and P. Westermark, *Medin-amyloid: a recently characterized age-associated arterial amyloid form affects mainly arteries in the upper part of the body.* Amyloid, 2005. **12**(2): p. 96-102.

139. Westermark, G.T., et al., *Noncerebral Amyloidoses: Aspects on Seeding, Cross-Seeding, and Transmission.* Cold Spring Harb Perspect Med, 2017. **8**(1): p. a024323.

140. Mucchiano, G.I., et al., *Apolipoprotein A-1-derived amyloid in atherosclerotic plaques of the human aorta*. Journal of Pathology, 2001. **193**(2): p. 270-5.

141. Mucchiano, G.I., et al., *Apolipoprotein A-I-derived amyloid in atherosclerosis. Its association with plasma levels of apolipoprotein A-I and cholesterol.* American journal of clinical pathology, 2001. **115**(2): p. 298-303.

142. Kristen, A.V., et al., *High prevalence of amyloid in 150 surgically removed heart valves--a comparison of histological and clinical data reveals a correlation to atheroinflammatory conditions.* Cardiovasc Pathol, 2010. **19**(4): p. 228-35.

143. Gordon, D.J., et al., *High-density lipoprotein cholesterol and cardiovascular disease. Four prospective American studies.* Circulation, 1989. **79**(1): p. 8-15.

144. Loavenbruck, A.J., et al., *Mass spectrometry analysis reveals non-mutated apolipoprotein A1 lumbosacral radiculoplexus amyloidoma.* Muscle and Nerve, 2012. **46**(5): p. 817-22.

145. Loquet, A., et al., *3D structure determination of amyloid fibrils using solid-state NMR spectroscopy*. Methods, 2018. **138-139**: p. 26-38.

146. Kar, K., et al., *beta-hairpin-mediated nucleation of polyglutamine amyloid formation.* J Mol Biol, 2013. **425**(7): p. 1183-97.

147. Tycko, R. and R.B. Wickner, *Molecular structures of amyloid and prion fibrils: consensus versus controversy.* Acc Chem Res, 2013. **46**(7): p. 1487-96.

148. Sezonov, G., D. Joseleau-Petit, and R. D'Ari, *Escherichia coli physiology in Luria-Bertani broth.* J Bacteriol, 2007. **189**(23): p. 8746-9.

149. Boatz, J.C., et al., *Cataract-associated P23T γD-crystallin retains a native-like fold in amorphous-looking aggregates formed at physiological pH.* Nature Communications, 2017. **8**: p. 15137.

150. Andrew, E.R., Bradbury, A., Eades, R. G., *Removal of Dipolar Broadening of Nuclear Magnetic Resonance Spectra of Solids by Specimen Rotation.* Nature, 1959. **183**(4678): p. 1802-1803.

151. Duer, M.J., *Solid-state NMR spectroscopy principles and applications*, 2002, Blackwell Science: Malden, MA. p. xvii, 567 p.

152. Keeler, J., *Understanding NMR spectroscopy*. 2nd ed2010, Chichester, U.K.: John Wiley and Sons. xiii, 511 p.

153. Metz, G., Wu, X. L., Smith, S. O., *Ramped-amplitude cross polarization in magic-angle-spinning NMR.* J. Magn. Reson. Ser. A, 1994. **110**(2): p. 219-227.

154. Hodgkinson, P. and L. Emsley, *The accuracy of distance measurements in solid-state NMR.* J Magn Reson, 1999. **139**(1): p. 46-59.

155. Laws, D.D., H.M. Bitter, and A. Jerschow, *Solid-state NMR spectroscopic methods in chemistry.* Angew Chem Int Ed Engl, 2002. **41**(17): p. 3096-129.

156. Morris, G.A. and R. Freeman, *Enhancement of nuclear magnetic resonance signals by polarization transfer.* J Am Chem Soc, 1979. **101**(3): p. 760-762.

157. Isas, J.M., R. Langen, and A.B. Siemer, *Solid-State NMR on the static and dynamic domains of huntingtin exon-1 fibrils.* Biochemistry, 2015. **54**: p. 3942−3949.

158. Takegoshi, K., S. Nakamura, and T. Terao, *^{13}C-^{1}H dipolar-assisted rotational resonance in magic-angle spinning NMR.* Chem Phys Lett, 2001. **344**(5-6): p. 631-637.

159. Hou, G., et al., *Spin diffusion driven by R-symmetry sequences: applications to homonuclear correlation spectroscopy in MAS NMR of biological and organic solids.* J Am Chem Soc, 2011. **133**(11): p. 3943-53.

160. Schutz, A.K., et al., *Atomic-resolution three-dimensional structure of amyloid beta fibrils bearing the Osaka mutation.* Angew Chem Int Ed Engl, 2015. **54**(1): p. 331-5.

161. Baldus, M. and B.H. Meier, *Total correlation spectroscopy in the solid state. The use of scalar couplings to determine the through-bond connectivity.* J Magn Reson Ser A, 1996. **121**(1): p. 65-69.

162. Baldus, M., et al., *Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems.* Mol Phys, 1998. **95**(6): p. 1197-1207.

163. Hong, M., *Resonance assignment of 13C/15N labeled solid proteins by two- and three-dimensional magic-angle-spinning NMR.* J Biomol NMR, 1999. **15**(1): p. 1-14.

164. Castellani, F., et al., *Determination of solid-state NMR structures of proteins by means of three-dimensional 15N-13C-13C dipolar correlation spectroscopy and chemical shift analysis.* Biochemistry, 2003. **42**(39): p. 11476-83.

165. Wishart, D.S. and B.D. Sykes, *The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data.* J Biomol NMR, 1994. **4**(2): p. 171-180.

166. Wishart, D.S., B.D. Sykes, and F.M. Richards, *The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy.* Biochemistry, 1992. **31**(6): p. 1647-51.

167. Shen, Y., et al., *TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts.* J Biomol NMR, 2009. **44**(4): p. 213-23.

168. Veshtort, M. and R.G. Griffin, *SPINEVOLUTION: a powerful tool for the simulation of solid and liquid state NMR experiments.* J Magn Reson, 2006. **178**(2): p. 248-282.

169. Costa, P.R., et al, *Solid-state NMR measurement of Psi in peptides: a NCCN 2Q-heteronuclear local field experiment.* Chemical Physics Letters, 1997. **280**(1-2): p. 95-103.

170. Gullion, T., Schaefer, J., *Rotational-echo double-resonance NMR (1969).* Journal of Magnetic Resonance, 1989. **81**(1): p. 196-200.

171. Feng, X., et al, *Direct determination of a molecular torsional angle by solid-state NMR.* Chemical Physics Letters, 1996. **257**(3-4): p. 314-320.

172. Zech, S.G., A.J. Wand, and A.E. McDermott, *Protein structure determination by high-resolution solid-state NMR spectroscopy: application to microcrystalline ubiquitin.* J Am Chem Soc, 2005. **127**(24): p. 8618-26.

173. Tang, M., G. Comellas, and C.M. Rienstra, *Advanced solid-state NMR approaches for structure determination of membrane proteins and amyloid fibrils.* Acc Chem Res, 2013. **46**(9): p. 2080-8.

174. Tuttle, M.D., et al., *Solid-state NMR structure of a pathogenic fibril of full-length human alpha-synuclein.* Nat Struct Mol Biol, 2016. **23**(5): p. 409-15.

175. Morag, O., et al., *The NMR-Rosetta capsid model of M13 bacteriophage reveals a quadrupled hydrophobic packing epitope.* Proc Natl Acad Sci U S A, 2015. **112**(4): p. 971-6.

176. Wang, S., et al., *Solid-state NMR (1)(3)C and (1)(5)N resonance assignments of a seven-transmembrane helical protein Anabaena Sensory Rhodopsin.* Biomol NMR Assign, 2013. **7**(2): p. 253-6.

177. Huber, M., Hiller, S., Schanda, P., Ernst, M., Bockmann, A., Verel, R., Meier, B.H., *A Proton-Detected 4D Solid-State NMR Experiment for Protein Structure Determination.* Chemphyschem, 2011. **12**: p. 915-918.

178. Nieuwkoop, A.J. and C.M. Rienstra, *Supramolecular protein structure determination by site-specific long-range intermolecular solid state NMR spectroscopy.* J Am Chem Soc, 2010. **132**(22): p. 7570-1.

179. Linser, R., *Solid-state NMR spectroscopic trends for supramolecular assemblies and protein aggregates.* Solid State Nucl Magn Reson, 2017. **87**: p. 45-53.

180. Meier, B.H., R. Riek, and A. Bockmann, *Emerging Structural Understanding of Amyloid Fibrils by Solid-State NMR.* Trends Biochem Sci, 2017. **42**(10): p. 777-787.

181. Rogawski, R. and A.E. McDermott, *New NMR tools for protein structure and function: Spin tags for dynamic nuclear polarization solid state NMR.* Arch Biochem Biophys, 2017. **628**: p. 102-113.

182. Tycko, R., *Physical and structural basis for polymorphism in amyloid fibrils.* Protein Sci, 2014. **23**(11): p. 1528-39.

183. Comellas, G. and C.M. Rienstra, *Protein structure determination by magic-angle spinning solid-state NMR, and insights into the formation, structure, and stability of amyloid fibrils.* Annu Rev Biophys, 2013. **42**: p. 515-36.

184. Ladizhansky, V., *Applications of solid-state NMR to membrane proteins.* Biochim Biophys Acta Proteins Proteom, 2017. **1865**(11 Pt B): p. 1577-1586.

185. Schanda, P. and M. Ernst, *Studying Dynamics by Magic-Angle Spinning Solid-State NMR Spectroscopy: Principles and Applications to Biomolecules.* Prog Nucl Magn Reson Spectrosc, 2016. **96**: p. 1-46.

186. Gras, S.L., L.J. Waddington, and K.N. Goldie, *Transmission electron microscopy of amyloid fibrils.* Methods Mol Biol, 2011. **752**: p. 197-214.

187. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis.* Nat Methods, 2012. **9**(7): p. 671-5.

188. Bunaciu, A.A., S. Fleschin, and H.Y. Aboul-Enein, *Biomedical investigations using Fourier transform-infrared microspectroscopy.* Crit Rev Anal Chem, 2014. **44**(3): p. 270-6.

189. Morris, K.L. and L.C. Serpell, *X-ray fibre diffraction studies of amyloid fibrils.* Methods Mol Biol, 2012. **849**: p. 121-35.

190. Sharma, D., et al., *Polyglutamine homopolymers having 8-45 residues form slablike beta-crystallite assemblies.* Proteins, 2005. **61**(2): p. 398-411.

191. Tam, S., et al., *The chaperonin TRiC blocks a huntingtin sequence element that promotes the conformational switch to aggregation.* Nat Struct Mol Biol, 2009. **16**(12): p. 1279-1285.

192. Choudhury, K.R. and N.P. Bhattacharyya, *Chaperone protein HYPK interacts with the first 17 amino acid region of Huntingtin and modulates mutant HTT-mediated aggregation and cytotoxicity.* Biochem Biophys Res Commun, 2015. **456**(1): p. 66-73.

193. Monsellier, E., et al., *Molecular interaction between the chaperone Hsc70 and the N-terminal flank of huntingtin exon 1 modulates aggregation.* J Biol Chem, 2015. **290**(5): p. 2560-2576.

194. Khoshnan, A., J. Ko, and P.H. Patterson, *Effects of intracellular expression of anti-huntingtin antibodies of various specificities on mutant huntingtin aggregation and toxicity.* Proc Natl Acad Sci USA, 2002. **99**(2): p. 1002-1007.

195. Wang, C.E., et al., *Suppression of neuropil aggregates and neurological symptoms by an intracellular antibody implicates the cytoplasmic toxicity of mutant huntingtin.* J Cell Biol, 2008. **181**(5): p. 803-816.

196. Southwell, A.L., et al., *Intrabodies Binding the Proline-Rich Domains of Mutant Huntingtin Increase Its Turnover and Reduce Neurotoxicity.* J Neurosci, 2008. **28**(36): p. 9013-9020.

197. Legleiter, J., et al., *Monoclonal antibodies recognize distinct conformational epitopes formed by polyglutamine in a mutant huntingtin fragment.* J Biol Chem, 2009. **284**(32): p. 21647-21658.

198. Chow, W.N.V., et al., *Degradation of mutant huntingtin via the ubiquitin/proteasome system is modulated by FE65.* Biochem J, 2012. **443**(3): p. 681-689.

199. Ehrnhoefer, D.E., L. Sutton, and M.R. Hayden, *Small changes, big impact: posttranslational modifications and function of huntingtin in Huntington disease.* Neuroscientist, 2011. **17**(5): p. 475-492.

200. De Genst, E., et al., *Structure of a Single-Chain Fv Bound to the 17 N-Terminal Residues of Huntingtin Provides Insights into Pathogenic Amyloid Formation and Suppression.* Journal of molecular biology, 2015. **427**(12): p. 2166-2178.

201. Muller, K., et al., *Review on the Processing and Properties of Polymer Nanocomposites and Nanocoatings and Their Applications in the Packaging, Automotive and Solar Energy Fields.* Nanomaterials (Basel), 2017. **7**(4): p. 74.

202. O'Nuallain, B., et al., *Kinetics and thermodynamics of amyloid assembly using a high-performance liquid chromatography-based sedimentation assay.* Methods in Enzymology, 2006. **413**: p. 34-74.

203. Krebs, M.R., E.H. Bromley, and A.M. Donald, *The binding of thioflavin-T to amyloid fibrils: localisation and implications.* J Struct Biol, 2005. **149**(1): p. 30-7.

204. Amdursky, N., Y. Erez, and D. Huppert, *Molecular rotors: what lies behind the high sensitivity of the thioflavin-T fluorescent marker.* Acc Chem Res, 2012. **45**(9): p. 1548-57.

205. Wetzel, R., et al., *An Aggregate Weight-Normalized Thioflavin-T Measurement Scale for Characterizing Polymorphic Amyloids and Assembly Intermediates.* Methods Mol Biol, 2018. **1777**: p. 121-144.

206. Kar, K., et al., *Critical nucleus size for disease-related polyglutamine aggregation is repeat-length dependent.* Nat Struct Mol Biol, 2011. **18**(3): p. 328-36.

207. DiFiglia, M., et al., *Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain.* Science, 1997. **277**(5334): p. 1990-3.

208. Thakur, A.K., et al., *Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism.* Nat Struct Mol Biol, 2009. **16**(4): p. 380-389.

209. Bhattacharyya, A., et al., *Oligoproline effects on polyglutamine conformation and aggregation.* J Mol Biol, 2006. **355**(3): p. 524-35.

210. Hoop, C.L., et al., *Polyglutamine amyloid core boundaries and flanking domain dynamics in huntingtin fragment fibrils determined by solid-state nuclear magnetic resonance.* Biochemistry, 2014. **53**(42): p. 6653-6666.

211. Choudhury, K.R. and N.P. Bhattacharyya, *Chaperone protein HYPK interacts with the first 17 amino acid region of Huntingtin and modulates mutant HTT-mediated aggregation and cytotoxicity.* Biochem Biophys Res Commun, 2015. **456**(1): p. 66-73.

212. Monsellier, E., et al., *Molecular interaction between the chaperone Hsc70 and the N-terminal flank of huntingtin exon 1 modulates aggregation.* J Biol Chem, 2015. **290**(5): p. 2560-76.

213. Duennwald, M.L., et al., *Flanking sequences profoundly alter polyglutamine toxicity in yeast.* Proc Natl Acad Sci U S A, 2006. **103**(29): p. 11045-50.

214. Tam, S., et al., *The chaperonin TRiC blocks a huntingtin sequence element that promotes the conformational switch to aggregation.* Nat Struct Mol Biol, 2009. **16**(12): p. 1279-85.

215. Mishra, R., et al., *Serine phosphorylation suppresses huntingtin amyloid accumulation by altering protein aggregation properties.* J Mol Biol, 2012. **424**(1-2): p. 1-14.

216. Sun, C.S., et al., *Conformational switch of polyglutamine-expanded huntingtin into benign aggregates leads to neuroprotective effect.* Sci Rep, 2015. **5**: p. 14992.

217. Duim, W.C., et al., *Sub-diffraction imaging of huntingtin protein aggregates by fluorescence blink-microscopy and atomic force microscopy.* Chemphyschem, 2011. **12**(13): p. 2387-90.

218. Wang, X., et al., *Gatekeeper residues in the major curlin subunit modulate bacterial amyloid fiber biogenesis.* Proc Natl Acad Sci U S A, 2010. **107**(1): p. 163-8.

219. Abeln, S. and D. Frenkel, *Disordered flanks prevent peptide aggregation.* PLoS Comput Biol, 2008. **4**(12): p. e1000241.

220. Fuxreiter, M., P. Tompa, and I. Simon, *Local structural disorder imparts plasticity on linear motifs.* Bioinformatics, 2007. **23**(8): p. 950-6.

221. Krishnan, R. and S.L. Lindquist, *Structural insights into a yeast prion illuminate nucleation and strain diversity.* Nature, 2005. **435**(7043): p. 765-72.

222. Lee, C.C., et al., *Design and Optimization of Anti-amyloid Domain Antibodies Specific for beta-Amyloid and Islet Amyloid Polypeptide.* J Biol Chem, 2016. **291**(6): p. 2858-73.

223. Poirier, M.A., et al., *Huntingtin spheroids and protofibrils as precursors in polyglutamine fibrilization.* J Biol Chem, 2002. **277**(43): p. 41032-7.

224. Gasteiger, E., et al., *ExPASy: The proteomics server for in-depth protein knowledge and analysis.* Nucleic Acids Res, 2003. **31**(13): p. 3784-8.

225. Meisl, G., et al., *Molecular mechanisms of protein aggregation from global fitting of kinetic models.* Nat Protoc, 2016. **11**(2): p. 252-72.

226. Chao, J.A., et al., *Inherent protein structural flexibility at the RNA-binding interface of L30e.* J Mol Biol, 2003. **326**(4): p. 999-1004.

227. Bennett, A.E., et al., *Heteronuclear decoupling in rotating solids.* J Chem Phys, 1995. **103**(16): p. 6951-6958.

228. Delaglio, F., et al., *Nmrpipe - a Multidimensional Spectral Processing System Based on Unix Pipes.* Journal of Biomolecular Nmr, 1995. **6**(3): p. 277-293.

229. Mandal, A., et al., *Structural Changes and Proapoptotic Peroxidase Activity of Cardiolipin-Bound Mitochondrial Cytochrome c.* Biophys J, 2015. **109**(9): p. 1873-84.

230. Wetzel, R., *Physical chemistry of polyglutamine: intriguing tales of a monotonous sequence.* J Mol Biol, 2012. **421**(4-5): p. 466-90.

231. Baias, M., et al., *Structure and Dynamics of the Huntingtin Exon-1 N-Terminus: A Solution NMR Perspective.* J Am Chem Soc, 2017. **139**(3): p. 1168-1176.

232. Newcombe, E.A., et al., *Tadpole-like Conformations of Huntingtin Exon 1 Are Characterized by Conformational Heterogeneity that Persists regardless of Polyglutamine Length.* J Mol Biol, 2018. **430**(10): p. 1442-1458.

233. Jayaraman, M., et al., *Slow amyloid nucleation via alpha-helix-rich oligomeric intermediates in short polyglutamine-containing huntingtin fragments.* J Mol Biol, 2012. **415**(5): p. 881-99.

234. Wagner, A.S., et al., *Self-assembly of Mutant Huntingtin Exon-1 Fragments into Large Complex Fibrillar Structures Involves Nucleated Branching.* J Mol Biol, 2018. **430**(12): p. 1725-1744.

235. Vieweg, S., et al., *An Intein-based Strategy for the Production of Tag-free Huntingtin Exon 1 Proteins Enables New Insights into the Polyglutamine Dependence of Httex1 Aggregation and Fibril Formation.* J Biol Chem, 2016. **291**(23): p. 12074-86.

236. Chiti, F. and C.M. Dobson, *Protein misfolding, functional amyloid, and human disease.* Annu Rev Biochem, 2006. **75**: p. 333-366.

237. Mainz, A., et al., *The chaperone αB-crystallin uses different interfaces to capture an amorphous and an amyloid client.* Nat Struct Mol Biol, 2015. **22**(11): p. 898-905.

238. Delaye, M. and A. Tardieu, *Short-range order of crystallin proteins accounts for eye lens transparency.* Nature, 1983. **302**(5907): p. 415-417.

239. Margittai, M. and R. Langen, *Fibrils with parallel in-register structure constitute a major class of amyloid fibrils: molecular insights from electron paramagnetic resonance spectroscopy.* Q Rev Biophys, 2008. **41**(3-4): p. 265-297.

240. Bloemendal, H., et al., *Ageing and vision: structure, stability and function of lens crystallins.* Prog Biophys Mol Bio, 2004. **86**(3): p. 407-485.

241. Ji, F., et al., *The human W42R γD-crystallin mutant structure provides a link between congenital and age-related cataracts.* J Biol Chem, 2013. **288**(1): p. 99-109.

242. Evans, P., et al., *The P23T cataract mutation causes loss of solubility of folded gammaD-crystallin.* J Mol Biol, 2004. **343**(2): p. 435-444.

243. Pande, A., et al., *Decrease in protein solubility and cataract formation caused by the Pro23 to Thr mutation in human gamma D-crystallin.* Biochemistry, 2005. **44**(7): p. 2491-2500.

244. Kardos, J., et al., *Structural studies reveal that the diverse morphology of beta(2)-microglobulin aggregates is a reflection of different molecular architectures.* Biochim. Biophys. Acta, 2005. **1753**(1): p. 108-120.

245. Yoshimura, Y., et al., *Distinguishing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation.* Proc. Natl. Acad. Sci. U.S.A., 2012. **109**(36): p. 14446-14451.

246. Winkler, J., et al., *Chaperone networks in protein disaggregation and prion propagation.* J Struct Biol, 2012. **179**(2): p. 152-160.

247. So, M., D. Hall, and Y. Goto, *Revisiting supersaturation as a factor determining amyloid fibrillation.* Curr Opin Struc Biol, 2016. **36**: p. 32-39.

248. Truscott, R.J.W., *Eye Lens Proteins and Cataracts*, in *Protein misfolding, aggregation, and conformational diseases. Part B: Molecular Mechanisms of Conformational Diseases.*, V.N. Uversky and A. Fink, Editors. 2007, Springer: New York, NY. p. 435-447.

249. Ecroyd, H. and J.A. Carver, *Crystallin proteins and amyloid fibrils.* Cell Mol Life Sci, 2009. **66**(1): p. 62-81.

250. Moreau, K.L. and J.A. King, *Cataract-causing defect of a mutant γ-crystallin proceeds through an aggregation pathway which bypasses recognition by the α-crystallin chaperone.* PLoS ONE, 2012. **7**(5): p. e37256.

251. Moran, S.D., et al., *Amyloid Fiber Formation in Human γD-Crystallin Induced by UV–B Photodamage.* Biochemistry, 2013. **52**(36): p. 6169-6181.

252. Wu, J.W., et al., *Comparative analysis of human γD-crystallin aggregation under physiological and low pH conditions.* PLoS ONE, 2014. **9**(11): p. e112309.

253. Sandilands, A., et al., *Altered aggregation properties of mutant gamma-crystallins cause inherited cataract.* EMBO J, 2002. **21**(22): p. 6005-6014.

254. Moran, S.D., T.O. Zhang, and M.T. Zanni, *An alternative structural isoform in amyloid-like aggregates formed from thermally denatured human γD-crystallin.* Protein Sci, 2014. **23**(3): p. 321-331.

255. Meehan, S., et al., *Characterisation of Amyloid Fibril Formation by Small Heat-shock Chaperone Proteins Human αA-, αB- and R120G αB-Crystallins.* J Mol Biol, 2007. **372**(2): p. 470-484.

256. Frederikse, P.H., *Amyloid-like protein structure in mammalian ocular lenses.* Curr Eye Res, 2000. **20**(6): p. 462-468.

257. Tycko, R., *Physical and structural basis for polymorphism in amyloid fibrils.* Protein Science, 2014. **23**(11): p. 1528-1539.

258. Li, J., et al., *Amyloid-like fibrils from a domain-swapping protein feature a parallel, in-register conformation without native-like interactions.* J Biol Chem, 2011. **286**(33): p. 28988-95.

259. Delaglio, F., et al., *NMRPipe: A multidimensional spectral processing system based on UNIX pipes.* J Biomol NMR, 1995. **6**(3): p. 277-293.

260. Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: development of a software pipeline.* Proteins, 2005. **59**(4): p. 687-696.

261. Harris, R.K., et al., *Further conventions for NMR shielding and chemical shifts (IUPAC Recommendations 2008).* Magn Reson Chem, 2008. **46**(6): p. 582-598.

262. Mandal, A., et al., *Structural Changes and Proapoptotic Peroxidase Activity of Cardiolipin-Bound Mitochondrial Cytochrome c.* Biophys. J., 2015. **109**(9): p. 1873-1884.

263. Shen, Y. and A. Bax, *SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network.* J Biomol NMR, 2010. **48**(1): p. 13-22.

264. Bassnett, S. and G. Duncan, *Direct measurement of pH in the rat lens by ion-sensitive microelectrodes.* Exp Eye Res, 1985. **40**(4): p. 585-590.

265. Helmus, J.J., et al., *Conformational flexibility of Y145Stop human prion protein amyloid fibrils probed by solid-state nuclear magnetic resonance spectroscopy.* Journal of the American Chemical Society, 2010. **132**(7): p. 2393-2403.

266. Zhang, H., S. Neal, and D.S. Wishart, *RefDB: a database of uniformly referenced protein chemical shifts.* J Biomol NMR, 2003. **25**(3): p. 173-195.

267. Kajava, A.V., U. Baxa, and A.C. Steven, *Beta arcades: recurring motifs in naturally occurring and disease-related amyloid fibrils.* FASEB J, 2010. **24**(5): p. 1311-1319.

268. van der Wel, P.C.A., *Domain swapping and amyloid fibril conformation.* Prion, 2012. **6**(3): p. 211-216.

269. McManus, J.J., et al., *Altered phase diagram due to a single point mutation in human gammaD-crystallin.* Proc Natl Acad Sci USA, 2007. **104**(43): p. 16856-16861.

270. Xi, Y.-B., et al., *Cataract-linked mutation R188H promotes βB2-crystallin aggregation and fibrillization during acid denaturation.* Biochem Biophys Res Commun, 2014. **447**(2): p. 244-249.

271. Surguchev, A. and A. Surguchov, *Conformational diseases: Looking into the eyes.* Brain Res Bull, 2010. **81**(1): p. 12-24.

272. Martin, R.W. and K.W. Zilm, *Preparation of protein nanocrystals and their characterization by solid state NMR.* J Magn Reson, 2003. **165**(1): p. 162-174.

273. Khago, D., et al., *Increased hydrophobic surface exposure in the cataract-related G18V variant of human γS-crystallin.* Biochimica et Biophysica Acta (BBA) - General Subjects, 2016. **1860**(1): p. 325-332.

274. Papanikolopoulou, K., et al., *Formation of amyloid fibrils in vitro by human gammaD-crystallin and its isolated domains.* Mol. Vis., 2008. **14**: p. 81-89.

275. Kosinski-Collins, M.S. and J. King, *In vitro unfolding, refolding, and polymerization of human gammaD crystallin, a protein involved in cataract formation.* Protein Sci, 2003. **12**(3): p. 480-490.

276. Zhao, L., et al., *Lanosterol reverses protein aggregation in cataracts.* Nature, 2015. **523**(7562): p. 607-611.

277. Makley, L.N., et al., *Pharmacological chaperone for α-crystallin partially restores transparency in cataract models.* Science (New York, NY), 2015. **350**(6261): p. 674-677.

278. Wild, E.J., et al., *Quantification of mutant huntingtin protein in cerebrospinal fluid from Huntington's disease patients.* J Clin Invest, 2015. **125**(5): p. 1979-1986.

279. Verghese, P.B., et al., *ApoE influences amyloid-β (Aβ) clearance despite minimal apoE/Aβ association in physiological conditions.* Proc. Natl. Acad. Sci. U.S.A., 2013. **110**(19): p. E1807-16.

280. Slingsby, C., *Protein interactions in the calf eye lens: interactions between beta-crystallins are repulsive whereas in gamma-crystallins they are attractive.* Eur Biophys J, 1992. **21**(1): p. 1-12.

281. Miettinen, M.S., et al., *Assessing Polyglutamine Conformation in the Nucleating Event by Molecular Dynamics Simulations.* J Phys Chem B, 2012. **116**(34): p. 10259-10265.

282. Johnson, S.M., et al., *The transthyretin amyloidoses: from delineating the molecular mechanism of aggregation linked to pathology to a regulatory-agency-approved drug.* J Mol Biol, 2012. **421**(2-3): p. 185-203.

283. Habchi, J., et al., *An anticancer drug suppresses the primary nucleation reaction that initiates the production of the toxic Aβ42 aggregates linked with Alzheimer's disease.* Sci. Adv., 2016. **2**(2): p. e1501244.

284. Center for Disease Control WONDER Online Database. *About Underlying Cause of Death, 1999-2016*. 2016 [cited 2018; Available from: https://wonder.cdc.gov/ucd-icd10.html.

285. Rafieian-Kopaei, M., et al., *Atherosclerosis: process, indicators, risk factors and new hopes.* Int J Prev Med, 2014. **5**(8): p. 927-46.

286. Vergeer, M., et al., *The HDL hypothesis: does high-density lipoprotein protect from atherosclerosis?* J Lipid Res, 2010. **51**(8): p. 2058-73.

287. Khuseyinova, N. and W. Koenig, *Apolipoprotein A-I and risk for cardiovascular diseases.* Curr Atheroscler Rep, 2006. **8**(5): p. 365-73.

288. Haghikia, A. and U. Landmesser, *Effects of Apolipoprotein A-I/High-Density Lipoprotein Cholesterol on Atherosclerotic Vascular Disease: Critical Impact of Atherosclerosis Disease Stage and Disease Milieu?* JACC Basic Transl Sci, 2018. **3**(2): p. 210-212.

289. Lau, D. and S. Baldus, *Myeloperoxidase and its contributory role in inflammatory vascular disease.* Pharmacol Ther, 2006. **111**(1): p. 16-26.

290. Podrez, E.A., H.M. Abu-Soud, and S.L. Hazen, *Myeloperoxidase-generated oxidants and atherosclerosis.* Free Radic Biol Med, 2000. **28**(12): p. 1717-25.

291. Shao, B., *Site-specific oxidation of apolipoprotein A-I impairs cholesterol export by ABCA1, a key cardioprotective function of HDL.* Biochim Biophys Acta, 2012. **1821**(3): p. 490-501.

292. Shao, B., et al., *Myeloperoxidase impairs ABCA1-dependent cholesterol efflux through methionine oxidation and site-specific tyrosine chlorination of apolipoprotein A-I.* J Biol Chem, 2006. **281**(14): p. 9001-4.

293. Shao, B., et al., *Methionine oxidation impairs reverse cholesterol transport by apolipoprotein A-I.* Proc Natl Acad Sci U S A, 2008. **105**(34): p. 12224-9.

294. Shao, B., et al., *Humans with atherosclerosis have impaired ABCA1 cholesterol efflux and enhanced high-density lipoprotein oxidation by myeloperoxidase.* Circ Res, 2014. **114**(11): p. 1733-42.

295. Peng, D.Q., et al., *Tyrosine modification is not required for myeloperoxidase-induced loss of apolipoprotein A-I functional activities.* J Biol Chem, 2005. **280**(40): p. 33775-84.

296. Peng, D.Q., et al., *Apolipoprotein A-I tryptophan substitution leads to resistance to myeloperoxidase-mediated loss of function.* Arterioscler Thromb Vasc Biol, 2008. **28**(11): p. 2063-70.

297. Huang, Y., et al., *An abundant dysfunctional apolipoprotein A1 in human atheroma.* Nature medicine, 2014. **20**(2): p. 193-203.

298. Shao, B., et al., *Tyrosine 192 in apolipoprotein A-I is the major site of nitration and chlorination by myeloperoxidase, but only chlorination markedly impairs ABCA1-dependent cholesterol transport.* J Biol Chem, 2005. **280**(7): p. 5983-93.

299. Brubaker, G., et al., *Apolipoprotein A-I lysine modification: effects on helical content, lipid binding and cholesterol acceptor activity.* Biochimica et biophysica acta, 2006. **1761**(1): p. 64-72.

300. Wu, Z., et al., *The refined structure of nascent HDL reveals a key functional domain for particle maturation and dysfunction.* Nat Struct Mol Biol, 2007. **14**(9): p. 861-8.

301. Bergt, C., et al., *The myeloperoxidase product hypochlorous acid oxidizes HDL in the human artery wall and impairs ABCA1-dependent cholesterol transport.* Proc Natl Acad Sci U S A, 2004. **101**(35): p. 13032-7.

302. Chan, G.K., et al., *Myeloperoxidase-mediated Methionine Oxidation Promotes an Amyloidogenic Outcome for Apolipoprotein A-I.* J Biol Chem, 2015. **290**(17): p. 10958-71.

303. Panzenbock, U., et al., *Oxidation of methionine residues to methionine sulfoxides does not decrease potential antiatherogenic properties of apolipoprotein A-I.* J Biol Chem, 2000. **275**(26): p. 19536-44.

304. Jayaraman, S., et al., *Impact of self-association on function of apolipoprotein A-I.* Journal of biological chemistry, 2011. **286**(41): p. 35610-23.

305. Burstein, E.A., N.S. Vedenkina, and M.N. Ivkova, *Fluorescence and the location of tryptophan residues in protein molecules.* Photochem Photobiol, 1973. **18**(4): p. 263-79.

306. Davidson, W.S., et al., *Structural organization of the N-terminal domain of apolipoprotein A-I: studies of tryptophan mutants.* Biochemistry, 1999. **38**(43): p. 14387-95.

307. Saito, H., et al., *Domain structure and lipid interaction in human apolipoproteins A-I and E, a general model.* J Biol Chem, 2003. **278**(26): p. 23227-32.

308. Brouillette, C.G., et al., *Forster resonance energy transfer measurements are consistent with a helical bundle model for lipid-free apolipoprotein A-I.* Biochemistry, 2005. **44**(50): p. 16413-25.

309. Chetty, P.S., et al., *Helical structure and stability in human apolipoprotein A-I by hydrogen exchange and mass spectrometry.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(45): p. 19005-10.

310. Hoofnagle, A.N., et al., *Multiple-reaction monitoring-mass spectrometric assays can accurately measure the relative protein abundance in complex mixtures.* Clinical chemistry, 2012. **58**(4): p. 777-81.

311. Tubb, M.R., L.E. Smith, and W.S. Davidson, *Purification of recombinant apolipoproteins A-I and A-IV and efficient affinity tag cleavage by tobacco etch virus protease.* Journal of lipid research, 2009. **50**(7): p. 1497-504.

312. Fritzsching, K.J., M. Hong, and K. Schmidt-Rohr, *Conformationally selective multidimensional chemical shift ranges in proteins from a PACSY database purged using intrinsic quality criteria.* J Biomol NMR, 2016. **64**(2): p. 115-130.

313. Mei, X. and D. Atkinson, *Crystal structure of C-terminal truncated apolipoprotein A-I reveals the assembly of high density lipoprotein (HDL) by dimerization.* The Journal of biological chemistry, 2011. **286**(44): p. 38570-82.

314. Wong, Y.Q., et al., *Methionine oxidation induces amyloid fibril formation by full-length apolipoprotein A-I.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(5): p. 1977-82.

315. Zandomeneghi, G., et al., *FTIR reveals structural differences between native beta-sheet proteins and amyloid fibrils.* Protein Sci, 2004. **13**(12): p. 3314-21.

316. Townsend, D., et al., *Heparin and Methionine Oxidation Promote the Formation of Apolipoprotein A-I Amyloid Comprising alpha-Helical and beta-Sheet Structures.* Biochemistry, 2017. **56**(11): p. 1632-1644.

317. Cohen, J.S., et al., *13C NMR analysis of methionine sulfoxide in protein.* J. Biochem. Biophys. Methods, 1979. **1**(3): p. 145-151.

318. Skvortsov, A.N., et al., *Molecular Structure and Spectral Properties of Methionine Sulfone, Product of Methionine Oxidation.* Russian Journal of Organic Chemistry, 2003. **39**(2): p. 170-175.

319. Zhao, L., J.N. Buxbaum, and N. Reixach, *Age-related oxidative modifications of transthyretin modulate its amyloidogenicity.* Biochemistry, 2013. **52**(11): p. 1913-26.

320. Kerschen, P. and V. Plante-Bordeneuve, *Current and Future Treatment Approaches in Transthyretin Familial Amyloid Polyneuropathy.* Curr Treat Options Neurol, 2016. **18**(12): p. 53.

321. Goldschmidt, L., et al., *Identifying the amylome, proteins capable of forming amyloid-like fibrils.* Proc Natl Acad Sci U S A, 2010. **107**(8): p. 3487-92.

322. Tartaglia, G.G. and M. Vendruscolo, *The Zyggregator method for predicting protein aggregation propensities.* Chem Soc Rev, 2008. **37**(7): p. 1395-401.

323. Sormanni, P., F.A. Aprile, and M. Vendruscolo, *The CamSol method of rational design of protein mutants with enhanced solubility.* J Mol Biol, 2015. **427**(2): p. 478-90.

324. Fernandez-Escamilla, A.M., et al., *Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins.* Nat Biotechnol, 2004. **22**(10): p. 1302-6.

325. Perutz, M.F., 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 beta-peptide of amyloid plaques.* Proc Natl Acad Sci U S A, 2002. **99**(8): p. 5596-600.