Mass spectrometry analysis of amyloid formation mechanisms

Patrick Daniel Knight

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

Astbury Centre for Structural Molecular Biology

University of Leeds

September 2017

Jointly Authored publications

Throughout this thesis the work directly attributable to the candidate is as follows:

(i) Literature research and compilation of the manuscript stated above.

(ii) The candidate performed all the experimental work and data analysis unless otherwise stated.

Details of jointly authored publications and the contributions of other authors to these manuscripts:

Chapter 4 contains work from the following manuscript published in the European Journal of Mass Spectrometry in 2017:

Identification of a novel site of interaction between ataxin-3 and the amyloid aggregation inhibitor polyglutamine binding peptide 1

Patrick D Knight, Theodoros K Karamanos, Sheena E Radford, Alison E Ashcroft

In this work I provided background information on polyglutamine protein aggregation and the effect of the polypeptide QBP1. I expressed and purified all of the proteins used in the investigation and performed all of the analysis involving the Thioflavin-T fluorescence assay, circular dichroism and ion mobility-mass spectrometry. T. K. Karamanos carried out the nuclear magnetic resonance experiments. I designed the experiments performed along with S. E. Radford and A. E. Ashcroft. All of the authors were involved in preparation of the manuscript.

Acknowledgments:

I would like to express my gratitude to all of the people who have helped with the production of the thesis, both in the lab and outside of it.

My first thank you is for Prof. Alison Ashcroft and Prof. Sheena Radford without whom this PhD could not have happened. Your advice and guidance throughout my PhD has been exceptional and you have helped me grow as a scientist.

I would also like to thank all of the members of the Ashcroft and Radford laboratories. On innumerable occasions your knowledge and assistance have been invaluable. In particular I would like to thank Dr Katie Stewart and Dr Antonio Calabrese for their editorial assistance on this and many other documents. I would also like to thank Dr Theo Karamanos for his interest in the project and for providing NMR expertise. Thank you as well to Dr James Ault and Mr Nasir Khan for persuading me that things are harder to break than they look.

Finally I would like to say a big thank you to everyone who has made my time in the lab so much fun. I have tremendously enjoyed all of the games, social events, trips and ridiculous conversations.

Abstract:

More than 50 human conditions are characterised by the deposition of aberrantly aggregated proteins into amyloid fibrils. These diseases range from neurodegenerative diseases associated with aging to systemic conditions associated with medical interventions. The protein(s) involved in aggregation varies in each condition. Few of these diseases currently have therapeutics available and they represent a growing health burden to society. Increasing the understanding of the processes of amyloid protein aggregation and the mechanisms by which it can be inhibited is vital to developing strategies to combat these diseases.

In this thesis, mass spectrometry and supporting biophysical and biochemical techniques were used to characterise the activity and interactions of two amyloid aggregation modulators, YDL085CA and the peptide QBP1, with amyloidogenic proteins.

YDL085CA, a small highly charged protein orthologous to the known modifier of amyloid aggregation MOAG-4/SERF1, was found to have an extremely buffer dependent effect on the aggregation of the amyloid protein Aβ40. In ammonium bicarbonate it was shown to accelerate amyloid formation while in sodium phosphate it was shown to act as an inhibitor. This observation suggests a prominent role for ionic strength and specific ion effects. The activity of YDL085CA was shown, by crosslinking followed by mass spectrometric analysis, to be mediated by interactions with regions of α-helical secondary structure in the YDL085CA protein.

QBP1, an 11-residue synthetic peptide, has been shown previously to inhibit the formation of amyloid fibrils by polyglutamine proteins, including ataxin-3. Native mass spectrometry studies on a range of ataxin-3 constructs revealed that QBP1 interacts with the monomeric ataxin-3. Unexpectedly, the site of interaction was localised not to the polyglutamine domain, but to a distal region lacking defined structure.

Overall, the results presented demonstrate two different mechanisms for the modulation of amyloid protein aggregation. In addition, the data illustrate the power and versatility of mass spectrometry as a technique for investigating protein interactions.

Table of contents

Contents

Jointly Authored publications ii

Acknowledgments: iii

Abstract: iv

Table of contents vi

Table of Figures: ix

Table of Tables: xv

Abbreviations: xvii

Amino acid abbreviations xxi

1 Introduction 1

1.1 Fundamentals of mass spectrometry 1

1.1.1 Ionisation 2

1.1.2 Mass determination and mass analysers 9

1.1.3 Mass spectrometry analysers – common terms 17

1.1.4 Detectors 19

1.1.5 Tandem mass spectrometry 20

1.1.6 Ion mobility spectrometry 26

1.1.7 Application of MS to the investigation of proteins 34

1.2 Amyloid 49

1.2.1 Protein Folding, misfolding and aggregation 50

1.2.2 Amyloid formation – mechanisms and structure 53

1.2.3 Determining amyloid structures 57

1.2.4 Amyloid in disease 62

1.2.5 Examples of amyloid diseases – Alzheimer's disease and Amyloid-β 67

1.2.6 Examples of amyloid diseases – Polyglutamine disease and ataxin-3 73

1.3 Selected other methods for studying amyloid proteins 79

1.3.1 Thioflavin T fluorescence assay 79

1.3.2 Transmission electron microscopy 80

1.3.3 1H-T1ρ NMR experiments: 81

1.3.4 Far UV-circular dichroism 82

1.4 Aims of this thesis 83
2 Materials and methods 84
2.1 Equipment 84
2.2 Chemicals 88
2.3 Enzymes for molecular biology and protein digest 89
2.4 Antibiotics 89
2.5 Markers and dyes 89
2.6 Kits 90
2.7 LB-agar plates 90
2.8 Buffers 90
2.9 Methods 92
2.9.1 Molecular biology 92
2.9.2 Expression and purification of proteins 94
2.9.3 Native IMS-MS 109
2.9.4 LC-MS 110
2.9.5 Crosslinking 112
2.9.6 ThT fluorescence assay 114
2.9.7 Transmission electron microscopy 115
2.9.8 Nuclear Magnetic Resonance spectroscopy 116
2.9.9 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) 116
2.9.10 Circular dichroism 117
2.9.11 Protein concentration determination 118
2.9.12 Data analysis and statistical methods for Thioflavin T data 119
3 YDL085CA – a modulator of amyloid protein aggregation 121
3.1 Introduction 121
3.1.1 Introduction to MOAG-4 121
3.1.2 Amyloid-β 125
3.1.3 Aims of this chapter: 126
3.2 YDL08CA modulates the aggregation of amyloid proteins 127
3.3 Kinetics of Aβ40 aggregation in the presence of YDL085CA 132
3.4 The effect of buffer on YDL085CA activity 139
3.5 Native MS investigation of the interaction between YDL085CA and amyloid proteins 144
3.6 Chemical crosslinking studies of the YDL085CA:Aβ40 complex 148
3.7 Discussion 167

4 Inhibition of the aggregation Ataxin-3 by polyglutamine binding peptide 1 (QBP1) 173

4.1 Introduction 173

4.2 Ataxin-3 constructs 178

4.3 QBP1 prevents polyQ dependent fibril maturation of ataxin-3 under MS-compatible conditions 180

4.4 Native MS reveals a novel site of interaction between ataxin-3 and QBP1 186

4.5 The ubiquitin interacting motifs of QBP1 do not mediate the interaction with QBP1 193

4.6 Ataxin-3 residues 183-221 are necessary and sufficient for binding QBP1 198

4.7 NMR provides evidence of the QBP1:ataxin-3 interaction in solution 202

4.8 The effect of the scrambled QBP1 variant SP1 upon binding to ataxin-3 205

4.9 QBP1 binding does not alter the conformation of ataxin-3 monomers 212

4.10 QBP1 interacts with other amyloid proteins 214

4.11 Discussion 220

5 Concluding remarks and future directions 224

5.1 Conclusions 224

5.2 Future directions 226

Appendix I: 228

Appendix I.I Details of plasmids: 228

Appendix I.2 DNA sequences: 231

Appendix I.3 Protein sequences: 237

Appendix I.4 Further examples protein of purifications: 239

Appendix II: 242

Appendix III: 258

References 267

Table of Figures:

Figure 1.1: A simple schematic of a mass spectrometer.....2
Figure 1.2: The mechanism of MALDI.....4
Figure 1.3: Droplet formation in the ESI/nESI process.....6
Figure 1.4: ESI mechanisms.....8
Figure 1.5: A schematic of a Quadrupole mass analyser.....9
Figure 1.6: A schematic of a simple ToF cell.....12
Figure 1.7: Delayed pulse extraction ToF.....14
Figure 1.8: A schematic of a reflectron ToF.....16
Figure 1.9: Orthogonal acceleration ToF.....17
Figure 1.10: Definitions of m/z resolution.....18
Figure 1.11: A schematic of a multichannel plate (MCP) detector.....20
Figure 1.12: Protein fragmentation nomenclature.....22
Figure 1.13: Scanning MS/MS modes.....25
Figure 1.14: Linear ion mobility spectrometry – mass spectrometry.....27
Figure 1.15: A schematic of a traveling-wave IMS cell.....30
Figure 1.16: The effect of a T-wave on ions.....31
Figure 1.17: MS techniques for studying proteins.....35
Figure 1.18: Non-native and native MS spectra.....37
Figure 1.19: Flow based UV-excimer laser FPOP experiments.....41
Figure 1.20: A simple schematic of a crosslinking MS method.....44
Figure 1.21: Classification of crosslinks.....46
Figure 1.22: The structure and reaction mechanism of BS^3.....47
Figure 1.23: Protein folding energy landscapes.....52
Figure 1.24: The kinetics of amyloid fibril formation and examples of species involved.....52
Figure 1.25: Primary and secondary nucleation and elongation in amyloid fibril formation.....55
Figure 1.26: Theoretical aggregation kinetics for an amyloid protein undergoing inhibition of different aggregation processes.....56

Figure 1.27: Typical characteristics of an amyloid fibril.....58
Figure 1.28: Cryo-EM density maps and atomic models of two morphologies of Tau fibrils.....60
Figure 1.29: Examples of possible amyloid toxic mechanisms.....66
Figure 1.30: Aβ processing at the cell membrane.....69
Figure 1.31: Aβ fibril structures.....71
Figure 1.32: An example of possible oligomer types observed in the aggregation of Aβ by IMS-MS.....72
Figure 1.33: Ataxin-3 structure.....75
Figure 1.34: Ataxin-3 aggregation occurs via a two-step mechanism.....77
Figure 2.1: The purification of ataxin-3 78Q.....98
Figure 2.2: The purification of the maltose binding protein – 183-222 construct.....99
Figure 2.3: The purification of Aβ40.....105
Figure 2.4: An example α-synuclein SEC A280 profile.....107
Figure 3.1: Structural relationships of the MOAG-4 protein family.....122
Figure 3.2: The sequence of Aβ40 and Aβ42.....125
Figure 3.3: Initial experiments suggest that YDL085CA does not modify the aggregation kinetics of Aβ40 at equimolar concentrations.....128
Figure 3.4: Intact LC-MS of YDL085CA indicates various amino acid substitutions.....129
Figure 3.5: Digestion sites of common enzymes on YDL085CA.....130
Figure 3.6: Digestion and sequencing of YDL085CA reveals the presence of several amino acid residue substitutions.....131
Figure 3.7: Intact LC-MS of the codon optimised YDL085CA reveals a single major species of the correct mass.....132
Figure 3.8: Example Aβ40 aggregation kinetics in the presence of YDL085CA.....133
Figure 3.9: Analysis of Aβ40 aggregation in the presence of YDL085CA.....135

Figure 3.10: Aβ40 aggregation kinetics at different concentrations in the presence of YDL085CA.....137
Figure 3.11: Aggregation kinetics for synthetic Aβ40 in the presence of YDL085CA.....138
Figure 3.12: Typical kinetics for Aβ40 in the presence of YDL085CA in ammonium bicarbonate.....140
Figure 3.13: Summary of Aβ40 aggregation kinetics in the presence of YDL085CA in ammonium bicarbonate.....141
Figure 3.14: Transmission electron microscopy reveals a morphological difference between fibrils formed by Aβ40 in the absence and presence of YDL085CA.....143
Figure 3.15: Native nESi-IMS-MS of a mixture Aβ40 and YDL085CA reveals a population of complexes.....145
Figure 3.16: Native MS reveals a 1:1 complex formed by α-synuclein and YDL085CA(S) 146
Figure 3.17: nESI-IMS-MS arrival time distributions for the $M+10H^{10+}$ ion of α-synuclein alone and the α-synuclein:YDL085CA(S) complex.....147
Figure 3.18: SDS-PAGE of Aβ40 and YDL085CA crosslinking at different concentrations of BS^3.....150
Figure 3.19: BS^3 crosslinking map for the 1:1 complex of Aβ40 and YDL085CA.....152
Figure 3.20: Crosslinking with isotopically labelled BS^3.....154
Figure 3.21: SDS-PAGE of three repeats of the Aβ40 and YDL085CA crosslinking experiment using a BS^3 D_0/D_4 mixture.....156
Figure 3.22: Example of data analysis for a crosslinked peptide from a YDL085CA:Aβ40 complex.....158
Figure 3.23: A crosslinking map of the 1:1 complex of Aβ40 and YDL085CA.....159
Figure 3.24: Helical wheel projections of helix 1 and helix 2 of YDL085CA.....161

Figure 3.25: The sequence of YDL085CA indicating the position of reactive residues.....161
Figure 3.26: A crosslinking map of Aβ40 monomer.....162
Figure 3.27: A crosslinking map of YDL085CA monomer.....164
Figure 3.28: Circular dichroism analysis of YDL085CA and Aβ40.....165
Figure 3.29: The position of charged residues in Aβ.....170
Figure 4.1: NMR derived 3D solution structures of QBP1.....176
Figure 4.2: Ataxin-3 constructs used in this investigation.....179
Figure 4.3: The aggregation characteristics of ataxin-3 are preserved in MS compatible buffer conditions.....181
Figure 4.4: QBP1 does not alter the polyQ independent step of ataxin-3 aggregation.....183
Figure 4.5: The presence of QBP1 inhibits the polyQ dependent transition from protofibrils to fibrils.....185
Figure 4.6: Native nESI-MS reveals a 1:1 complex between ataxin-3 78Q and QBP1.....187
Figure 4.7: Native nESI-MS reveals the 1:1 complex between ataxin-3 and QBP1 is maintained for ataxin-3 constructs with sub-pathogenic length polyQ regions.....189
Figure 4.8: The interaction between ataxin-3 and QBP1 is mediated by a region outside of the polyQ domain.....190
Figure 4.9: The Josephin domain alone does not interact with QBP1.....192
Figure 4.10: The sequences of ataxin-3 derived UIM1 and UIM12 peptides.....193
Figure 4.11: Far CV CD data reveals that secondary structure is maintained in the isolated UIM peptides.....195
Figure 4.12: Denaturing MS of the peptide UIM1.....196
Figure 4.13: Native nESI-MS shows no evidence of a UIM1:QBP1 complex.....197

Figure 4.14: Native nESI-MS of UIM12 show no evidence of a complex with QBP1.....198
Figure 4.15: The addition of residues 183-221 restores QBP1 binding in ataxin-3.....199
Figure 4.16: Ataxin-3 residues 183-221 enable QBP1 binding in the context of other proteins.....201
Figure 4.17: ^{1}H-T1ρ NMR confirms that the interactions of QBP1 with ataxin-3 observed in the gas phase are reflective of interactions in solution.....203
Figure 4.18: The sequences of QBP1 and the scrambled peptide, SP1.....205
Figure 4.19: The scrambled peptide SP1 does not inhibit polyQ dependent fibril formation by ataxin-3 78Q.....206
Figure 4.20: The scrambled peptide SP1 interacts with ataxin-3.....207
Figure 4.21: Native nESI-MS of JDU1 in the presence of both QBP1 and SP1G.....208
Figure 4.22: Collision induced dissociation shows that JDU1 forms a more stable complex with QBP1 than with SP1G.....210
Figure 4.23: Collision induced unfolding of the JDU1:QBP1 and JDU1:SP1G complex M+10H^{10+} reveals that the conformational stability of the two complexes are comparable.....211
Figure 4.24: nESI-IMS-MS analyses reveals that the CCS of ataxin-3 constructs are unchanged by QBP1 binding.....213
Figure 4.25: QBP1 does not inhibit the formation of fibrils by α-synuclein.....216
Figure 4.26: QBP1 forms a complex with the compact conformation of α-synuclein.....218
Figure 4.27: Potential amyloidogenic regions of ataxin-3 identified by consensus prediction.....221
Figure AI.1: Plasmid map of PetSac Aβ40.....228
Figure AI.2: Plasmid map of pET23a α-synuclein.....228
Figure AI.3: Plasmid map of pET28b.....229
Figure AI.4: Plasmid map of pDEST17.....229

Figure AI.5: Plasmid map of pMAL Cx5 used in the expression of the MBP fusion protein.....230
Figure AI.6: SEC elution profile for ataxin-3 14Q.....239
Figure AI.7: SEC elution profile for JDU1.....240
Figure AI.8: SEC elution profile for JD+.....240
Figure AI.9: SEC elution profile for JD.....241
Figure AII.1: Analysis of varying Aβ40 concentration upon Aβ40 aggregation in the presence of YDL085CA.....243
Figure AII.2: Analysis of varying YDL085CA concentration upon aggregation of synthetic Aβ40.....244
Figure AIII.1: IMS drift plots for ataxin-3 78Q and ataxin-3 78Q + QBP1.....260
Figure AIII.2: IMS drift plots for ataxin-3 14Q and ataxin-3 14Q + QBP1.....261
Figure AIII.3: IMS drift plots for JDU1 and JDU1 + QBP1.....262
Figure AIII.4: IMS drift plots for JD+ and JD+ + QBP1.....263
Figure AIII.5: IMS drift plots for JD.....264
Figure AIII.6: Aggregation kinetics of ataxin-3 78Q in the presence of SP1.....264
Figure AIII.7: Amylpred 2.0 analysis of the sequence of ataxin-3 78Q.....266

Table of Tables:

Table 1.1: Gas phase fragmentation of peptides product ions generated.....24
Table 1.2: Examples of amyloid diseases.....63
Table 1.3: Lengths of polyQ expansion in the polyQ diseases.....74
Table 2.1: Chemicals used in this investigation.....88
Table 2.2: Enzymes used in this investigation.....89
Table 2.3: Antibiotics used in this investigation.....89
Table 2.4: Gel standards and dyes used in this investigation.....89
Table 2.5: Kits used in this investigation.....90
Table 2.6: LB-Agar plate recipe.....90
Table 2.7: Details of the ataxin-3 auto induction media.....91
Table 2.8: Components of the ataxin-3 auto induction media.....91
Table 2.9: Buffers used in the Nickel affinity chromatography.....92
Table 2.10: Primers used in the generation of novel truncated ataxin-3 constructs.....93
Table 2.11: Primers used in the generation of MBP-183-221.....93
Table 2.12: The sequence of the YDL085CA coding region of the plasmids.....94
Table 2.13: Details of the AKTA program used in the NiNTA purification of ataxin-3.....96
Table 2.14: Details of the AKTA program used to purify ataxin-3.....97
Table 2.15: Solutions conditions used in the purification of Aβ40.....102
Table 2.16: Details of the AKTA run for the purification of Aβ40.....103
Table 2.17: Instrumental conditions for native IMS-MS of proteins in this study.....110
Table 2.18: Typical LC-MS conditions for intact mass analysis.....111
Table 2.19: Typical LC-MS/MS conditions for sequencing of digested peptides.....112
Table 2.20: Components of a 30% Tris-trycine gel for SDS-PAGE.....116
Table 2.21: Extinction coefficients for the protein used in this study.....118

Table 4.1: Kinetic analyses of the data presented in Figure 4.3……182
Table 4.2: Analysis of the aggregation kinetics presented in Figure 4.4……184
Table 4.3: Helical secondary structure content of the peptides UIM1 and UIM12……195
Table 4.4: Decrease in signal intensity in the methyl region upon the addition of ataxin-3 derived constructs relative to QBP1 alone……204
Table AI.1: DNA sequences for the proteins used in this investigation……231
Table AI.2: Predicted masses of the proteins and peptides used in this investigation……237
Table AI.3: Example cross linking data from a 1:1 complex of Aβ40 and YDL085CA……239
Table AII.1: Data relating to Figure 3.3……245
Table AII.2: Data relating to Figure 3.8 and Figure 3.9……246
Table AII.3: Data relating to Figure 3.10……248
Table AII.4: Data relating to Figure 3.11……249
Table AII.5: Data relating to Figure 3.12 and Figure 3.13……250
Table AII.6: Example cross linking data from a 1:1 complex of Aβ40 and YDL085CA……251
Table AIII.1: CCS of ataxin-3 78Q and ataxin-3 78Q:QBP1 complexes……258
Table AIII.2: CCS of ataxin-3 14Q and ataxin-3 14Q:QBP1 complexes……258
Table AIII.3: CCS of JDU1 and JDU1:QBP1 complexes……259
Table AIII.4: CCS of JD+ and JD+:QBP1 complexes……259
Table AIII.5: CCS of JD……260

Abbreviations:

2A	An amyloid-β fibril morphology
3Q	An amyloid-β fibril morphology
APP	Amyloid precursor protein
Aβ	Amyloid-β
BS^3	Bis(sulfosuccinimidyl)suberate
C	Coulomb
CCS	Collisional cross section
CD	Circular dichroism
CEM	Chain Ejection Model
CFR	The central flexible region of ataxin-3, residues 183-291
CI	Chemical ionisation
CID	Collision induced dissociation
CIU	Collision induced unfolding
CRM	Charged Residue Model
CSD	Charge state distribution
DC	Direct current
DTT	1,2-dithioltheritol
e	Elementary charge
ECD	Electron capture dissociation
EI	Electron impact
$E_{kinetic}$	Kinetic energy
EM	Electron microscopy
EMD	Electron multipliers

EOAD	Early onset Alzheimer's disease
ESI	Electrospray ionisation
ETD	Electron transfer dissociation
ExD	Electron-x-dissociation
FAB	Fast atom bombardment
FAIMS	Field asymmetric ion mobility spectrometry
FPOP	Fast photo oxidative labelling of proteins
FWHM	Full width half maximum
GST	Glutathione S-transferase
HDMS	High Definition Mass Spectrometer
HDX	hydrogen deuterium exchange
HPLC	High pressure liquid chromatography
HT	His-tag
IEM	Ion Evaporation Model
IMS	Ion mobility spectrometry
IMS-MS	Ion mobility spectrometry-mass spectrometry
IPTG	Isopropyl-β-D-thiogalacto-pyranoside
IR	infra-red spectroscopy
JD	Joesphin domain
L	Length
LC-MS	Liquid chromatography-mass spectrometry
LOAD	Late onset Alzheimer's disease
m/z	Mass to charge ratio
MALDI	Matrix assisted laser desorption ionisation

MBP	Maltose binding protein
MCP	Multichannel plate
MJD	Machado Joseph disease
MOAG-4	Modifier of aggregation 4
MS	Mass spectrometry
MWCO	Molecular weight cut-off
nESI	Nanoelectrospray ionisation
NMR	Nuclear magnetic resonance spectroscopy
PA	Projection approximation
PDB	Protein Data Bank
PHF	Paired helical filament
PMSF	Phenylmethylsulfonyl fluoride
polyQ	Polyglutamine
ppm	Parts per million
PSA	Projection superposition approximation
PTM	Post-translational modification
QBP1	Polyglutamine binding peptide 1
RF	Radio frequency
SCA3	Spinocerebellar ataxia type-III
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SERF	Small EDRK rich factor
SF	Straight filament
SID	Surface induced dissociation

SP1	A scrambled form of polyglutamine binding peptide 1
ssNMR	Solid state nuclear magnetic resonance spectroscopy
TAT	Trans-activator of transcription
Te	Tetramer
TEV	*Tobacco Etch Virus nuclear-inclusion-a endopeptidase*
ThT	Thioflavin T
TM	Trajectory method
ToF	Time of flight
TTR	Transthyretin
T-wave	Travelling wave
TWIMS	Travelling-wave ion mobility spectrometry
UIM	Ubiquitin interacting motif
ULP1	Ubiquitin-like specific protease 1
UPLC	Ultra-performance liquid chromatography
v	Velocity
Vs	Acceleration potential
z	Charge

Amino acid abbreviations

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartate
E	Glu	Glutamate
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

1 Introduction

Mass spectrometry (MS) is a powerful analytical technique capable of answering a wide array of scientific questions in many fields. In essence a mass spectrometer is a device for the measurement of the mass-to-charge ratio (*m/z*) of a given ion or ions in the gas phase. The earliest applications of MS were in fundamental particle physics in the late 19th century (Thomson 1897) and since then the field has grown to include chemistry, biology and a wide array of academic and industrial applications. The continuing development and refinement of MS technologies and methodologies (some of which are discussed in in the following text) push the boundaries of what is possible. It seems likely that when J. J. Thomson began his work on cathode rays, eventually leading to his being awarded the Nobel prize for Physics, he would have given little thought to the application of his instruments' successors to the analysis of complex biological molecules, as will be the primary concern of this thesis.

1.1 Fundamentals of mass spectrometry

Broadly, a mass spectrometer consists of three regions each with a specific function (Figure 1.1). First the analyte must be converted from its initial (usually neutral) state, be that a solution as is common in liquid chromatography-MS (LC-MS), a gas or even a solid into gas phase charged ions. This process can be achieved through a multitude of different mechanisms each with its own positive and negative attributes. After ionisation, the gas phase ions enter one or more analysers. Analysers in MS separate ions based on their *m/z*. A wide variety of different analysers have been developed and implemented in MS; often instruments have multiple analysers in different arrangements (e.g. tandem quadrupole, quadrupole-time of flight analysers, etc.). Once the ions have been separated based on their *m/z* they must be detected, generally by measuring the impact of the charged particle on a detector plate. In modern mass spectrometers the resulting signal is amplified and converted to a digital signal (if it is not already) before being passed to a data system which then combines the series of m/z and intensity data to form the mass spectra that are typically the output

observed by a user. Each stage of this process will be described in detail in subsequent sections.

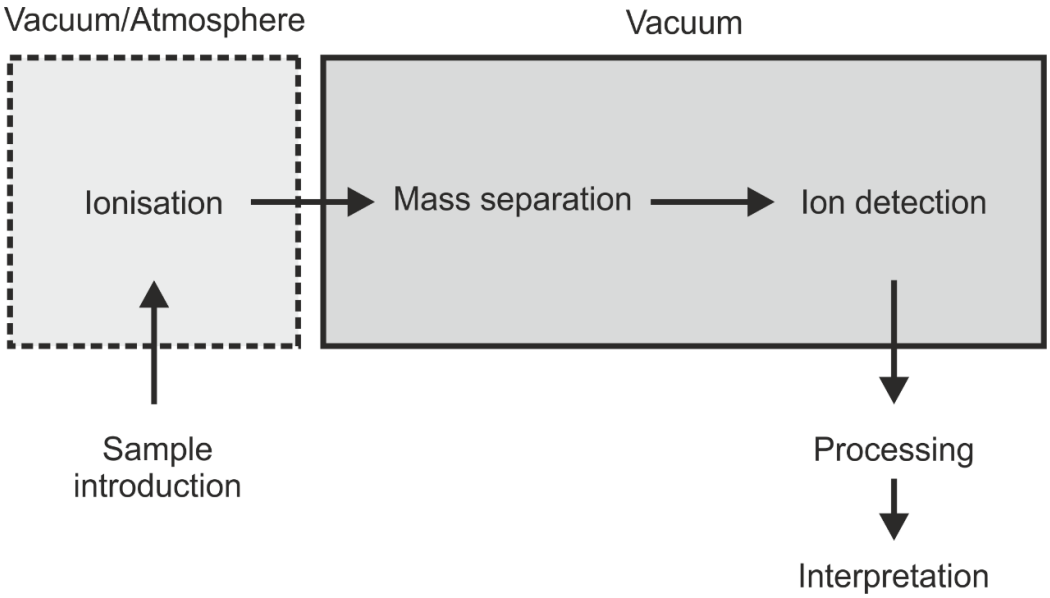


Figure 1.1: A simple schematic of a mass spectrometer. A mass spectrometer consists of a method of sample introduction, such as direct infusion or high pressure liquid chromatography (HPLC) followed by the generation of gas phase ions of the analyte. The ionisation process can occur at atmospheric pressure or under vacuum. Subsequent to ionisation the analyte ions are passed through one or more mass analysers. Both mass analysis and detection occur under vacuum. After separation the ions are detected, the signal processed and the outcome interpreted either by software or by the user.

1.1.1 Ionisation

The first step in analysing a sample by MS is to generate gas phase ions of the analyte. The ionisation source both generates the ions and introduces them into the rest of the mass spectrometer. A variety of ionisation techniques have been developed for MS. Some of the most common are electron impact (EI), chemical ionisation (CI), electrospray ionisation (ESI) and matrix assisted laser desorption ionisation (MALDI).

One of the earliest methods of ionisation, pioneered by Dempster *et al.* (Dempster 1918) in the early part of the 20th century, was electron impact or electron ionisation (EI). In EI, in an orthogonal direction to the sample ions, electrons are generated from a heated filament and pass through a sample

chamber to an anode. The sample chamber is maintained under vacuum. The electrons collide with the sample with high energy and lead to the expulsion of an electron from the analyte and the generation of a radical cation. The nature of EI means that it can only be applied to gaseous samples. In addition the EI process is efficient only for small volatile molecules and as such it is unsuitable for analysis of larger molecules. The high energy of the electron collisions in the source and the inherent instability of the radical cations formed results in fragmentation of the analyte ions, a process which can from a useful fingerprint of the analyte, or which may be problematic in some sample analyses.

Chemical ionisation (CI) overcomes some of the issues surrounding analyte ion fragmentation (Munson and Field 1966). In CI, a reagent gas is first ionised by a process similar to EI before the resulting ions are used to protonate the analyte molecule(s). Both the collisions experienced by the analyte ions and the protonated ions formed by this process are more likely to result in unfragmented molecular ions due to their lower energy and greater stability, respectively. CI allows for the observation of intact molecular ions making it more suitable than EI for many analyses, however analyte fragmentation can still occur. Large molecules and those which are easily fragmented during the collisions with the reagent gas ions can be extremely difficult to analyse with CI.

EI, CI and other early ionisation techniques, while suitable for the analysis of small molecules and very stable compounds, fare poorly when applied to large, non-volatile and relatively unstable molecules such as biological polymers. The high energies involved in the generation of ions typically result in fragmentation of the analyte molecules, rendering analysis problematic. Fast atom bombardment (FAB) overcame some of these issues when it was developed in the 1980s (Barber *et al.* 1981); however the analysis of large biological molecules really took off with the development of new soft ionisation techniques in the late 20th century. Soft ionisation techniques are generally low energy and result in no inherent fragmentation of the sample upon ionisation.

Matrix assisted laser desorption ionisation (MALDI) was developed by Hillenkamp and co-workers (Figure 1.2) (Karas *et al.* 1987; Karas and Hillenkamp 1988). In MALDI an analyte (in solution) is mixed with a matrix, usually consisting of an organic aromatic acid compound, and co-crystallised on a target plate. The sample is exposed to a high intensity UV laser under vacuum. The matrix absorbs the energy from the laser and transfers it to the analyte molecules resulting in both the analyte and the matrix forming gas phase species. Once in the gas phase a proton will be transferred from the matrix to the analyte resulting in the formation of analyte ions. Generally the species generated by this method are solely $(M+H)^+$ ions making interpretation of the spectra relatively easy. The generation of a predominantly mono-protonated population of ions can limit the resolution of MALDI for larger molecules. While MALDI does not cause significant fragmentation of covalent polymers, the use of an organic matrix and the exposure to the laser can disrupt many non-covalent complexes limiting the application of MALDI with regard to native complex analysis. MALDI is widely used for the analysis of small peptides and increasingly in the field of imaging mass spectrometry (Schwamborn, Kriegsmann and Weichert 2017).

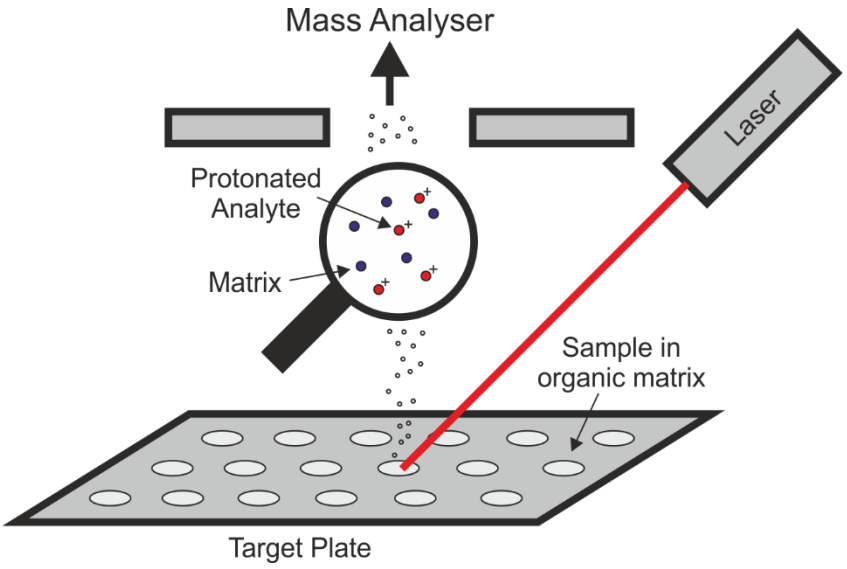


Figure 1.2: The mechanism of MALDI. Samples are prepared in an organic matrix and spotted onto a target plate. The sample spots are dried, inserted into the source of the mass spectrometer and exposed to a laser which generates protonated analyte ions. The ions are then extracted into the analyser region of the mass spectrometer for analysis.

Electrospray ionisation (ESI), developed by Fenn and co-workers around the same time as MALDI (Fenn *et al.* 1989; Yamashita and Fenn 1984), is one of the most widely used methods especially for the analysis of large intact biomolecular complexes. Unlike EI, CI and MALDI, ESI occurs at atmospheric pressure thus reducing the complexity of the source region. In ESI an electric potential is applied to a conductive capillary that terminates in a tapered tip, with the sample solution within, generally in either aqueous or aqueous-organic solvent. This leads to a build-up of charge at the surface of the solution and the formation of a Taylor cone at the tip of the capillary. At the end of the Taylor cone, the stream of solution elongates and then breaks up into droplets of solvent containing analyte molecules (Kebarle 2000). The droplets generated are highly charged and so continue to be mobile in the electric field. As the droplets continue to travel, solvent evaporation occurs and as the volume of the droplets decreases there is an increasing concentration of the charge at the surface of the droplet. Once a critical threshold is reached, the Raleigh limit (Rayleigh 1882), the repulsion between charges overcomes the surface tension of the droplet causing the droplet to split in an event called Coulombic fission. The resulting droplets then undergo this cycle repeatedly resulting in progressively smaller droplets (Figure 1.3).

ESI results typically in analyte molecules acquiring a distribution of charges resulting in a charge state distribution (CSD) (Kebarle 2000). This is especially true for large biomolecules, such as proteins, which have multiple solvent accessible to pronation sites. A number of factors can affect the number of charges and the distribution of charge states including the ionisation voltage applied, the presence of proton donors in the solution and the degree to which the protonation sites on the sample are exposed.

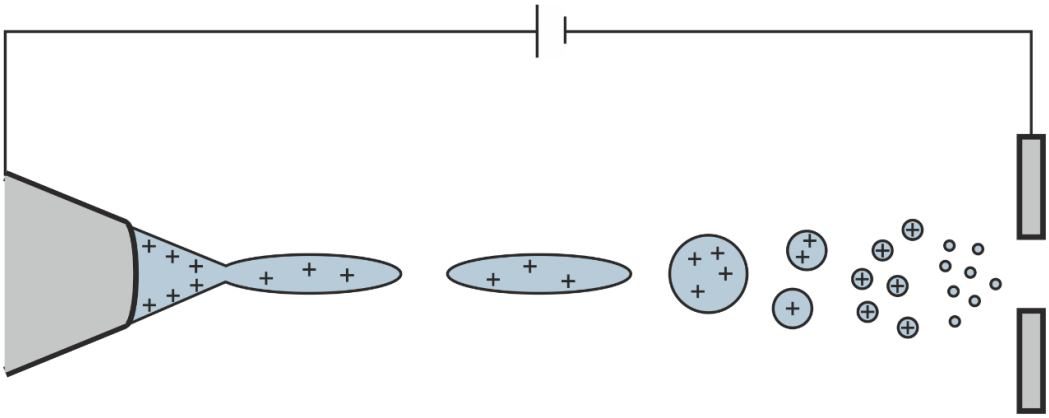


Figure 1.3: Droplet formation in the ESI/nESI process. The sample, in solution, is introduced via a capillary. An electrical potential is maintained a capillary and an electrode at the entrance aperture of the MS. The charge present on the surface of the solution at the end of the capillary results in the formation of a Taylor cone. Droplets form at the end of the cone and undergo sequential fission. As fission progresses the droplets become smaller and the charge more concentrated. The process by which these terminal droplets produce a gas phase ion is described below in Figure 4.

The mechanism by which the analyte molecules escape the terminal droplets and become charged gas phase ions is hotly debated with three main mechanisms proposed and each supported by scientific evidence (Figure 1.4).

The Charged Residue Model (CRM) (Figure 1.4a), proposed by Dole *et al.* in 1968 (Dole *et al.* 1968), suggests that as the solvent-analyte droplet becomes smaller during progressive rounds of Coulombic fission the analyte ion will eventually be left desolvated. The charge remaining in the terminal droplet will be transferred to the analyte resulting in its ionisation. Polar molecules such as proteins are likely to localise in the centre of the droplet and thus evade expulsion by other methods such as Ion Evaporation Model (IEM). The CRM explains the generation of differently charged populations of the same analyte molecule by differences in the initial size of the terminal droplet.

The IEM (Figure 1.4b), first suggested by Iribrane and Thompson in 1976 (Iribarne and Thomson 1976), proposes that once a droplet reaches a critical threshold, analyte ions that are near the surface are expelled as charge carriers to reduce Coulombic repulsion. This mechanism is thought to dominate for small

hydrophobic molecules with a low number of charges, which will preferentially populate the edge of droplets.

The more recently proposed Chain Ejection Model (CEM) (Figure 1.4c) builds on the other two suggested mechanisms (Konermann *et al.* 2013). CEM suggests that individual sections of a protein polymer chain can function independently as charge carriers and undergo a process not unlike IEM. However, being tethered to the droplet by the rest of the polymer chain, the charged areas remain associated in the gas phase. As progressive rounds of IEM-like events occur, the polymer chain is eventually entirely outside of the solvent droplet and is thus free to move away as a gas phase ion. Gas phase measurements of protein structure using ion mobility spectrometry (IMS)-MS show that protein ions formed by ESI maintain native-like structures suggesting that the expulsion of a nascent ion in a residue by residue manner is unlikely for folded proteins; however, CEM may occur for unfolded polymer chains, which lack secondary and tertiary structure.

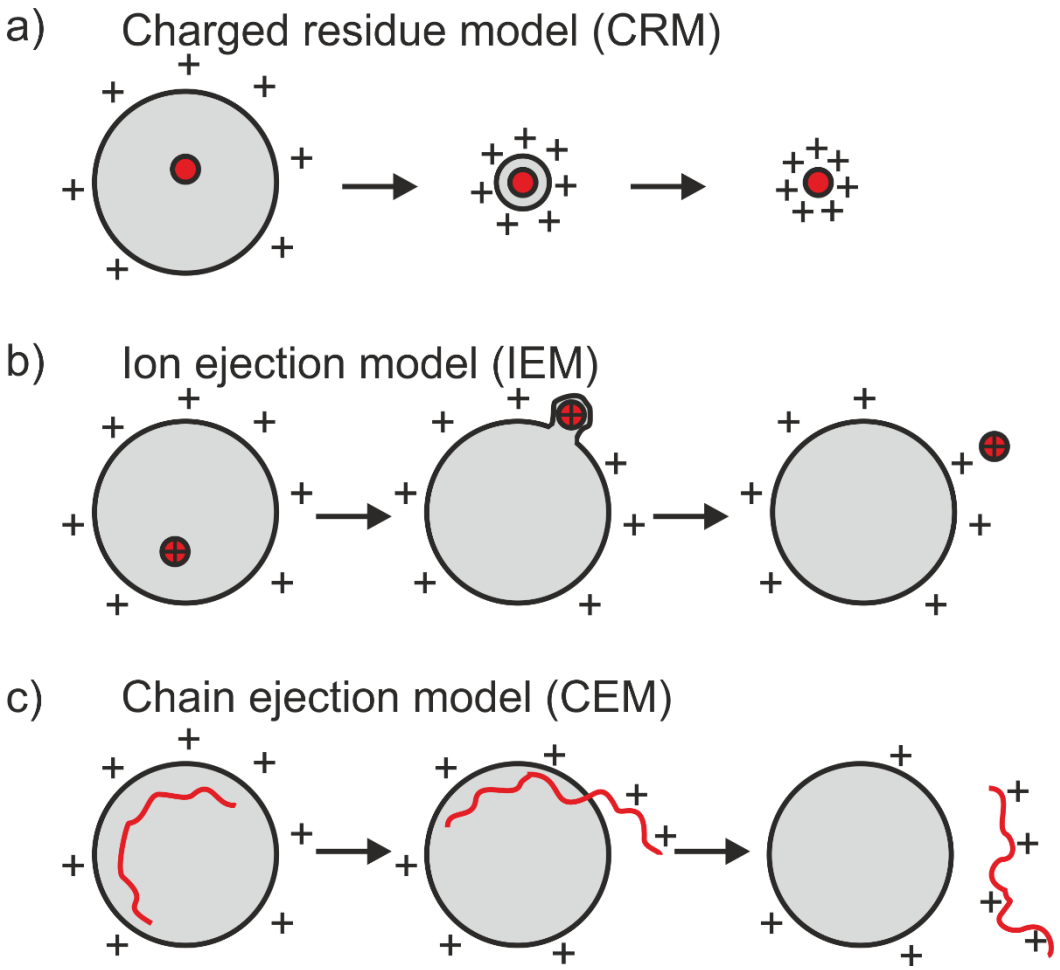


Figure 1.4: ESI mechanisms. Three theorised mechanisms for the formation of gas phase ions from solvent droplets are described. (a) The CRM in which fission leads to a situation where all solvent is lost leaving an analyte ion as the sole charge carrier. (b) The IEM in which an analyte is ejected as a charge carrier to relieve the crowding of charge in the droplet. (c) The CEM involves polymeric ions undergoing partial ejection from the droplet with polymer subunits acting as charge carriers.

NanoESI (nESI) (Wilm and Mann 1996; Wilm and Mann 1994) works by similar principles to ESI, however the flow rates and hence amount of sample consumed are much lower. In addition signal intensity is maintained or improved compared to ESI and there is an increased tolerance for salt (Wilm and Mann 1996).

ESI can generate both both positive and negative ions. Positively charged ions are formed by the transfer of protons to the analyte ion as described above. Negatively charged, deprotonated ions $(M-H)^-$ are generated by deprotonation of the analyte. The polarity of the potential applied to the mass spectrometer determines which of these sets of ions are analysed. It has been suggested that,

in the case of a protein ion in positive mode, the basic amino acid side chains become protonated and carry the charge (Krusemark *et al.* 2009) (Chowdhury, Katta and Chait 1990).

1.1.2 Mass determination and mass analysers

Once gas phase ions have been generated in the source region they are manipulated by a variety of means in order to move them to one or more analysers. Analysers separate ions based on their *m/z*. Many different types of analyser can be employed in a mass spectrometer either individually or in series, as is commonly seen in commercial instrumentation.

1.1.2.1 Quadrupole

One of the most commonly encountered mass to charge analysers is the quadrupole (Figure 1.5) (Paul and Steinwedel 1953). Quadrupoles, along with variants such as the hexapole and octopole, can be used as analysers (with DC and RF voltages applied) or as ion guides (with only RF voltages applied). These analysers (hereafter collectively referred to as quadrupole mass analysers), subject ions to an oscillating electric field and separate ions based on their movement through the analyser under these conditions.

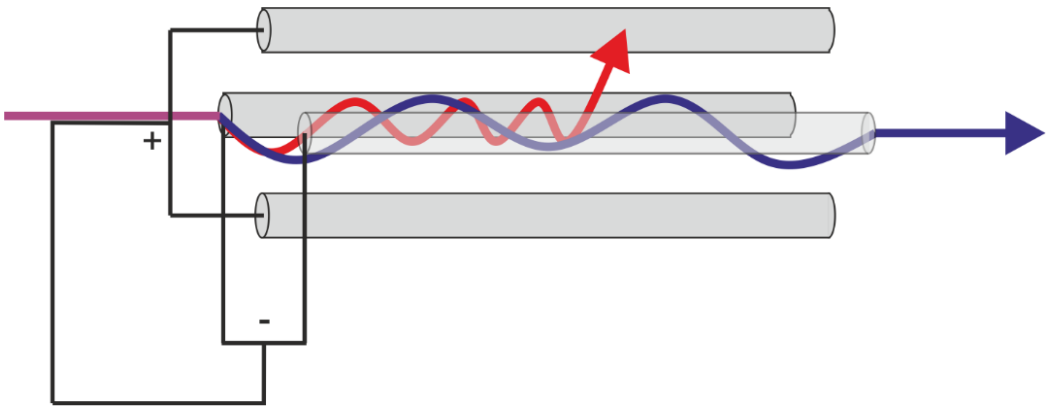


Figure 1.5: A schematic of a Quadrupole mass analyser. The two pairs of poles each carrying a constant DC potential and an oscillating radio frequency with an alternating current. The net effect of these electric fields results in only ions of a certain *m/z* having a stable trajectory (blue) while all other ions have an unstable trajectory (red) and undergo collisions with the poles preventing transmission to the detector.

Quadrupoles consist of electromagnetic poles arranged in pairs with each pair of poles having the opposite polarity. A direct current is applied across the poles to drive the ions towards the exit aperture. A second current is applied to the poles which oscillates in the radio frequency range. The net effect of these potentials on the trajectory an ion is dependent upon the *m/z* of the ion.

Upon entering a quadrupole an ion can have two fates. If the ion oscillates in a manner that allows it to exit the analyser it will be able to reach the detector. These ions are said to have a stable trajectory. Alternatively an ion may oscillate in such a way that it will collide with one of the poles and so fail to reach the detector. These ions have unstable trajectories.

Quadrupole mass analysers are scanning mass analysers. Under a given set of potentials only ions of a specific *m/z* will have stable trajectories and exit the analyser towards the detector. The quadrupole can either be tuned to a single *m/z* window (mass filtering) or can be set to scan through a range of *m/z* sequentially. In scanning mode the quadrupole acts as an ion guide allowing all ions to pass through allowing quadrupoles to be implemented in tandem MS as part of the ion path.

The mass filtering or scanning process of a quadruple mass analyser reduces the duty cycle and sensitivity of the instrument. However, it can also be extremely useful in tandem analyser instruments as it allows for the selection of a specific *m/z* ion for further interrogation in a later analyser. Quadrupoles are generally very robust and can function efficiently at much lower vacuum than other analysers.

1.1.2.2 Time of flight

Time of flight (ToF) mass analysers separate ions of different *m/z* based on the time taken to traverse a field free region after acceleration by an electric potential (Cameron and Eggers 1948; Fredrickson 1946). A heavier ion or an ion with a lower charge will be accelerated less than an ion with lower mass or higher charge. The electrical potential experienced by the ions is uniformly applied and

so all ions of a given charge receive the same amount of energy. Therefore, upon conversion of the potential to kinetic energy ($E_{kinetic}$), the velocity of ions (*v*) of the same charge will be dependent upon their respective masses. This can also be expressed in terms of the charge of the ion (*z*), the elementary charge, *e (1.601 x 10^{-19} C*), and the acceleration potential (V_s)

Equation 1.1:

$$E_{kinetic} = \frac{mv^2}{2} = zeV_s$$

Rearranged to make velocity the subject as shown in equation 1.2.

Equation 1.2:

$$v^2 = \frac{2E_{kinetic}}{m} = \frac{2zeV_s}{m}$$

Ions with greater mass have a lower velocity and therefore take longer to traverse the analyser (length *L*), as described by the equation 1.3.

Equation 1.3

$$t = \frac{L}{v}$$

As v^2 can be defined as shown in equation 1.2, the equation 1.3 can be presented as shown in equation 1.4.

Equation 1.4:

$$t^2 = \frac{m}{z}\left(\frac{L^2}{2eV_s}\right)$$

The *m/z* of an ion can therefore be related to its time of arrival at the end of the analyser (generally at a detector).

Equation 1.5:

$$\frac{m}{z} = \frac{2V_s \times e \times t^2}{L^2}$$

Equation 1.5 is key to the function of all time-of-flight mass to charge analysers. L and V_s can vary between systems and experiments.

In a linear ToF the ions enter the analyser and are immediately subjected to the acceleration potential (the polarity of which depends on the ionisation mode). The ions then undergo separation in the field free region as described above and in Figure 1.6, with the lower *m/z* ions traversing the analyser more quickly than their higher *m/z* counterparts. The ions discharge on the detector as they reach it with the lower *m/z* ions being detected first. The times of arrival at the detector are then used to determine the *m/z*.

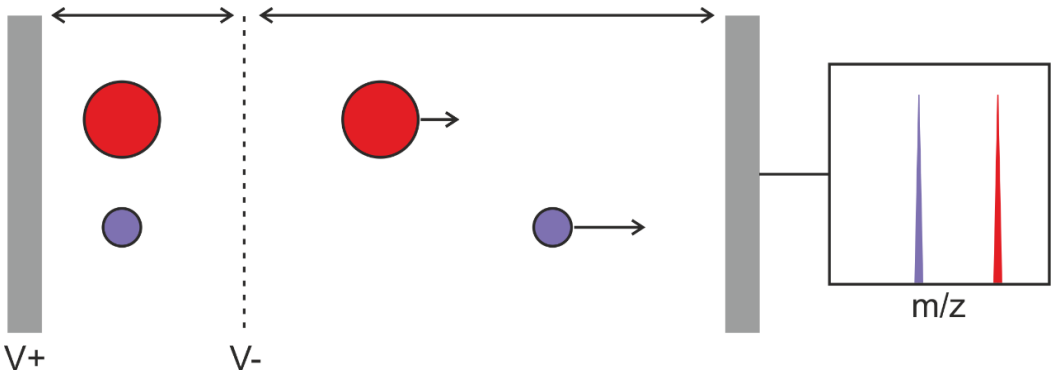


Figure 1.6: A schematic of a simple ToF cell. The analyte ions are exposed to an accelerating potential prior to entering a field free region. Ions with a lower *m/z* (blue) will experience a greater acceleration and thus cross the field free region more quickly than an ion with a higher *m/z* (red).

Linear ToF analysers have a number of limitations. Firstly they are ill suited to continuous ionisation techniques, such as ESI, as each ion will reach the acceleration region at a different time and so the flight times of ions will be impossible to determine. Linear ToF can be used with pulsed ionisation techniques such as MALDI; however, the technique still suffers from poor

resolution. Small differences in the ions generated from a single species entering the detector form the source result in “smearing” of the arrival time and a resultant loss in m/z resolution. This smearing can make the separation of species with only small mass differences impossible.

Three properties contribute to the differences in initial energy for a population of ions with the same *m/z*. The first issue is temporal distribution. Ions arriving in the acceleration region at different times will arrive at the detector at different times, as alluded to above. The ions arriving in the analyser will also have a distribution of kinetic energies resulting from their location during ionisation. The kinetic energy an ion has prior to acceleration will affect its eventual velocity and thus arrival time at the detector. Finally, the spatial distribution of ions during the acceleration process will result in different ions experiencing the acceleration potential for different time periods. The resulting distribution of accelerations will cause a further loss of resolution.

The low resolution of a linear ToF can be mitigated by a variety of instrumental design features. The relative contribution of the distributions discussed above can be made smaller by increasing the length of the drift tube and so increasing the separation of each ion from another of different *m/z*. However the inherent issues of energy distributions are not solved by this process and the increased length of the ToF cell has practical considerations including increased area under vacuum and increased instrument length.

Delayed pulse extraction (Figure 1.7) improves the resolution of the ToF analyser as it corrects for the differences in initial kinetic energy between ions (Wiley and McLaren 1955). This is achieved by using a source in which the ions are generated or collected in a field free region. Ions with greater kinetic energy will move further along the instrument before the field is applied and thus experience the field for a shorter period of time. The result is that the ions with lower initial kinetic energies are accelerated more than those that had high initial energies and as a result, if applied correctly, the isobaric ions should arrive at the detector

at the same time. The narrower arrival time distribution leads to better m/z resolution.

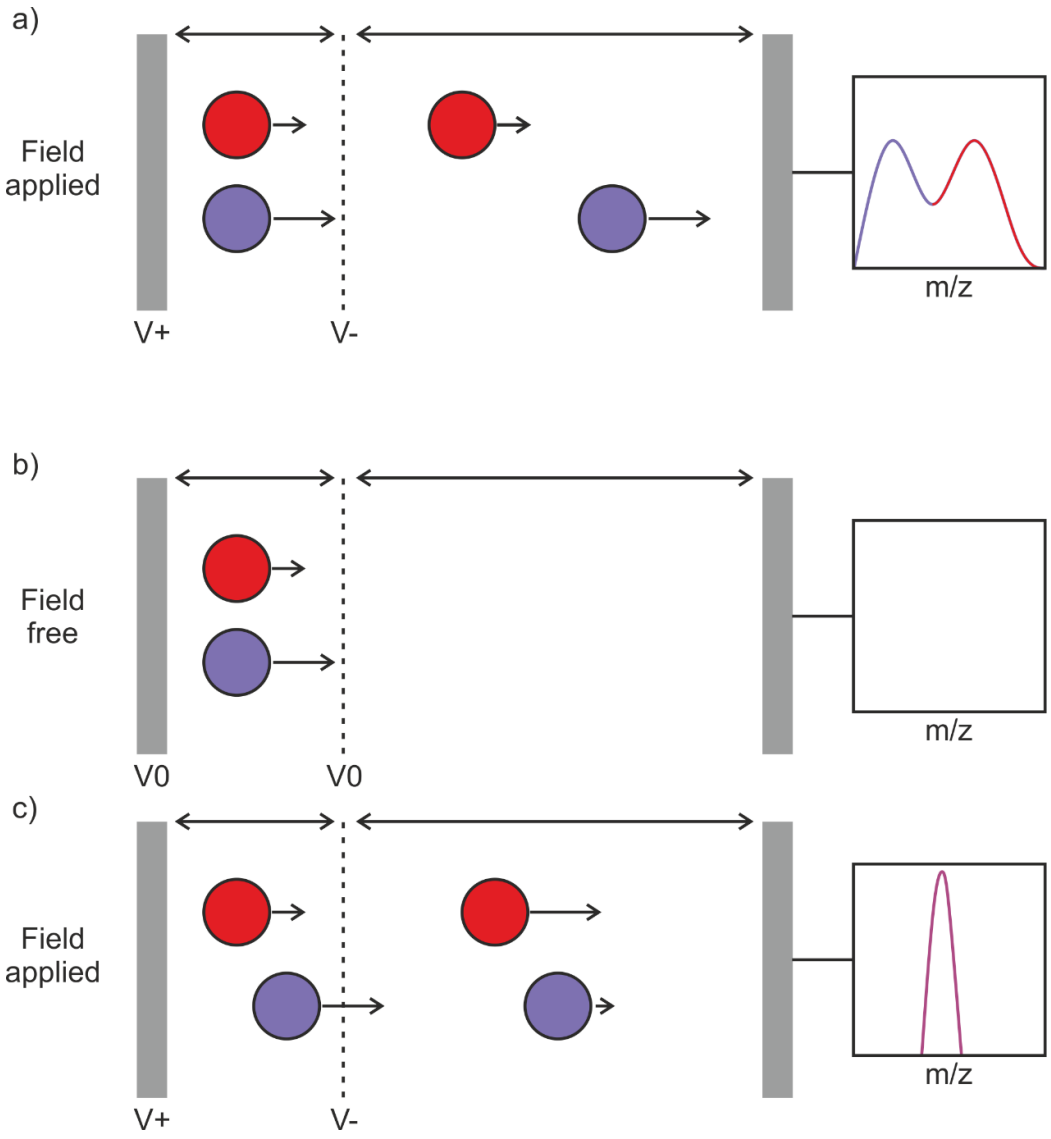


Figure 1.7: Delayed pulse extraction ToF. In a simple ToF (a) ions of the same *m/z* may reach the accelerating region at high velocity (blue) or low velocity (red). This results in poor *m/z* resolution as the high velocity ion will maintain its increased velocity after acceleration. The introduction of delayed pulse extraction reduces this issue. The ions are trapped in the accelerating region without the application of an electrical potential (b). Ions with higher velocity will traverse further across the cell. After a period of ion collection the electrical potential is applied (c). Ions which have traversed further across the accelerating region due to their higher initial velocity will be exposed to the accelerating potential for a shorter period of time than those with a low initial velocity. The ions with the lower initial velocity will therefore undergo more acceleration resulting in a higher final velocity. With correct calibration the two ions should reach the detector at the same time. The mass resolution is thus improved.

Reflectrons (Figure 1.8) correct for variations in ion velocity during the ions' movement through the analyser. The reflectron consists of a set of stacked ring electrodes which function to reflect ions into a detector that is situated outside of the initial ion beam line (Mamyrin 1973). The ions of a given *m/z* entering the analyser with different kinetic energies will interact differently with the field generated by the reflectron. Ions with higher velocity will interact with the reflectron before their slower counterparts. They will, due to their higher kinetic energy, travel further into the reflectron's electric field before being reflected. As a result the ions with higher energy will have a longer flight path than ions with a lower energy. As the ions leave the reflectron they are accelerated back to their initial velocities. The higher energy ions will then catch up with the lower energy ions which have travelled a shorter path and both will reach the detector simultaneously. The reflectron acts as a focussing device increasing the resolution of the analyser by reducing the smearing from the kinetic energy distribution. Reflectrons have the added advantage of effectively doubling the path length of a ToF analyser of a given size and thus increasing the resolution.

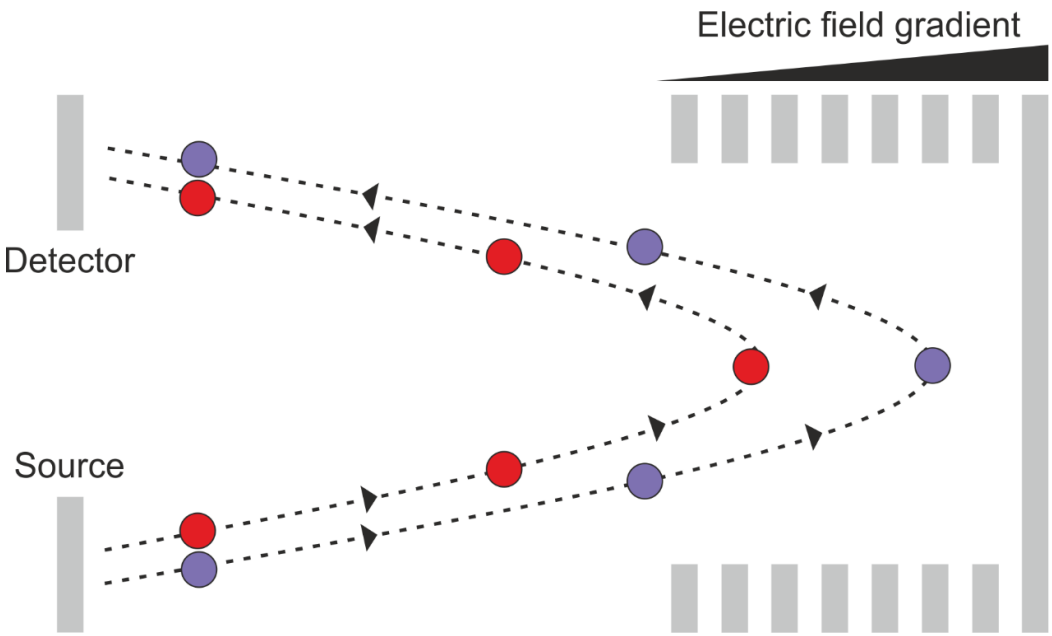


Figure 1.8: A schematic of a reflectron ToF. Two ions of identical *m/z* are generated with low velocity (red) and high velocity (blue). The high velocity ion penetrates further into the electric field of the reflectron and are overtaken by the low velocity ion. The electric field then accelerates the ions back to their initial velocities. The high velocity ion then catches up with the low velocity ion at the detector.

While a ToF instrument with the modifications discussed above has good resolution and can be used to analyse very high mass compounds they can only be paired with pulsed ionisation sources such as MALDI which provide the precise time resolution required for *m/z* determination. However, continuous flow ionisation techniques are beneficial in many cases. Pairing continuous flow ionisation techniques with good temporal resolution is achieved through orthogonal acceleration ToF.

In orthogonal acceleration ToF (Figure 1.9) the continuous ion beam from the source runs perpendicular to the ToF analyser (Dawson and Guilhaus 1989; Coles 1993) (Guilhaus, Selby and Mlynski 2000). A pusher, consisting of a field free compartment and an accelerator, is positioned at an angle to the main beam. Ions from the main beam can be collected in the pusher region for a period of time and then a pulsed potential is applied to accelerate ions into the ToF.

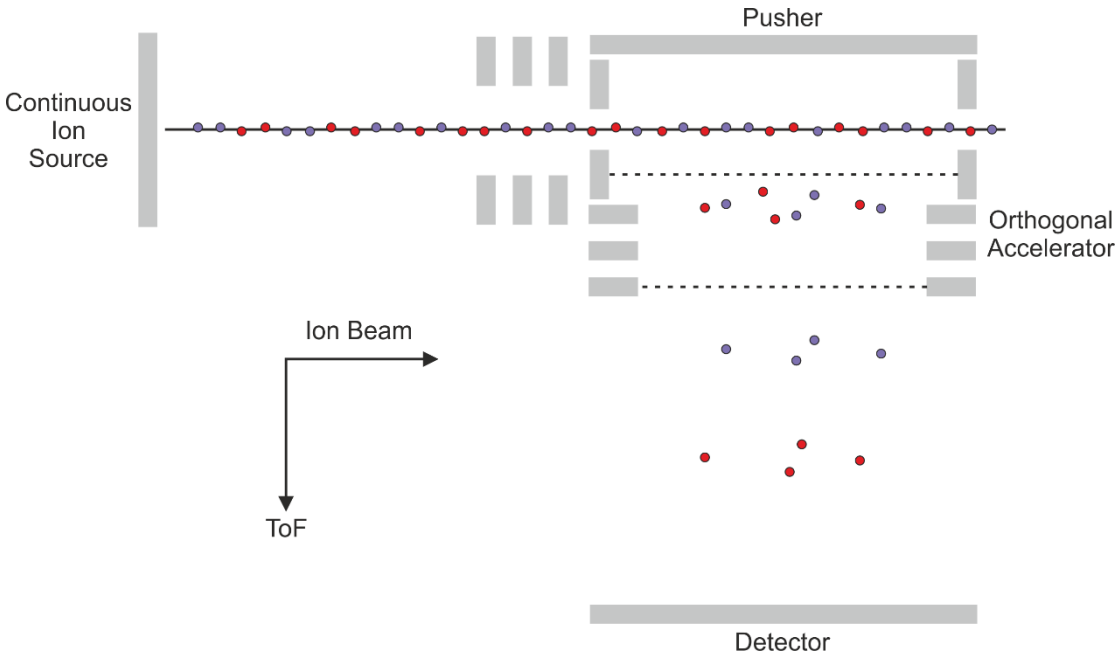


Figure 1.9: Orthogonal acceleration ToF. Ions generated from a continuous ionisation source enter a field free region called the pusher. After a period of ion collection a potential is applied to move the ions into the orthogonal accelerator. The ions can then be exposed to a potential and enter the ToF cell. Red circles represent ions with a low m/z and blue circles ions with a high m/z.

Orthogonal acceleration ToF is commonly used in instruments where MS/MS is applied. It allows the packets of ions that are produced from a quadrupole filtering experiment for example to be collected and investigated.

1.1.3 Mass spectrometry analysers – common terms

Four terms commonly used to compare mass analysers are mass range, sensitivity, mass accuracy and mass resolution. Mass range is simply the *m/z* range over which the analyser(s) can separate ions. ToF instruments have a theoretically unlimited mass range while most commercial quadrupoles are limited to <4000 *m/z* (though they can be modified for higher masses). Sensitivity describes that proportion of ions entering the spectrometer that are detected. Sensitivity is determined by many elements of the instrument.

Mass accuracy is the amount by which a measured mass deviates from the calculated theoretical mass of a known compound. Mass accuracy is often quoted in parts per million (ppm) and is determined by equation 1.6.

Equation 1.6:

$$Mass\ Accuracy = \left(\frac{Mass_{theoretical} - Mass_{observed}}{Mass_{theoretical}}\right) \times 10^6$$

Resolving power is a key characteristic of a mass spectrometer. It determines the ability of the mass spectrometer to separate two species with a similar molecular weight as well as contributing to the accuracy of the *m/z* value determined from the data. There are two definitions of resolution. The first states that the resolution, R, is described by equation 1.7.

Equation 1.7:

$$R = \left(\frac{m/z}{\Delta m/z}\right)$$

Where *Δm/z* is the smallest difference between two signals that can be observed at a given *m/z*. This definition relies on the definition of a peak separation which varies between analyser types. A second method which does not rely on this definition is the full width half maximum (FWHM) method. In FWHM the *Δm/z* is the peak width at 50% of the total signal. These two definitions are compared in Figure 1.10.

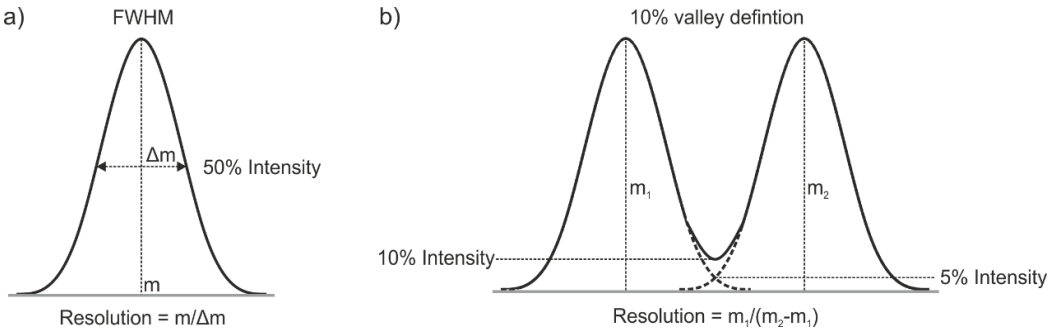


Figure 1.10: Definitions of *m/z* resolution. (a) The full width half maximum (FWHM) definition is based on the width of a single peak. (b) The 10% valley definition is based on the *m/z* difference between two peaks that can be resolved to 10% of the maximum signal. Definitions published by the Royal Society for Chemistry (V. Barwick 2006).

1.1.4 Detectors

Once ions have been separated in the mass analyser they must be detected. There are a range of different detector types available, though most commercial instruments use electron multipliers (EMD) of one of three types: discrete dynode detectors, continuous dynode detectors and microchannel plate detectors. Electron multipliers amplify the signal generated from an ion collision and increase the signal to a detectable level.

Discrete dynode detectors consist of an array of dynodes arranged such that the electrons ejected from one dynode are subject to a high voltage potential which ensures that they collide with the next in a chain resulting in signal amplification along the longitudinal axis. A continuous dynode detector works in a very similar manner, however, instead of a series of individual dynodes, a single curved tube is used with the electron cascade carried along the length of the tube and experiencing secondary interactions with different parts of the same dynode.

Microchannel plate (MCP) detectors are the most commonly used detector in modern instruments. MCP consist of an array of small continuous dynode detectors arranged in a metal plate (Figure 1.11). The length of the channel in a MCP is shorter that of a normal continuous dynode and as such the arrival times of ions can be more accurately measured allowing for greater resolution and mass accuracy.

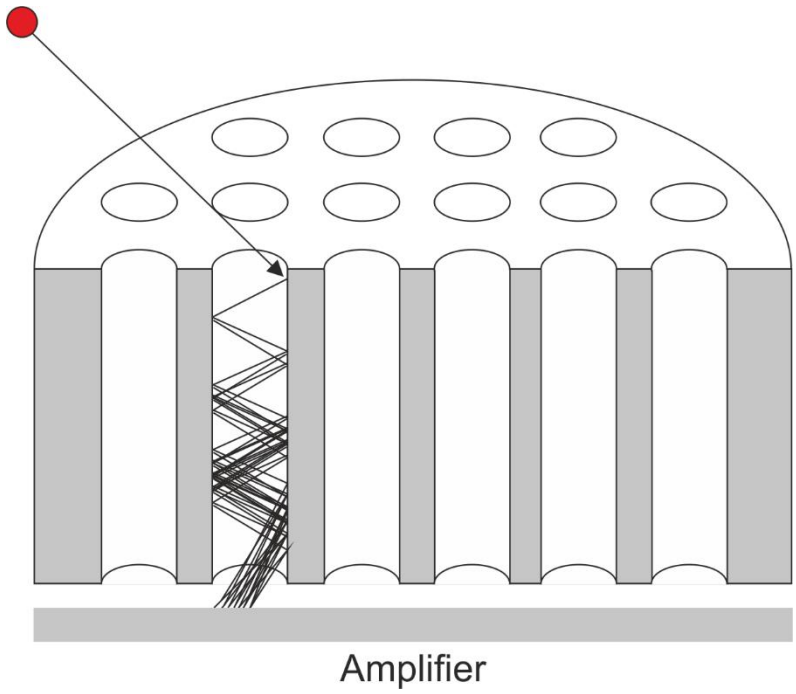


Figure 1.11: A schematic of a multichannel plate (MCP) detector shown in cross-section. The MCP consists of an array of continuous diode detectors. When an ion enters a channel it undergoes a collision with the surface. The collision results in the release of a number of electrons which are directed further along the channel by the application of a potential gradient. The electrons then undergo further collisions resulting in signal amplification.

In addition to electron multiplier detectors, photon multiplier detectors are also used. Instead of an electron these devices generate a photon upon interaction with an ion. The photon subsequently collides with the detector and generates a secondary cascade of electrons. The photon multiplier allows for faster response times than a pure electron multiplier.

1.1.5 Tandem mass spectrometry

While soft ionisation mass spectrometry can provide the mass of a molecule sometimes more information is required, such as when identifying unequivocally the sequence of a chemical polymer. Harder ionisation modes can, through in-source dissociation/fragmentation, provide information on the substituent parts of a molecule, but these techniques often require a pure sample of a single species. Tandem mass spectrometry or MS/MS involves arranging two or more mass analysers in series, the first of which isolates an ion which can then be released into the second analyser and analysed in isolation. The inclusion of a

fragmentation cell between the two analysers allows for the generation of product ions which can provide a great deal of information on the chemical makeup and structure of molecules, ranging from small organic molecules through to large biomolecules.

A simple reaction scheme for the fragmentation of an ion, called a precursor ion, in MS/MS can be represented as:

$$Precursor\ ion \rightarrow Product\ ion + Neutral\ loss$$

It is not necessarily always the case that the product of the fragmentation will result in a single charged, and therefore detectable, species and in some cases more than one product ion will be observed.

There is an enormous array of ion fragmentation techniques available either commercially or in home built instruments. Some of the most common are collision induced dissociation (CID), surface induced dissociation (SID), electron capture dissociation (ECD) and electron transfer dissociation (ETD). Each technique has its own advantages and disadvantages though for many applications the major deciding factor is ease of access.

Due to the wide application of MS/MS techniques in the investigation and sequencing of proteins a standard nomenclature has been developed to describe the fragmentation of these ions. The protein fragmentation nomenclature was first codified by Roepstroff and Fohlman in 1984 (Roepstorff and Fohlman 1984) and developed further by Biemann (Biemann 1988). This system is shown in Figure 1.12.

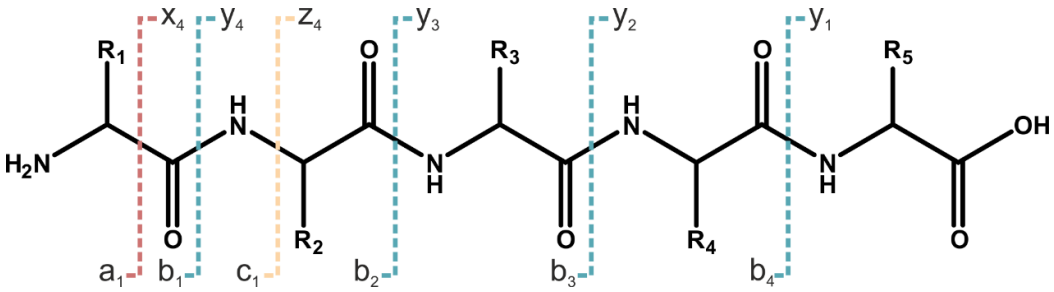


Figure 1.12: Protein fragmentation nomenclature. The fragmentation of proteins in the gas phase can proceed via a variety of mechanism resulting in different fragments forming. Commonly observed are *b*-*y* and *c*-*z* ion pairs (Biemann 1988; Roepstorff and Fohlman 1984). Note that the b_1 ion is not normally observed.

One of the most commonly available fragmentation techniques is CID (Jennings 1968). CID involves the collision of the ion(s) of interest with an inert buffer gas such as argon. Additional energy is provided by acceleration of the ions in an electric field. The tuning of this field can, to an extent, be used to tune the fragmentation of the ions. The collisions lead to the transfer of kinetic energy to the precursor ions where the energy is then converted into internal energy. This internal energy is distributed across the vibrational modes of the ions, a process called activation. The activated ion, should the energy prove great enough to break chemical bonds, will fragment with the weakest bonds undergoing fission first (Brodbelt 2016). In a typical protein the first bond to break will be the peptide bond resulting in the formation of b/y ions.

CID is a robust technique and has been utilised extensively in a variety of methodologies (Brodbelt 2016). CID can be used in the investigation of non-covalent interactions, protein folding/unfolding and protein sequencing. The fragmentation of peptides in CID is an ergodic process. That is, there is sufficient time between energy being gained by the peptide and the fragmentation to allow for the rearrangement of energy to the weakest bond. As such there can be significant rearrangement of chemical groups such as post-translational modifications (PTMs) or deuterium scrambling in hydrogen deuterium exchange experiments.

ETD (Syka *et al.* 2004), unlike CID, is a non-ergodic technique. Due to the rapid transmission of energy and fragmentation of a chemical bond in ergodic techniques there is little time for energy to be rearranged within the molecule. Consequently the bond that is broken will be the one where the energy is first absorbed rather than the weakest bond. As such, positional information for groups such as PTMs is maintained post fragmentation. In addition, ETD (along with ECD, together referred to as ExD) generates c and z ions as opposed to b and y ions.

Broadly, the method for ETD is as follows. Multiply charged analyte ions are generated by ESI (as is typical for protein MS experiments). A second ions source, often a glow discharge source, is used to generate anions of an electron rich reagents such as nitrosobenzene, which will provide a supply of low energy electrons. These negative ions are transferred to a holding cell. Switching the polarity of the front end of the instrument and manipulating the wave form generated in the ion guides allows the two populations to mix. Upon mixing the analyte cations and the reagent anions react, with an electron being transferred to the analyte, generating a radical cation (with a net charge one lower than the precursor analyte ion). The fragmentation mechanism of ExD is extremely complicated, may involve nearby residues side chains and is not currently well understood. The most common result is the formation of c and z ions by cleavage of the bond between the α-carbon and nitrogen. The addition of further energy may be required to dissociate the fragment ions.

ECD (Zubarev, Kelleher and McLafferty 1998) proceeds via a similar process with much higher energy electrons.

A range of peptide fragmentation techniques and the ions formed are described in Table 1.1.

Table 1.1: Gas phase fragmentation of peptides product ions generated.

Fragmentation technique	Ergodic/non-ergodic	Product ions
Collison induced dissociation	**Ergodic**	**b- and y-**
Surface induced dissociation	**Ergodic**	**b- and y-**
Electron transfer dissociation	**Non-ergodic**	**c- and z-**
Electron capture dissociation	**Non-ergodic**	**c- and z-**
Charge transfer dissociation	**Non-ergodic**	**a- and x-**

1.1.5.1 Arrangement of analysers

A typical arrangement of analysers in an MS/MS instrument would be a quadruople followed either by another quadrupole (the ubiquitously applied tandem or triple quadrupole instrument) or a ToF. The scanning mode of quadrupoles makes them especially well suited to the role of a *m/z* filter. Other instrument sets ups are available such as ToF/ToF and coupled ion trap instruments.

There are a number of scan modes in which an MS/MS instrument can function. These are described in Figure 1.13. In all cases the analysers are arranged in sequence, separated by a fragmentation region.

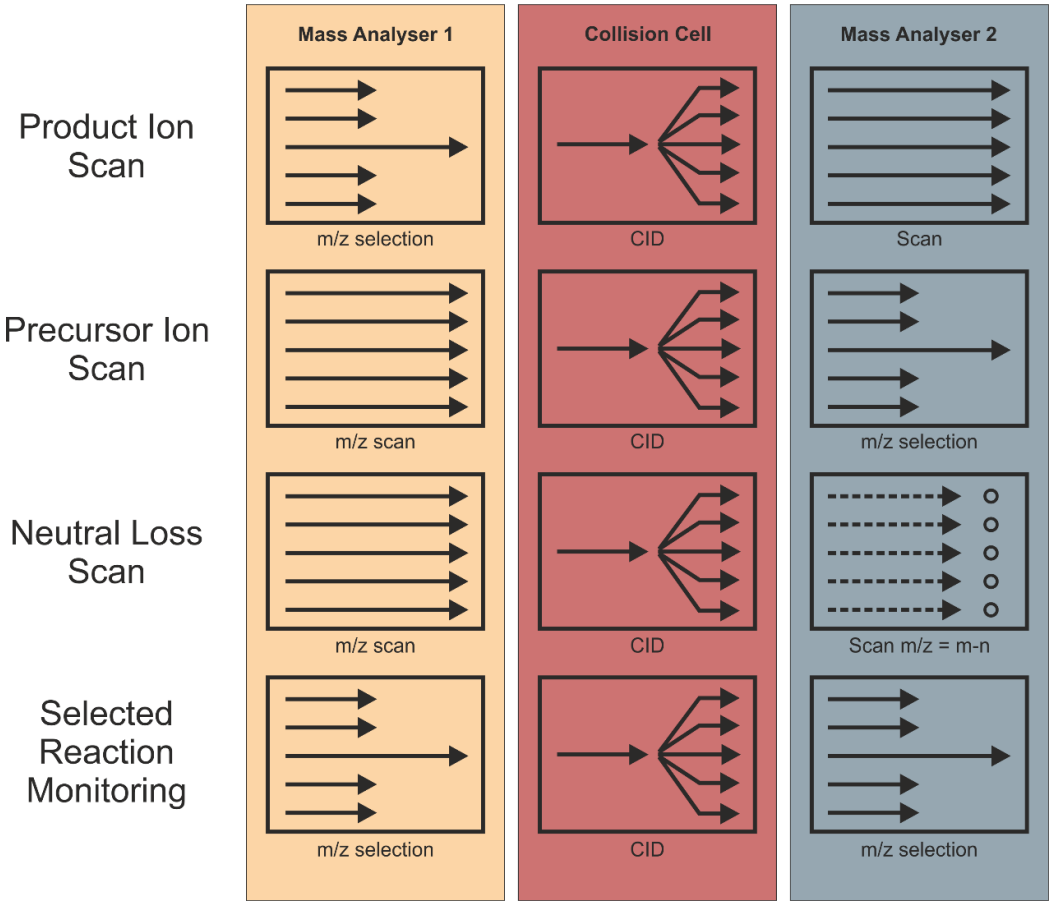


Figure 1.13: Scanning MS/MS modes. MS/MS typically involves two mass analysers separated by a collision cell. In product ion scanning an ion is selected in the first analyser and all product ions recorded in the second mass analyser. In precursor ion scanning all ions are scanned and a single product ion recorded in the second mass analyser. In neutral loss scanning all ions are recorded in the first mass analyser and a specific mass reduction recorded in the second mass analyser. In selected reaction monitoring a specific precursor and a specific product ion are selected.

In product ion scanning a single *m/z* species is selected in the first analyser and subsequently fragmented in the collision cell. The second analyser is employed to observe all resulting product ions. This mode is often used in peptide sequencing mass spectrometry and structural studies.

Precursor ion scanning involves the first analyser operating in scanning mode, with a wide *m/z* range and the second analyser is set to a specific *m/z*. This allows screening for precursors which generate a specific product e.g from a post-translational modification.

For a neutral loss scan, both analyser 1 and analyser 2 are set to scan. However analyser 2 is set to detect ions which correspond to the loss of a specific *m/z* from

the ions detected in the first scan. As neutral loss events generating uncharged molecules that cannot be observed by MS are common for the loss of small chemical groups, this method can be used to identify compounds with specific functional or chemical groups e.g. a series of carboxylic acids.

Selected reaction monitoring scans require both analysers to be set to a specific *m/z*. The ion selected in the first analyser is fragmented and the fragments observed in the second analyser. This process maintains very high selectivity as required in the identification of specific analytes in populations of closely related compounds such as in metabolomics and drug monitoring in sporting events.

1.1.6 Ion mobility spectrometry

Traditional mass spectrometry allows the measurement of the *m/z* of an ion and thus the elucidation of the mass of the molecule from which the ion was generated. This provides information on the chemical structure of a molecule. However, increasingly there is interest in not just the molecular formula or chemical identity of a molecule but also in its size and conformation. Ion mobility spectrometry (IMS) can provide this additional information.

Ion mobility spectrometry, which can be implemented in a variety of forms, separates ions based on the time they take to cross a gas-filled chamber. This drift time is dependent upon the charge of the ion (which determines how it experiences the electric field), the ion's mass (which determines its velocity) and the collisional cross-section of the ion (which determines how often it interacts with the buffer gas). The drift time of an ion can be converted into a cross-sectional area which can provide information about the three dimensional structure of the ion analysed.

While linear IMS (Figure 1.14) has existed in its own right for many years, in combination with MS it has become a very powerful tool for both structural analysis of molecules and as an orthogonal separation technique for complex samples. A variety of IMS-MS instrument set ups have been implemented including (high) field asymmetric ion mobility spectrometry (FAIMS), linear drift

tube IMS and Travelling-wave IMS (TWIMS). The last two of these techniques have been used extensively in the investigation of protein native structure and are discussed in detail below. FAIMS has also seen some use in structural and conformation studies but is not discussed here.

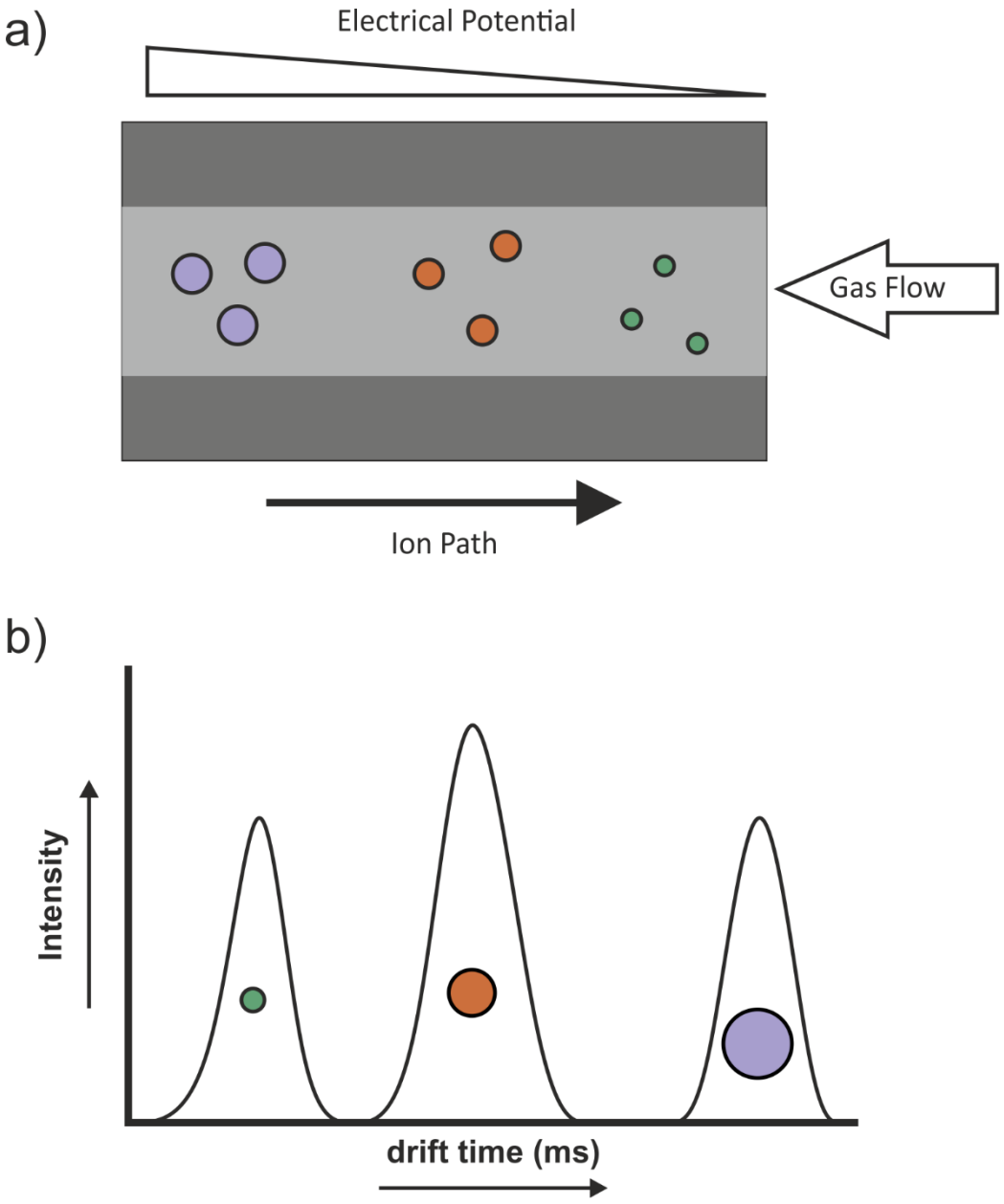


Figure 1.14: Linear ion mobility spectrometry – mass spectrometry. (a) A linear drift cell IMS. An electrical potential is applied across a gas filled region and the time taken for ions to cross it recorded. With the same mass mass and charge, more expanded ions undergo a greater number of collisions and thus their progress is retarded to a greater extend that more compact ions. (b) A drift time plot for the three ions shown in (a). The coloured circles represent ions of the same m/z but differing CCS (from largest to smallest blue, red and green).

A linear drift tube IMS device consists of a gas-filled chamber across which a constant electrical potential is applied. Ions are trapped at the entrance to the IMS cell. When the trapped ions are released the time taken for them to cross the chamber is measured. The mobility of an ion, K, is related to its velocity, V_D and the electric field it experiences, E.

Equation 1.8:

$$K = \frac{V_D}{E}$$

The length of the drift tube, L, is known and the drift time, t_D, can be observed so the equation becomes:

Equation 1.9:

$$K = \frac{L}{t_D E}$$

To account for differences in temperature and pressure the mobility is often expressed as reduced mobility, K_0, the value normalised to standard pressure (760 Torr) and temperature (273.3 K).

Equation 1.10:

$$K_0 = K\frac{273.2}{T} \times \frac{P}{760}$$

$$= \frac{L}{t_D E} \times \frac{273.2}{T} \times \frac{P}{760}$$

Where T is temperature in Kelvin and P is pressure in Torr.

The Mason-Schamp equation (Equation 1.11) (Mason and Schamp 1958) describes the relationship between the mobility of an ion and its collisional cross section, Ω. Also taken into account are the charge of the ion, *z*, the elementary charge, e, the number density of the buffer gas used, N, μ the reduced mass of the ion and buffer gas and the Boltzmann constant, k_B.

Equation 1.11:

$$K = \left(\frac{3ze}{16N}\right)\left(\frac{2\pi}{\mu k_B T}\right)^{\frac{1}{2}}\frac{1}{\Omega}$$

For reduced mobility the equation can be written as:

Equation 1.12:

$$\Omega' = \frac{(18\pi)^{\frac{1}{2}}}{16}\frac{ze}{(k_B T)^{\frac{1}{2}}}\left[\frac{1}{m_i}+\frac{1}{m_n}\right]^{\frac{1}{2}}\frac{760}{P}\frac{T}{273.2}\frac{1}{N}\frac{t_D E}{L}$$

m_i and m_n represent the masses of the analyte ion and the buffer gas respectively. It can be seen from the equation that the larger the drift time the larger the collisional cross section.

The ability of linear drift tube IMS to determine the cross-section of an ion from experimental data from first principles is a powerful use of IMS and has been widely employed in the laboratory (Barr *et al.* 2016; Bleiholder *et al.* 2013b; Servage *et al.* 2015; Shelimov and Jarrold 1995). However, linear IMS devices have suffered historically from poor ion transmission and consequently reduced sensitivity when compared to their non-IMS counterparts. In addition, the primary limit of the resolving power of a linear IMS is the length of the chamber, much like ToF. Finally, until recently, commercial linear IMS devices have been very rare resulting in most systems being homemade (though Agilent now produce a commercial linear drift tube IMS).

In the early 2000s TWIMS (Figure 1.15) was developed (Giles *et al.* 2004; Pringle *et al.* 2007). A TWIMS cell consists of a series of stacked ring ion guides. Each ring can have a different potential applied to it allowing for the electrical field to be altered both along the axis of the ion beam and over time. Indeed the stacked rings also allow for shaping of the wave to alter the electrical conditions the ions experience.

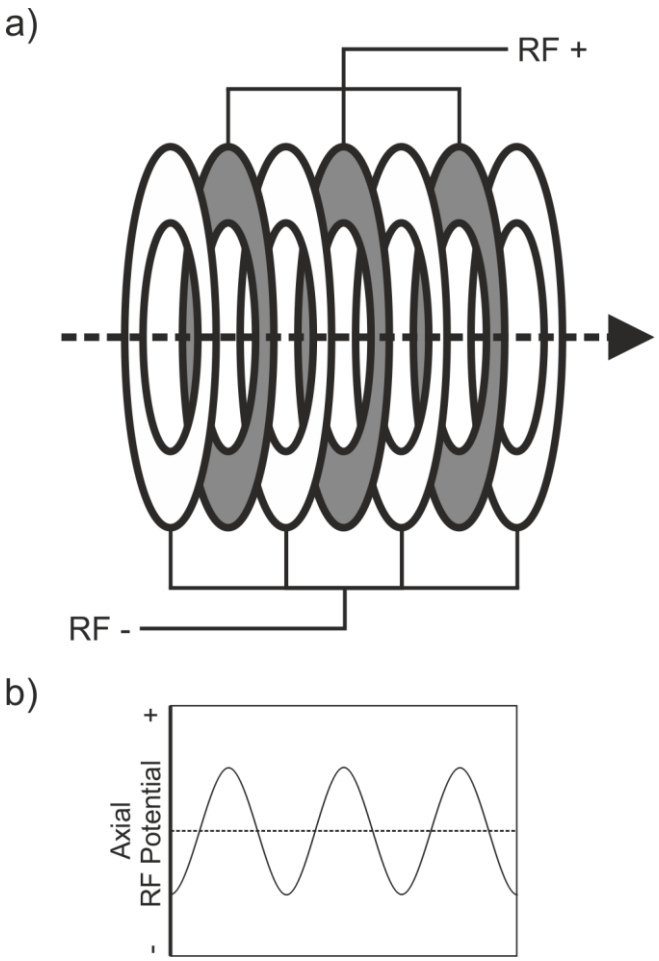


Figure 1.15: A schematic of a traveling-wave IMS cell. The TWIMS cell is a stacked ring ion guide. Alternating rings are paired and an RF alternating electrical potential is applied. This generates a wave form shown in (b) with ions being pushed along by the wave and impeded by the buffer gas.

A constant RF potential is applied to the TWIMS cell to focus and confine the ions and thus minimise the loss of ions that occurs in some other IMS techniques. A DC voltage can be applied to the stacked rings in sequence generating a wave of electrical potential that travels down the cell. This travelling wave (T-wave) provides the ions with energy to escape the potential wells created by the radially confining RF potential. The amplitude of this wave, often referred to as wave height determines whether an ion will successfully traverse the TWIMS cell.

The TWIMS cell is filled with an inert buffer gas, usually argon or nitrogen, at a pressure of around 0.5 mbar (Pringle *et al.* 2007). As the analyte ions are driven through the cell by the T-wave they interact with this gas and experience drag due to the collisions (Figure 1.16a). As in a linear IMS cell, this drag retards the

movement of the ions, however, unlike in a linear IMS cell the ions can roll over the top of the T-wave. The ions which roll over can then interact with the next wave. Ions with a more extended conformation will experience a greater number of rollover events than a similarly charged ion with a smaller collisional cross section. Therefore the more extended ions will take longer to traverse the TWIMS cell and reach the detector (Figure 1.16b).

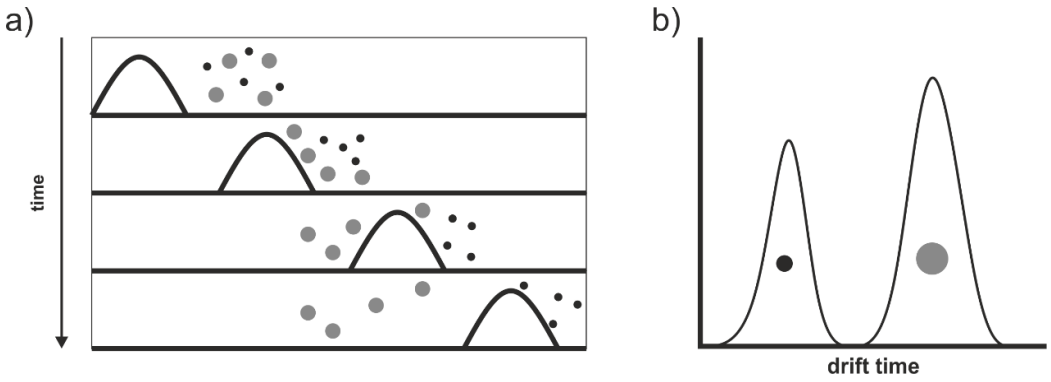


Figure 1.16: The effect of a T-wave on ions. (a) As the T-wave moves along the cell ions are pushed ahead of it. More expanded ions experience more collisions with buffer gas and thus "roll over" the wave and are separated from more compact ions. These ions will then be accelerated by the next wave. The result is a drift time plot shown in (b). Grey circles represent ions with the same m/z but an expanded and compact CCS for the grey and black circles, respectively.

One major drawback of TWIMS is that the non-uniform nature of the T-wave process makes it very difficult to determine mathematically the collisional cross-section of an ion based on its mobility (Shvartsburg and Smith 2008). To overcome this issue a calibration with ions with known CCS is required (Ruotolo *et al.* 2008; Smith *et al.* 2009). In the case of proteins these ions are typically commonly available proteins which have had their CCS determined by linear IMS. Recent work has proposed that it is possible to determine the CCS of a protein without the requirement for calibration (Mortensen, Susa and Williams 2017).

The equation for reduced mobility can be modified to take into account the non-uniform field as shown in equation 1.13.

Equation 1.13:

$$\Omega = \frac{(18\pi)^{\frac{1}{2}}}{16}\frac{ze}{(k_B T)^{\frac{1}{2}}}\left[\frac{1}{m_i}+\frac{1}{m_n}\right]^{\frac{1}{2}}\frac{760}{P}\frac{T}{273.2}\frac{1}{N}X{t_D}^{y}$$

Where *X* and *y* are correction factors experimentally determined from the CCS standards described above. *X* accounts for the non-uniform nature of the electric field while *y* accounts for its non-linearity. The CCS can be made independent of mass and charge by dividing by *ze* and the reduced mass

Equation 1.14:

$$\Omega' = \frac{(18\pi)^{\frac{1}{2}}}{16}\frac{ze}{(k_B T)^{\frac{1}{2}}}\frac{760}{P}\frac{T}{273.2}\frac{1}{N}X{t_D}^{y}$$

The equation can be further simplified to have a single constant, X':

Equation 1.15:

$$\Omega' = X'{t_D}^{y}$$

This equation can be treated in the format y=mx+c as a linear relationship. If Ω' is plotted against t_D for the CCS standards a calibration curve can be constructed for analytes of unknown CCS. The reduced CCS from this calibration can then be converted into a CCS

Equation 1.16:

$$\Omega = ze\left[\frac{1}{m_i}+\frac{1}{m_n}\right]^{\frac{1}{2}}\Omega'$$

TWIMS has been used in many investigations of protein structure. The large number of protein structures determined in atomic detail by x-ray

crystallography and other techniques can be compared to the experimentally determined CCS data by estimating a theoretical CCS from the Protein Data Bank (PDB) files.

For proteins with a known structure or with a structural model (including intrinsically disordered proteins), the expected CCS can be calculated computationally. This allows comparison between the structure of a gas phase ion of a protein and a known structure from the PDB. In addition, models can be generated for theoretical complex arrangements and thus validated or rejected based on observed data (Ruotolo *et al.* 2005; Smith, Radford and Ashcroft 2010).

There are several CCS estimation methods available which vary in the manner in which they model interactions with the buffer gas (Benesch and Ruotolo 2011; Knapman *et al.* 2010; Marklund *et al.* 2015; Mesleh *et al.* 1996).

The most accurate method of predicting the CCS of a given protein from its structure is the trajectory method (TM) (Mesleh *et al.* 1996). The TM models the collisions of individual atoms of the protein with the buffer gas. The trajectory method is very computationally intensive and thus is impractical to use in many instances.

The exact hard spheres scattering (EHSS) model treats all atoms (both the buffer gas and the protein) as spheres (Shvartsburg and Jarrold 1996). The interactions of these spheres are then modelled. The result is a much reduced computational load.

The projection approximation (PA) algorithm is far less computationally intensive and is reported to provide a similar level of accuracy to the TM (Benesch and Ruotolo 2011). CCS estimation by this method works by calculating a series of 2D projections of a structure at various rotations. These projections are then averaged to produce a rotationally averaged CCS. This process is based on the widely accepted theory that ions tumble randomly in the gas phase.

The projection superposition approximation (PSA) method developed by Bleiholder and co-workers functions in a similar manner to the PA however it also applies corrections for atomic potentials and takes into account the effect of the molecules shape on collision frequency (Bleiholder, Contreras and Bowers 2013; Bleiholder *et al.* 2013a; Bleiholder, Wyttenbach and Bowers 2011). Likewise the IMPACT software generates estimated CCS in a much shorter time than the TM (Marklund *et al.* 2015).

1.1.7 Application of MS to the investigation of proteins

Mass spectrometry can be applied to the investigation of a wide variety of molecules. One such area and the focus of the work described here, is the study of protein complexes and structural biology. With modern instrumentation the study of protein primary structure is, in many cases, fairly routine. Techniques for the study of the higher order structure of proteins have been developed more recently and rely primarily on advances in electrospray ionisation to make them possible.

MS has been employed to study proteins from small peptides through to multi-megadalton complexes (Leney and Heck 2017; Snijder *et al.* 2013), a mass range that compares favourably with the limited ranges of both electron microscopy (EM) and nuclear magnetic resonance spectroscopy (NMR), which are limited to very large and relatively small proteins respectively. In addition, MS is capable of accurately assigning species in highly heterogeneous and dynamic systems which can cause issues for EM and X-ray crystallography. Finally, the sensitivity of modern MS techniques facilitates the investigation of proteins with low expression levels or from small samples, thus overcoming the large sample requirements of some other structural techniques.

In general the MS techniques employed to investigate the structure of proteins can be split into two categories: native MS and non-native MS techniques, often referred to as chemical footprinting techniques (Figure 1.17).

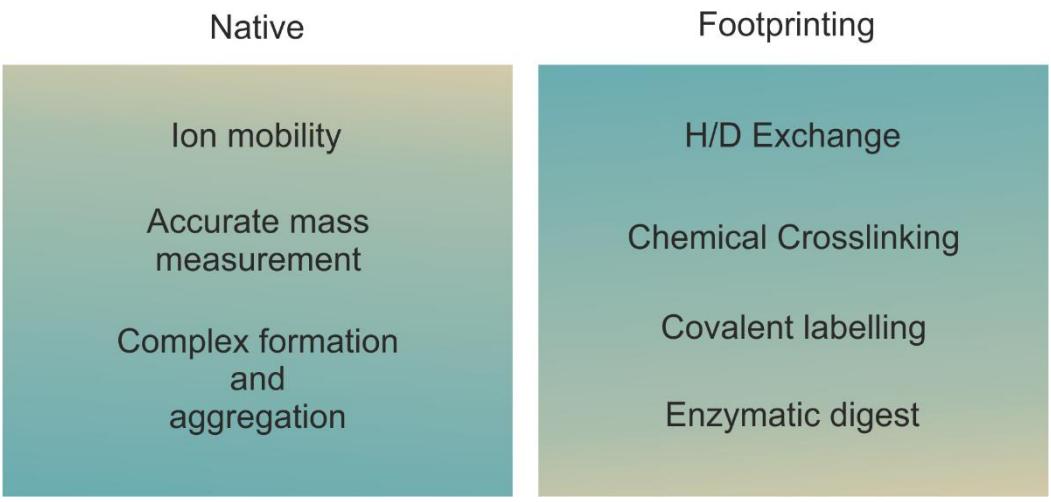


Figure 1.17: MS techniques for studying proteins. MS techniques for the investigation of proteins can be broken down into two broad groups. Native MS techniques attempt to maintain the native like structure of the protein in the gas phase. Foot-printing techniques label the protein in its native form in solution and then investigate the modifications in the gas phase.

1.1.7.1 Native MS techniques

ESI of proteins, as with other molecules, usually generates multiple charge states of a single molecular species. Several groups have demonstrated that there is a link between the solvent accessible surface area of a protein and the average number of charges an ion of that protein will carry (Kaltashov and Mohimen 2005; Konermann 2007; Li *et al.* 2016; Testa, Brocca and Grandori 2011; Testa *et al.* 2013). The more extended a protein is the greater the accessible surface area and the greater the number of sites exposed which can carry charge.

Historically, most protein samples were examined under denaturing conditions (Figure 1.18a). The addition, for example, of organic solvents and acids to the protein solution can aid in the ionisation of the analyte. Denaturing conditions are still used in experiments that aim to determine accurate mass or for the investigation of peptides in LC-MS experiments.

Data more relevant to the analysis of biological systems can be acquired under native or non-denaturing conditions (Figure 1.18b). By carefully controlling the pH and avoiding denaturants, such as organic solvents, native-like structures can be maintained in the gas phase. The exact conditions required will vary between

proteins and the native environment of the protein must be considered, for example a lysosomal protein may be considered native at pH 4, while the same conditions for an extracellular protein would likely be considered denaturing. Exactly how similar the gas phase native ion structure is to that of the solution species is still debated. Recent developments such as IMS, gas phase infra-red spectroscopy (IR) (Seo *et al.* 2016; Seo *et al.* 2017) and gas phase HDX (Beeston *et al.* 2015; Rand *et al.* 2009) provide evidence that the solution phase structures, tertiary and secondary structure respectively, of at least some species are maintained in the gas phase under appropriate native conditions. There is, however, a growing body of evidence that not all proteins maintain their solution structure in the gas phase and as such CCS data should always be viewed with at least some degree of caution (Devine *et al.* 2017).

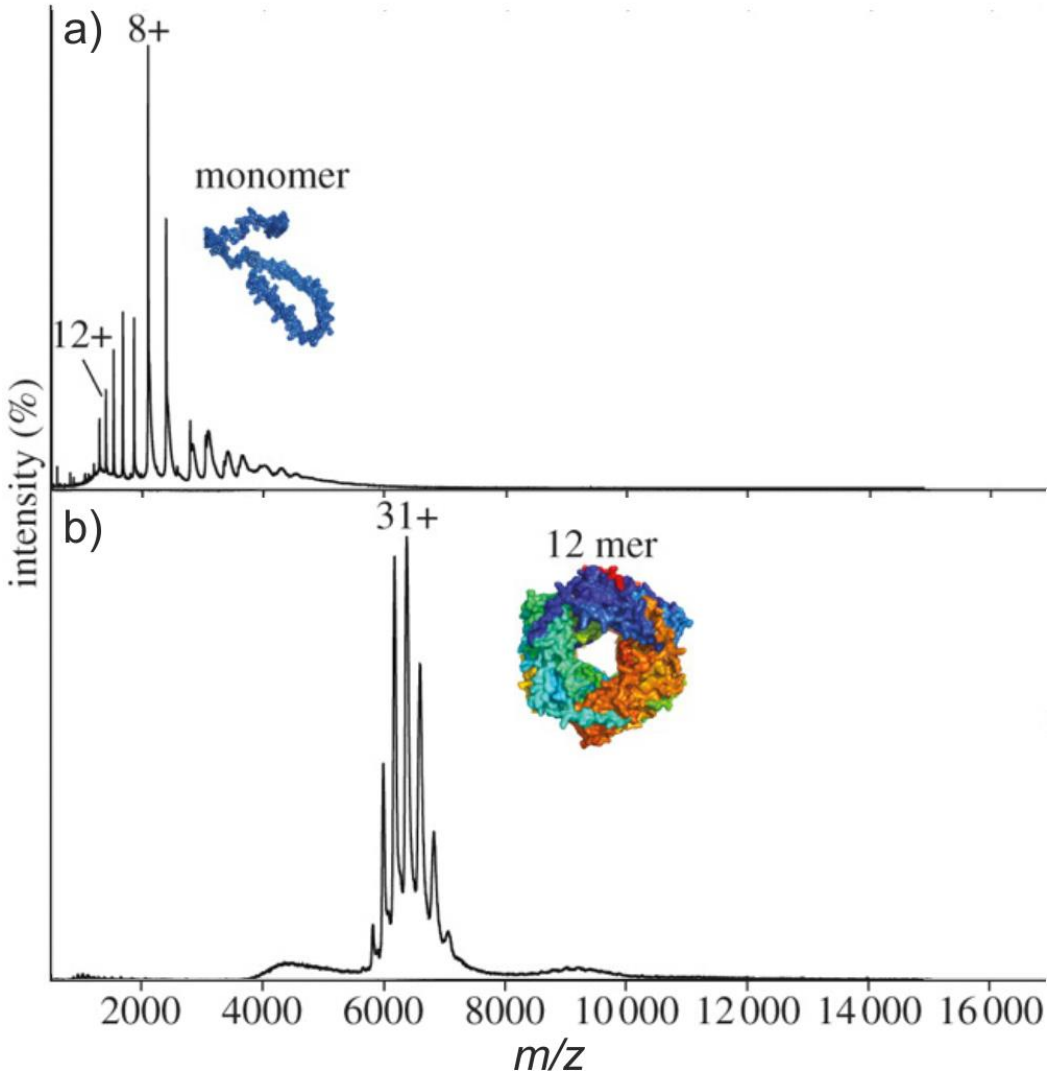


Figure 1.18: Non-native and native MS spectra of the protein TaHSP16.9. (a) Non-native MS results in the observation of a monomeric charge state distribution. (b) Native-MS results in the observation of an intact complex formed from 12 TaHSP16.9 subunits. Figure adapted from (Hilton and Benesch 2012).

Native MS in particular requires careful sample preparation. Folded or partially folded proteins have a tendency to adduct small ions such as sodium and small molecules. These adducts can result in the broadening of a peak for a given ion making accurate mass determination difficult. Dialysis or buffer exchange of samples prior to MS can alleviate this issue to some degree. Once a sample has entered the mass spectrometer the application of increased collision energies can also dissociate these adducts, however, this should be done sparingly to prevent the additional energy from altering the structure of the protein ions.

The link between the average charge state of an ion and the structure of the protein has led to the suggestion that charge state distributions can provide information on a protein's structure. It is common to observe multiple charge state distributions for a single protein. Deconvolution of the charge state distributions can reveal otherwise hidden conformations and can also be used to assign relative abundances to each conformation (Borysik, Radford and Ashcroft 2004) (Dobo and Kaltashov 2001). IMS is increasingly replacing this kind of analysis where it is available.

In addition to observations on protein conformation, native MS can be used to investigate non-covalent interactions. Native MS has been used to investigate protein:protein and protein:ligand interactions extensively (Benjamin *et al.* 1998; Rosati *et al.* 2014; Shepherd *et al.* 2013; Soper-Hopper, Eschweiler and Ruotolo 2017; Yen *et al.* 2017). Considerable thought must be given to the conditions under which such experiments are undertaken. For example the collision energy in an instrument can cause dissociation of non-covalent complexes and thus cause them not to be observed.

This phenomenon of CID can be harnessed in the investigation of the relative strengths of protein interactions. By increasing the collision energies incrementally and recording the abundance of the complex ion at each energy, a dissociation curve can be constructed. A similar curve can be generated for protein unfolding and is known as collision induced unfolding (CIU) (Ruotolo *et al.* 2007).

Native MS screening methodologies have been established which harness the considerable advantages of MS in terms of sample requirements and processing speed, to look for compounds which bind to a given substrate or to identify compounds which cause a specific effects such as preventing oligomer formation and so aid drug discovery (Chen *et al.* 2015; Young *et al.* 2016; Young *et al.* 2015).

IMS-MS investigations of proteins and protein complexes were pioneered in the 1990s. However, with the commercialisation of IMS-MS technologies the field has

grown rapidly. The IMS derived collisional cross-sections of many proteins have been demonstrated to be compatible with those predicted from the native structures of the protein (Bush *et al.* 2010; Smith *et al.* 2009). A general trend in which higher charge states exhibit larger cross-sections has been attributed to Coulombic repulsion between the charges on these ions (Tolmachev *et al.*).

Recent examples of the use of IMS-MS in the investigation of proteins and protein complexes include studies of the aggregation of fibril forming proteins (Cole *et al.* 2015; Smith *et al.* 2011; Wang *et al.* 2016a), chaperone activity (Lai *et al.* 2017; Schiffrin *et al.* 2016) and protein:protein complex formation (Niu *et al.* 2016).

More recent work has shown that while many proteins' CCS do agree with the predicted values for a sub-population of proteins the CCS predicted is much larger than is observed by IMS (Devine *et al.* 2017). This phenomenon is believed to be due to the collapse of certain structures in the gas phase. As a general rule compact globular proteins are the least likely to undergo large scale gas phase collapse while more extended proteins are more likely to experience this effect.

1.1.7.2 Non-native/Footprinting MS techniques

Native MS techniques can provide information on the conformation of a protein and constituents of a protein complex in the gas phase. Protein foot printing techniques can provide information on the characteristics of the protein or complex in solution. These techniques include hydrogen deuterium exchange (HDX), covalent labelling techniques, enzymatic digests and chemical cross linking. The resolution of these techniques varies from residue level in some cases to peptide level in others. While not the atomic level resolution of some other structural biology techniques, protein footprinting can provide beneficial information in many cases.

In general, although not exclusively, in order to get the best resolution possible for a footprinting technique some form of enzymatic digest is performed to fragment the protein into shorter peptides which can be more easily analysed by MS. The digested peptides are then separated by liquid chromatography (LC), a

technique that is easily paired with continuous ion sources such as ESI. LC systems commonly include high pressure-LC (HPLC) or ultra-performance liquid chromatography (UPLC) separation although size exclusion chromatography (SEC) and other chromatography techniques have also been used.

Hydrogen-deuterium exchange (HDX) exploits the difference in mass between H and D and the exchange of exposed H/D positions with solvent H/D over time. The mass shift that results from an exchange event, for example +1 Da if a single H is exchanged for a solvent D, can be detected by MS. While this shift can be observed in intact proteins it is generally more useful to examine peptides so that the exchanged position can be localised. There are three types of exchangeable H/D in proteins. Side chain hydrogens exchange on a very fast time scale and so are not generally of interest in the investigation of dynamics. Heteroatom-bound hydrogens are also exchangeable. The hydrogens on backbone amides are the most interesting in HDX MS as they exchange reversibly on a slower time scale than the side-chain hydrogens (Cao, Burke and Dennis 2013). The rate of exchange of backbone amide hydrogens is measurable on the timescale of a ESI-MS experiment allowing real-time and temporally defined experiments to be undertaken.

The reversibility of the H/D exchange can be problematic as information can be lost if the exchanged position exchanges back prior to the peptide being analysed. Most HDX experiments are monitored by using LC-MS. Once the exchange step has occurred the samples must be handled carefully and quickly to minimise back exchange. Precautions include keeping the sample at low temperature and pH. These requirements for temperature and pH mean that the choice of enzyme for digestion of the proteins is limited; pepsin is a common choice due to its high activity at pH 3 where the rate of exchange for labile hydrogens is low, though other enzymes have also been used.

Changes in HDX do not necessarily indicate binding sites of ligands as allosteric changes are largely indistinguishable from such structural rearrangements by HDX. In addition, while residue level information is theoretically possible it is

more common to present peptide level data. This is in part because many of the techniques for fragmenting peptides, such as CID, can cause rearrangement of the labile H/D atoms, scrambling the original locations. Despite this limitation and those described above, HDX-LC-MS/MS is a very popular technique for examining protein conformation and interactions.

Many of the limitations of HDX-MS (sample conditions post-labelling, scrambling of H/D position) can be overcome by using a covalent label of solvent exposure. One method of achieving this is fast photo oxidative labelling of proteins (FPOP) (Xu and Chance 2007). In FPOP, reactive ●OH radicals are generated in solution e.g. from hydrogen peroxide with the protein of interest (Figure 1.19). The ●OH radicals can then interact with the side chains of the residues which are exposed to the solvent.

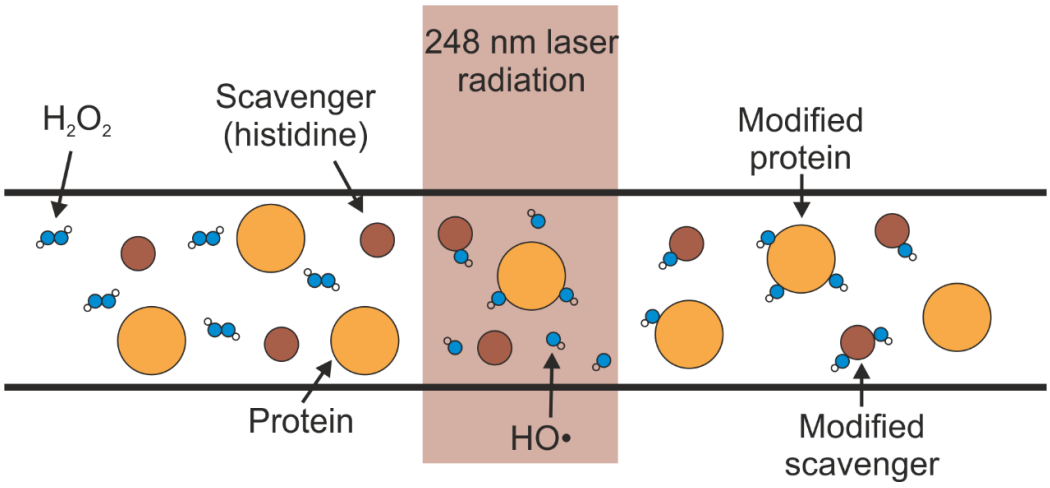


Figure 1.19: A schematic of a flow based UV-excimer laser FPOP experiment. A protein in solution with H_2O_2 and a scavenger molecule, such as histidine, flows through a capillary. At a certain point the capillary is exposed to a UV-excimer laser which causes photolysis of the H_2O_2 generating ●OH radicals. The ●OH radicals react with exposed amino acid side chains or with the scavenger molecules. The reaction time is generally in the µs range although it can be tuned based on scavenger concentration and other characteristics. The covalently modified protein can be collected and analysed offline by MS.

●OH radicals can be generated in a number of ways. The most common are laser photolysis of H_2O_2 with a 248 nm UV excimer laser (Sharp, Becker and Hettich 2004) and X-ray radiolysis of water (Asmus 1984; Hayes, Kam and Tullius 1990;

Sclavi *et al.* 1998). In the case of laser photolysis, the generation of radicals takes only a few nanoseconds and can be performed in low mM concentrations of hydrogen peroxide (10-20 mM). A typical experimental set up places the sample flowing through a thin capillary with a pulsed laser exposing each region to a single pulse. The life time of the ●OH radicals in solution can be tuned by the addition of scavenger molecules, such as glutamine or histidine. It is believed that most FPOP setups are probing µs windows of protein conformation.

Subsequent to their generation, the ●OH radicals can react with the side-chains of amino acids. The reactivity of each side-chain is different with a significant drop of in reactivity after the five most reactive (cysteine, methionine, tryptophan, tyrosine and phenylalanine) (Konermann, Tong and Pan 2008; Takamoto and Chance 2006; Xu and Chance 2007). ●OH radical reactions can lead to a number of different modifications including +16, +14, +32 and a series of complex reactions with histidine residues. All of these reaction products can be detected by MS. When residues which are either very poorly reactive or undergo side reactions are excluded, 14 residues are able to be labelled by FPOP and provide coverage of a predicted 65% of an average protein (Wang and Chance 2011). The large size of these modifications relative to HDX makes observation possible even in intact proteins. The stable nature of the covalent modifications means that the subsequent digestion and analysis of the protein can be more varied than is the case for HDX. As the chemical groups generated in FPOP are generally not rearranged during fragmentation, FPOP can provide residue level information far more easily the HDX.

1.1.7.3 Introduction to the study of protein structure and interactions by means of chemical crosslinking and mass spectrometry

While native mass spectrometry has been used extensively in the investigation of protein structure and complexes (as described in the Chapter 1.1.7.1), it is limited in its resolution when examining sites of interaction. In addition, in some cases such as described below, the complex of interest may be difficult to manipulate in the gas phase. In instances such as these, chemical crosslinking followed by

mass spectrometry (crosslinking MS) strategies can be a powerful tool to investigate the structure of a protein or protein complex.

Chemical cross-linkers include a wide variety of different molecules. In general they are bifunctional (i.e. reactive at two sites) and form covalent bonds to one or more reactive sites of a peptide. The cross-linker can react to form a bridge between two residues that are in close proximity in solution, either within a single protein or between two different proteins. This can provide structural information on the protein and can also indicate sites of close interaction between binding partners. Each cross-linker has a specific distance between the reactive groups which determines the maximum distance apart that two cross linked residues can be in space. In some cases this length can be modified with changes to the length of the linker region. The linker lengths can form a distance restraint when modelling a protein or protein complex.

As with other covalent techniques, cross linked samples are fairly robust once the initial reaction is complete. Usually a quenching solution is used to prevent any further reactions. The nature of the quench is dependent upon the cross linker used. There can be issues with false negatives due to the absence of reactive residues in the site of interest. This issue can be alleviated by use of non-specific cross-linkers or cross-linkers with reactivity complementary to the first reagent used.

In a general crosslinking MS investigation of a protein complex, the protocol used may resemble that shown in Figure 1.20. A sample of the protein or proteins of interest is prepared in a suitable buffer to which is added a crosslinking reagent, a selection of which are discussed in detail below. After a reaction period (in some cases under specific conditions, such as exposure to UV light for photoactivatable crosslinkers) the reaction is quenched and the samples separated by SDS-PAGE. For a simple complex, A+B → AB, a band of the molecular weight of the complex will be observed. Depending upon the concentration and efficiency of the

crosslinking reagent there may also be monomer and higher order structures observed. The bands of interest (typically alongside the monomer subunits or uncrosslinked controls) are extracted, digested with a protease and subjected to other treatment depending upon the proteins in question (commonly reduction and methylation of disulphide bonds). The resulting peptides are then analysed by LC-MS/MS and the sites of any crosslinks mapped.

Figure 1.20: A simple schematic of a crosslinking MS method. A protein or proteins are mixed with a crosslinking reagent in solution. The crosslinker may require activation by exposure to UV light or other methods. The solution is incubated for a period of time and the reaction quenched, often by addition of an excess of reactive groups in a buffer component. The resulting crosslinked protein mixture is run on SDS-PAGE to separate the crosslinked proteins from each other and from non-crosslinked protein. The gel bands of interest are extracted and digested, often with trypsin. The resulting peptides are investigated by LC-MS/MS in order to identity the proteins involved and to map the site of interaction.

While a monomer band may appear in a crosslinked sample this does not mean that it is unmodified protein. There are in fact a range of possible crosslinking reactions that can occur as described in Figure 1.21. The nomenclature of these crosslinked species was formalised by Schilling *et al.* in the early 2000s (Schilling *et al.* 2003). After the initial formation of a crosslinker-protein conjugate several types of reaction can occur. The free end of the crosslinker may interact with the solvent or buffer molecules, generating type-0 or dead-end crosslinks. While these crosslinks do not provide information on the site of an interaction they can provide information on surface exposure much like other surface labelling techniques. If there is a second reactive residue close enough to the site of the initial interaction, the free reactive group may form an intra-peptidal crosslink

also known as a type-1 crosslink. If the free end of the crosslinker is close to the site of an interaction with another peptide or protein which also contains a reactive group a type-2 crosslink may form. This kind of crosslink is the type generally used in mapping the sites of interactions between subunits in a complex.

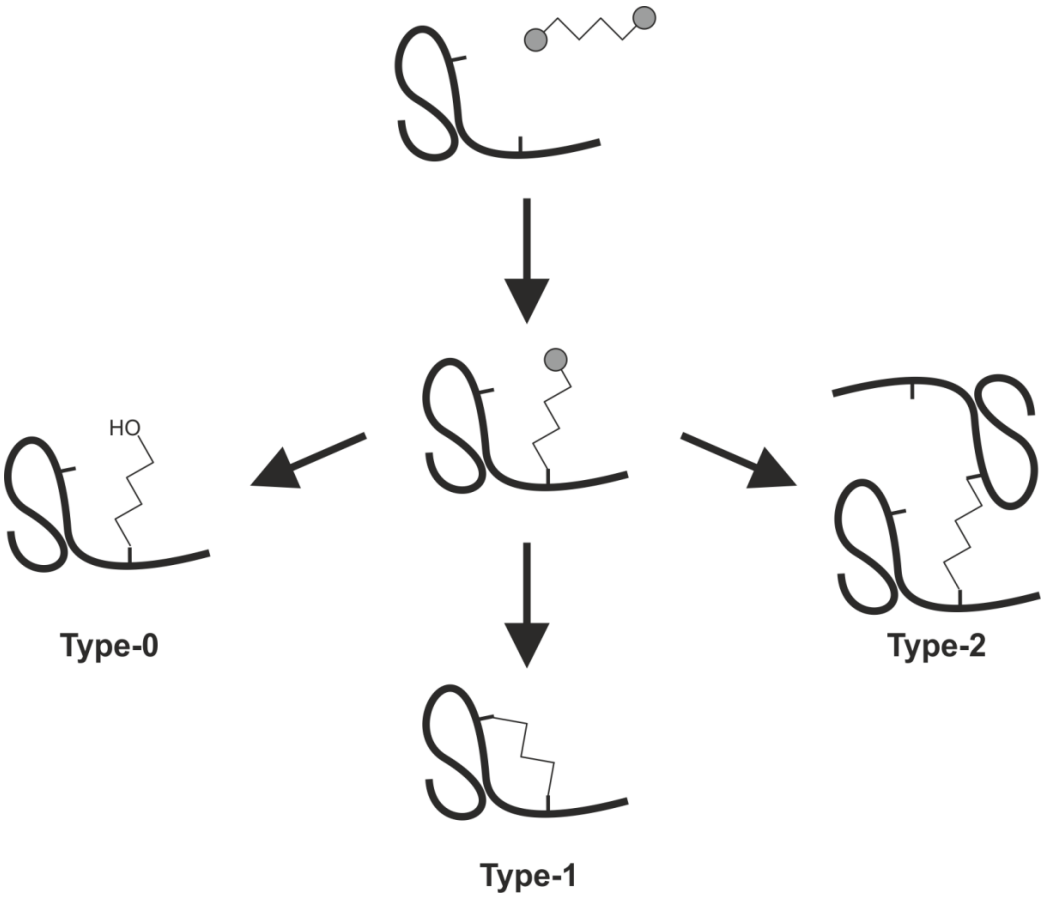


Figure 1.21: Classification of crosslinks. There are three types of crosslinks that can occur in a crosslinking experiment after the initial interaction. The reaction of the remaining reactive group with buffer components or solvent results in so called dead-end crosslinks, also known as type-0 crosslinks. The reactive group interacting with a second reactive site within the same peptide generates type-1 crosslinks. Type-1 crosslinks can provide information on protein conformation. Type-2 crosslinks occur when the reactive group interacts with a second peptide molecule, either of the same protein or a different one. Type-2 crosslinks are useful for identifying and investigating interactions between proteins and domains (Schilling *et al.* 2003).

The crosslinking reagent bis(sulfosuccinimidyl)suberate (BS^3), whose structure is shown in Figure 1.22, has been utilised extensively in the investigation of protein structure and protein:protein interactions (Leitner *et al.* 2016; Murakami *et al.* 2013; Nguyen-Huynh *et al.* 2015; Politis *et al.* 2015; Schmidt and Robinson 2014b).

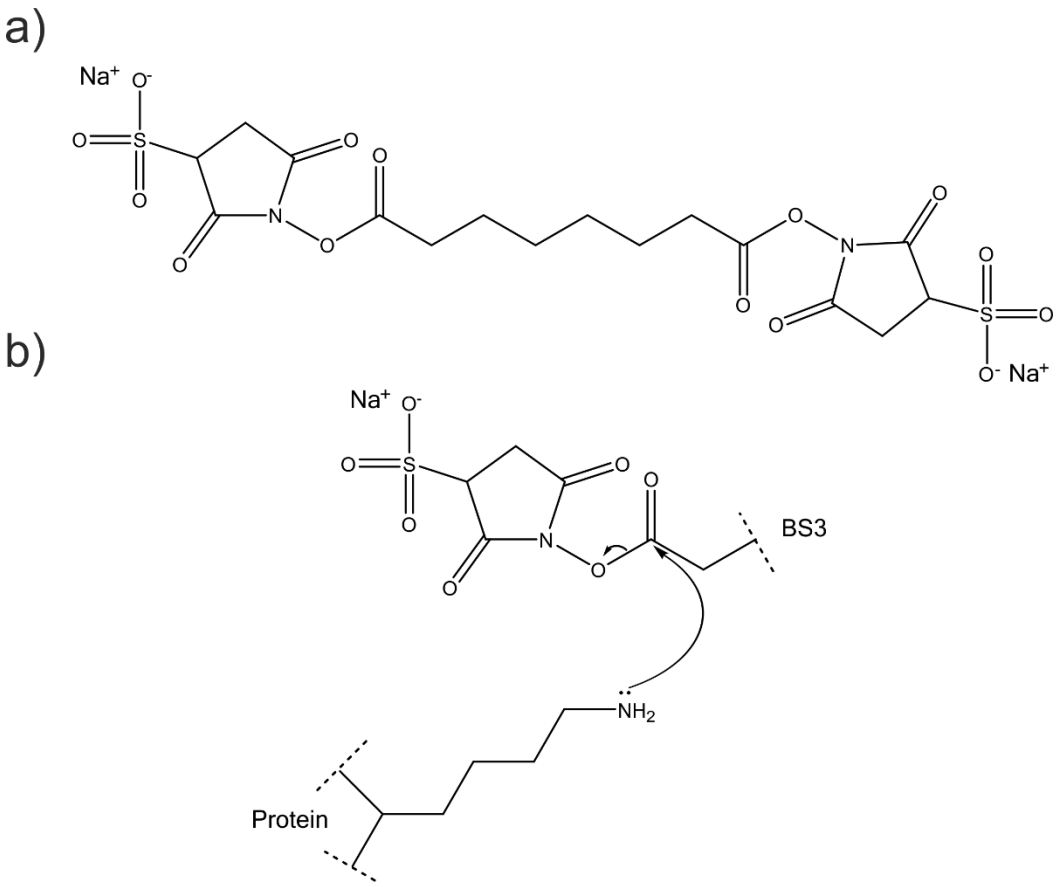


Figure 1.22: The structure and reaction mechanism of the heterobifunctional crosslinking reagent BS^3. (a) The chemical structure of BS^3. The two reactive groups are separated by a linker region, $(CH_2)_6$, 11.4 Å in length. (b) The reaction mechanism of covalently modifying a primary amine group, the most common reaction between BS^3 and proteins.

BS^3 is a homobifunctional crosslinker which reacts primary amines (i.e. lysine side chains and N-termini). The reactivity of BS^3 comes from a pair of N-hydroxysulfosuccinimide (NHS) esters at either end of the molecule. NHS esters undergo nucleophilic attack upon amide groups, resulting in the formation of an amide bond between the amide containing group and the crosslinker minus NHS (Figure 1.22b) (Leitner *et al.* 2016).

While most literature concerns itself with the reactions of NHS-esters and primary amines, there is significant evidence for other reactions. The first reaction is the hydrolysis of the NHS-ester in aqueous solutions, which has a half-life of the order of hours at pH 7.5 (Kalkhof and Sinz 2008). The outcome of this process is the generation of non-reactive species which effectively reduce the

concentration of crosslinking reagent. The rate of hydrolysis is dependent upon both pH and temperature.

NHS-esters can also react with non-amine containing amino acid side chains, in particular those containing hydroxyl groups. In one study it was estimated that as much as 20% of the total observed crosslinked species were the result of reactions with hydroxyl containing residues (Madler *et al.* 2009; Kalkhof and Sinz 2008). Madler *et al.* present a comprehensive investigation of the reactivity of NHS-esters with a range of amino acids in the context of peptides (thus removing concerns regarding the relative surface exposure of different classes of side chains) (Madler *et al.* 2009). Their work demonstrates that, in addition to amines, the hydroxyl groups of serine, threonine and tyrosine are also reactive. The abundance of these reactions with non-amine groups is dependent upon a wide range of factors including pH (generally higher pH favours amine reactions) and the peptide sequence surrounding the reactive residues (histidine and arginine in particular). Finally, the broader context of structure in which a residue finds itself also seems likely to alter the reactivity of the residue (e.g. H-bonding).

The abundance of primary amines in many proteins and the tendency of charged amino acid residues to be exposed on the surface of soluble proteins make reagents that react with these groups attractive for mapping interactions. This is also true of hydroxyl group containing residues which are often projected into the surrounding solution. The reactivity of BS^3 along with the combination of good water solubility and activity in the native pH range of many systems make it a common choice for crosslinking experiments (Leitner *et al.* 2016; Murakami *et al.* 2013; Nguyen-Huynh *et al.* 2015; Politis *et al.* 2015; Schmidt and Robinson 2014b). Due to its extensive use in the laboratory, many robust protocols exist for crosslinking with BS^3 (Rosenberg 1996; Schmidt and Robinson 2014a; Scientific 2012; Shi *et al.* 2017).

1.2 Amyloid

The aging population of the United Kingdom and other countries in the developed world has, along with other developments in medicine and public health, led to a paradigm change in the list of common causes of death. As of 2016 the most common cause of death in England and Wales was “Dementia and Alzheimer disease”, accounting for 11.6% of the recorded deaths for 2015 (Statistics 2015). It is predicted that the number of Alzheimer’s disease cases will triple in the next 40 years to reach a worldwide occurrence of 80 million new cases (International 2015).

Alzheimer’s disease is part of a group of diseases known as amyloidosis or amyloid diseases. There are more than 50 human amyloid diseases characterised by the aggregation into amyloid fibrils of nearly as many proteins (Sipe *et al.* 2016; Sipe *et al.* 2012). In each case the causative protein, clinical pathologies and outcomes are different.

The processes by which many of these diseases arise are poorly understood. In some cases mutations are present which predispose an individual to develop the disease (such as the expansion of polyQ domains in polyQ diseases (Fan *et al.* 2014)), while in others lifestyle factors appear to be the likely cause (for example the aggregation of iAPP in Type II Diabetes Mellitus (Westermark, Andersson and Westermark 2011)). In the case of Alzheimer’s disease for example, age is the strongest correlating factor although particular mutations may also affect onset (Weggen and Beher 2012). In general the pathway of amyloid formation is the misfolding and aggregation of a normally soluble protein to form increasingly large and eventually insoluble, aggregates (Glabe 2008). A conversion occurs to alter the conformation of the proteins in the aggregates which then continue to grow and are deposited as fibrils. The identity of the disease-causing agent in this milieu of species is a subject of much debate (Goure *et al.* 2014; Kayed and Lasagna-Reeves 2013; Sengupta, Nilson and Kayed 2016; Verma, Vats and Taneja 2015).

Despite these differences in cause, pathology and outcome, the structure of the amyloid fibrils which give this class of disease its name is remarkably preserved across diseases caused by otherwise unrelated proteins (Eichner and Radford 2011). The initial feature that identified amyloid fibrils was the observation that abnormal fibrillar structures observed in post mortem brain tissue stained blue with iodine, a behaviour normally associated with starch (Westermark, Johnson and Westermark 1999). In addition amyloid fibrils bind to the dye Congo red and cause apple green bifrifingance when viewed under cross-polarised light (Bennhold 1922; Divry 1927). More recent work has shown that amyloid fibrils have a similar structure, being well ordered and having characteristic cross-β architecture (Eichner and Radford 2011; Geddes *et al.* 1968). However, the solving of high resolution structures has revealed highly complex architectures with distinct morphologies even within the fibrils of a single protein (Fitzpatrick *et al.* 2013; Fitzpatrick *et al.* 2017; Paravastu *et al.* 2008; Petkova *et al.* 2005; Petkova, Yau and Tycko 2006).

The lack of understanding of the pathological process of amyloid diseases has hampered the development of therapeutics (Folch *et al.* 2016; Sacchettini and Kelly 2002; Yiannopoulou and Papageorgiou 2013). Indeed there are very few effective therapeutics currently available (Hammarstrom *et al.* 2003) and many seek to alleviate the symptoms of the diseases rather than to prevent the diseases process since the identity of the disease-causing species remains unclear (Folch *et al.* 2016; Sacchettini and Kelly 2002; Yiannopoulou and Papageorgiou 2013). Given the growing economic and social burden of care for sufferers of this class of disease, it is vital to increase the knowledge of the disease processes involved.

1.2.1 Protein Folding, misfolding and aggregation

Many proteins have a highly specific native structure to which they must fold to achieve their function. All of the information required to determine this final fold is contained in the amino acid sequence of the peptide chain (Anfinsen *et al.* 1961). The mechanism by which the nascent chain reaches its final structure is via a conformational search for the lowest accessible free energy conformation

(the structure with the lowest free energy and therefore highest thermodynamic stability). However, the Levinthal paradox (Levinthal 1969), shows that it would be impossible for even a short peptide chain to sample every possible conformation on a biological timescale. In fact, even a modest estimate suggests that such a random conformational search would take longer than the age of the universe (Martínez 2014). It was proposed therefore that the folding of a protein must occur via specific kinetically controlled folding routes. This process reduces the number of conformations that must be sampled for the protein to fold. The formation of regions of native-like structure early in the folding process can bias the rest of the conformational search in such a way as to make folding highly efficient (Levinthal 1968; Levinthal 1969; Zwanzig, Szabo and Bagchi 1992).

The so-called energy landscape of protein folding (Figure 1.23) represents the many structures possible for the nascent chain of a newly synthesised or unfolded protein being reduced over time as lower energy structures are formed with the native structure situated at the bottom of the energy well (Dill and Chan 1997; Schonbrun and Dill 2003). As ordered regions occur, they represent a local minima in the energy landscape and thus any transitions that result in a less folded structure are unfavourable.

In reality, the energy landscape is not at all smooth as depicted in Figure 1.23a. There are many local minima into which the partially folded protein can fall resulting in a rough energy landscape (Figure 1.23b). Situations arise in which the next step in folding requires the input of energy to escape from a low free energy state. These metastable species are called folding intermediates (Schonbrun and Dill 2003). It is common for a protein to have many intermediates some or all of which will be populated for a given molecule's folding process (Jahn and Radford 2008; Schonbrun and Dill 2003). Some of the local minima are in fact not on the folding pathway (Kim *et al.* 2013). Proteins that cannot gain sufficient energy to escape from one of these off pathway intermediates states can become trapped in a misfolded state (Jahn and Radford 2008). The misfolded form of a protein is stabilised by aberrant interactions between regions of the protein that do not interact in the native state (Feige *et al.* 2008; Moulick, Goluguri and Udgaonkar

2017; Neudecker *et al.* 2012). The misfolded protein often fails to perform its intended role and may display markedly different biological properties and structure to the correctly folded protein.

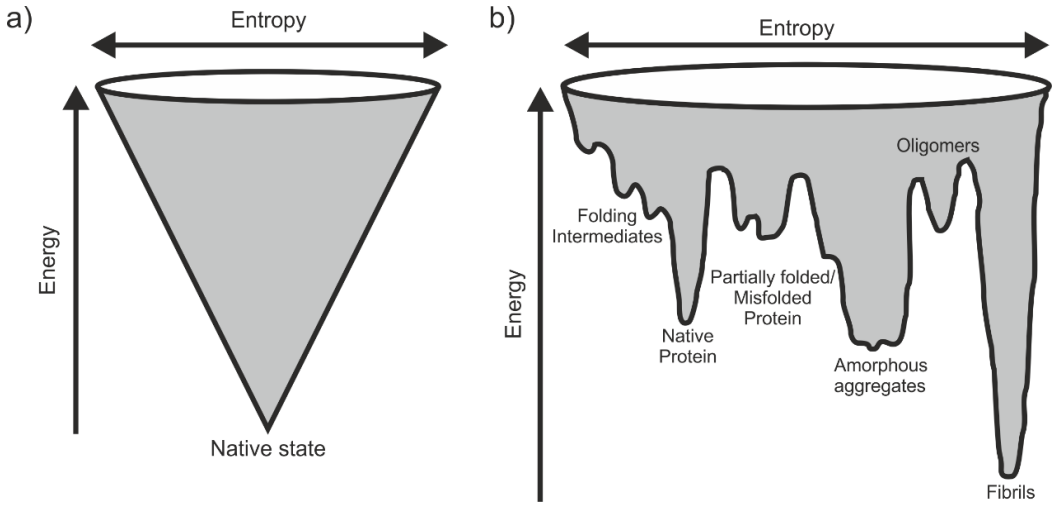


Figure 1.23: Schematic representations of protein folding energy landscapes. (a) A simplified energy landscape showing that as the protein folds towards its native state the free energy is reduced. (b) In reality, the topology of the protein folding landscape is more complicated and involves folding intermediates separated by energy barriers. In addition, the protein can become misfolded or aggregate. In many cases these aggregated species have a lower free energy than the folded state and so the protein cannot escape these energy wells making aggregates very stable. Figure adapted from (Jahn and Radford 2008) and (Schonbrun and Dill 2003).

Factors which alter the thermodynamic environment or interfere with the folding kinetics of the protein can disrupt this process and increase the likelihood of misfolded proteins being formed. Mutations in the protein or changes in the environment such as pH or temperature are examples of these factors (Chiti and Dobson 2006).

Misfolded proteins and trapped intermediates often lack the stringently evolved structure of the natively folded protein. It is common for hydrophobic regions that are normally buried to be exposed (Feige *et al.* 2008; Moulick, Goluguri and Udgaonkar 2017; Neudecker *et al.* 2012). In the aqueous environment of the cell or extracellular spaces where proteins normally find themselves, these exposed hydrophobic patches have a tendency to form interactions in order to shield themselves from solvent. These aberrant interactions can cause the protein to

aggregate, in particular with other misfolded proteins (Moreno-Gonzalez and Soto 2011).

The aggregates formed in these aberrant interactions can take a number of forms. They can consist of a single aggregating species or a mixture with other aggregation-prone proteins. Often the aggregates formed are unstructured and highly amorphous. These amorphous aggregates can self-assemble into large plaques or deposits. In other cases the aggregates self-assemble into well-ordered amyloid fibrils. In recent years it has been shown that a wide range of proteins, not just those associated with amyloid diseases, can form amyloid fibrils under certain conditions (Goldschmidt *et al.* 2010; Auer *et al.* 2008; Chiti *et al.* 1999; Guijarro *et al.* 1998; Urbanc *et al.* 2004). These proteins, like the amyloid disease-causing proteins, show no sequence homology. It has been suggested, therefore, that the amyloid fibril may represent a stable structure accessible by any polypeptide chain (Auer *et al.* 2008; Chiti *et al.* 1999; Guijarro *et al.* 1998; Urbanc *et al.* 2004) though not all studies agree (Goldschmidt *et al.* 2010).

1.2.2 Amyloid formation – mechanisms and structure

Amyloid fibril formation is a multistage process of aggregation from a soluble protein to a variety of oligomeric intermediates and eventually insoluble fibrils (Eichner and Radford 2011; Gillam and MacPhee 2013; Jucker and Walker 2011). There is evidence that the fibril is not an inert, terminal point in this process and can undergo shedding – the loss of smaller species from the end of the fibril (Tipping *et al.* 2015) (Carulla *et al.* 2005).

Fibril formation occurs via a nucleated growth process (Figure 1.24) (Arosio *et al.* 2014; Buell, Dobson and Knowles 2014). The formation of a nucleus from misfolded monomers is a rate limiting step in aggregation (Gillam and MacPhee 2013). The formation of the nucleus itself is thought to be thermodynamically unfavourable and as such is an uncommon event. After the formation of the nucleus, further addition of misfolded monomers is a thermodynamically favourable process and elongation can occur very quickly (Eichner and Radford 2011; Gillam and MacPhee 2013; Jucker and Walker 2011). The stochastic

interactions required for the formation of the nucleus result in a lag phase in fibril formation (Figure 1.24). The characteristic sigmoidal kinetic curve of amyloid fibril formation is a product of the nucleation-dependent lag phase and the monomer concentration-dependent elongation phase (Buell, Dobson and Knowles 2014).

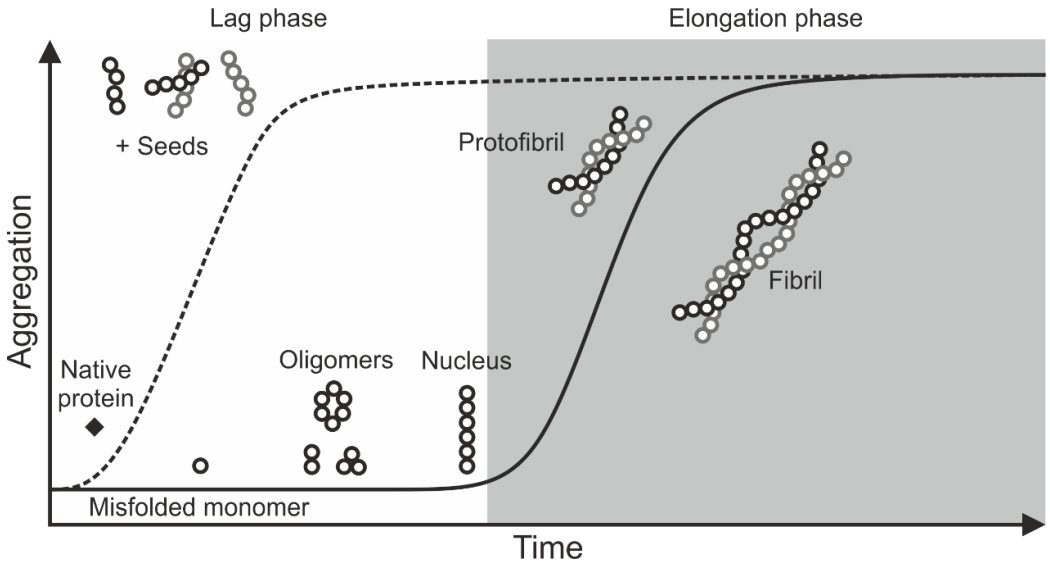


Figure 1.24: The kinetics of amyloid fibril formation and examples of species involved. Amyloid aggregation proceeds via a sigmoidal growth curve. In the lag phase misfolded monomers and/or oligomers may form. Once a critical nucleus is formed the elongation and/or propagation of the protofibrils and fibrils occurs rapidly until equilibrium is reached. The addition of seeds generated from the fragmentation of fibrils (dotted line) reduces the length of the lag phase considerably by removing the contribution of primary nucleation. Figure adapted from (Glabe 2008).

The addition of preformed nuclei or seeds (fragments of fibrils) to a population of monomer shortens the lag phase significantly (Buell, Dobson and Knowles 2014; Chatani *et al.* 2010; Hortschansky *et al.* 2005; Jucker and Walker 2011; Pinotsi *et al.* 2016). The presence of preformed nuclei means that the kinetics of seeded fibril formation are largely dependent upon elongation rather than stochastic interactions required for nucleus formation (Buell, Dobson and Knowles 2014).

The addition of fibril fragments or the fragmentation of existing fibrillar species can promote secondary nucleation (Figure 1.25) (Jucker and Walker 2011). It can be imagined that each end of a growing fibril may be extended by the addition of

a new subunit. For a population of x fibrils there will be 2x ends. If this population is fragmented in such a way that each fibril is broken into two pieces the number of available extension sites will increase the (2x)2=4x. The increase the in number of extension sites will result in a rapid increase in the amount of protein included in fibrils. Secondary nucleation can also occur in conditions where fibril fragmentation is unlikely. It is thought that the surface of existing fibrils can act as a catalyst for the formation of new fibrils (Buell, Dobson and Knowles 2014; Knowles, Vendruscolo and Dobson 2014). This process leads to a positive feedback loop where the formation of fibrils promotes the formation of further fibrils (Cohen *et al.* 2013).

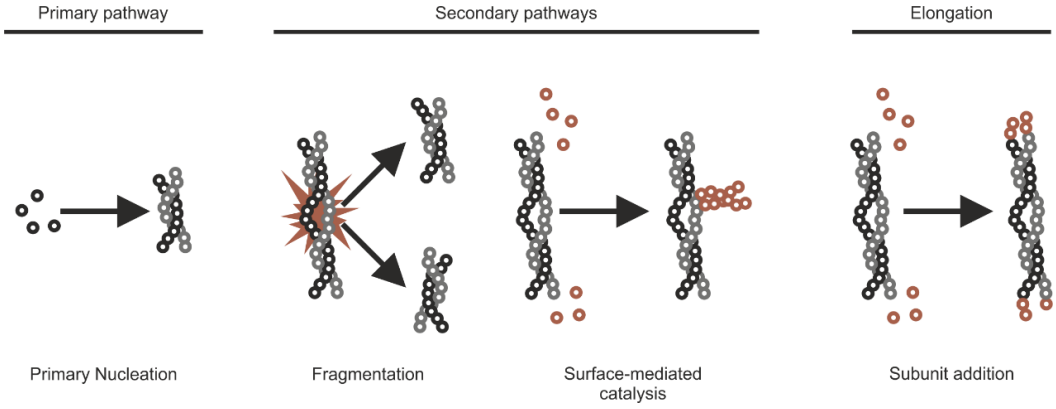


Figure 1.25: Primary and secondary nucleation and elongation in amyloid fibril formation. The primary nucleation of amyloid fibril formation involves the generation of a nucleus from soluble proteins and its subsequent elongation. Secondary nucleation involves the generation of fibrils or nuclei from existing fibrilar structures by a number of possible mechanisms. Elongation is the addition of aggregating units to the end(s) of a fibril. Figure adapted from (Cohen *et al.* 2012).

As described above, the three major processes that determine the kinetics of amyloid fibril formation for a given protein under a certain set of conditions are primary nucleation, secondary nucleation and elongation (Buell, Dobson and Knowles 2014; Cohen *et al.* 2012). The contribution of each of these elements to aggregation can be observed in the aggregation kinetics profile of the protein (Figure 1.26) (Arosio *et al.* 2014; Cohen *et al.* 2012). For example the lag phase of aggregation is primarily determined by the formation of critical nuclei (Eden *et al.* 2015). Therefore, a change in the lag phase length for a protein as a condition changes will indicate that the altered condition has affected primary nucleation

(Arosio *et al.* 2014; Buell, Dobson and Knowles 2014). However, if the gradient of the elongation phase has changed then the effect is likely occurring in the secondary nucleation process (Arosio *et al.* 2014; Buell, Dobson and Knowles 2014). The rate of elongation under a fixed concentration of protein is dependent upon the number of sites for sub-unit addition. Alterations in the elongation rate alters both the apparent lag phase length and the rate of elongation (Arosio *et al.* 2014; Buell, Dobson and Knowles 2014).

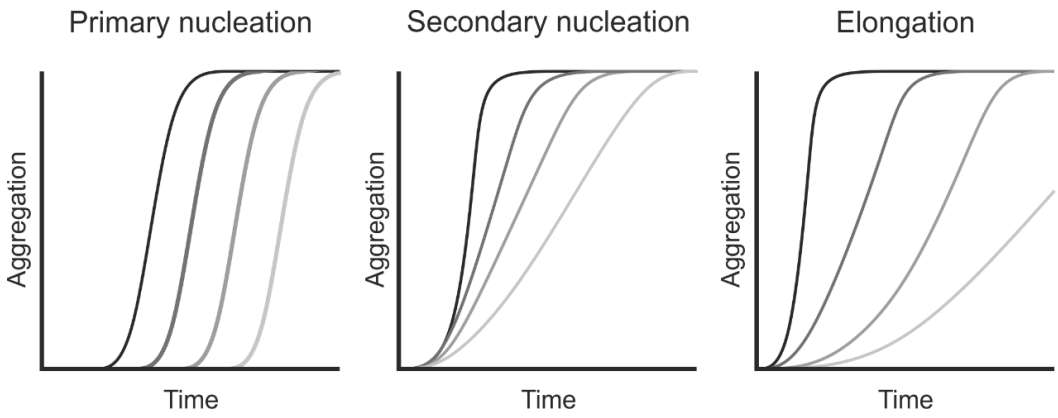


Figure 1.26: Theoretical aggregation kinetics for an amyloid protein undergoing inhibition of different aggregation processes. In each case the grey scale lines represent increasing concentration of inhibitor as the line becomes lighter. Inhibition of primary nucleation results in an increase in the lag time compared to the protein alone. The inhibition of secondary nucleation results in a change in the gradient of the elongation phase. Inhibition of elongation alters both the lag phase and the elongation rate. Figure adapted from (Arosio *et al.* 2014).

A variety of factors can affect the aggregation process by changing one or more of the processes described above. Physical factors such as temperature and agitation, chemical factors such as pH and the presence of other molecules are all known to impact aggregation of amyloid proteins (Buell, Dobson and Knowles 2014; Stewart and Radford 2017). The co-incubation of amyloid proteins with other molecules including metal ions (Abelein, Graslund and Danielsson 2015; Bush *et al.* 1994; Liu *et al.* 2011a; Sarell, Wilkinson and Viles 2010), carbohydrates (Castillo *et al.* 1999; Stewart *et al.* 2016), other proteins (including chaperones)(Wright *et al.* 2015) and small molecules (Amijee *et al.* 2012; Ehrnhoefer *et al.* 2008; Ma *et al.* 2014; Richard *et al.* 2011; Sinha *et al.* 2011; Yang *et al.* 2005) has been shown to affect aggregation significantly (Stewart and

Radford 2017). In some cases these molecules are intended as possible therapeutics (Amijee *et al.* 2012; Ehrnhoefer *et al.* 2008; Ma *et al.* 2014; Richard *et al.* 2011; Sinha *et al.* 2011; Yang *et al.* 2005). However, in other cases they represent molecules that the amyloid proteins are likely to encounter *in vivo* (Abelein, Graslund and Danielsson 2015; Bush *et al.* 1994; Liu *et al.* 2011a; Sarell, Wilkinson and Viles 2010; Castillo *et al.* 1999; Stewart *et al.* 2016); (Wright *et al.* 2015).

1.2.3 Determining amyloid structures

The structure of the fibrils formed at the end of amyloid aggregation is well defined and, on a macroscopic scale, is shared across different amyloid proteins (Knowles, Vendruscolo and Dobson 2014; Makin and Serpell 2005). The typical morphology of an amyloid fibril when examined by transmission electron microscopy or atomic force microscopy is a long unbranched fibril in the micron length range (Chiti and Dobson 2006; Knowles, Vendruscolo and Dobson 2014).

The superstructure of the fibril can consist of multiple protofilaments associated with one another. Evidence for the existence of several variations on this structure can be found even within a single sample (Meinhardt *et al.* 2009; Saiki *et al.* 2005) with single, double and triple or larger fibrils observed, as well as fibrils with alternative lateral associations and “twists” (Meinhardt *et al.* 2009). The protofilaments can be either wrapped round one another or lie side by side to form ribbons (Meinhardt *et al.* 2009; Nugent, Kaminski and Kaminski Schierle 2017; Seuring *et al.* 2017; Sneideris *et al.* 2015; Tycko 2014). There is variation in the structure of the fibril even within these classifications.

Common to all amyloid fibrils is the high abundance of β-sheet content in the fibrils, which is often not present in the soluble protein (Eanes and Glenner 1968; Makin and Serpell 2005). The fibril core consists of stacked β-strands which can be either parallel or antiparallel in arrangement (Madine *et al.* 2008; Nielsen *et al.* 2009). This highly ordered structure is maintained by large numbers of hydrogen bonds between the backbones of the residues in the fibril core (Saiki *et al.* 2005). This cross-β structural motif is a defining characteristic of amyloid

fibrils. When observed by X-ray fibre diffraction, reflections are observed at 4.7 Å and 10 Å (Makin and Serpell 2005) (Figure 1.27). These signals, indicative of the regular spacing of the intra-strand β-strand packing and the defined inter-sheet distances between adjacent β-sheets in the fibril are diagnostic of amyloid structure (Makin and Serpell 2005).

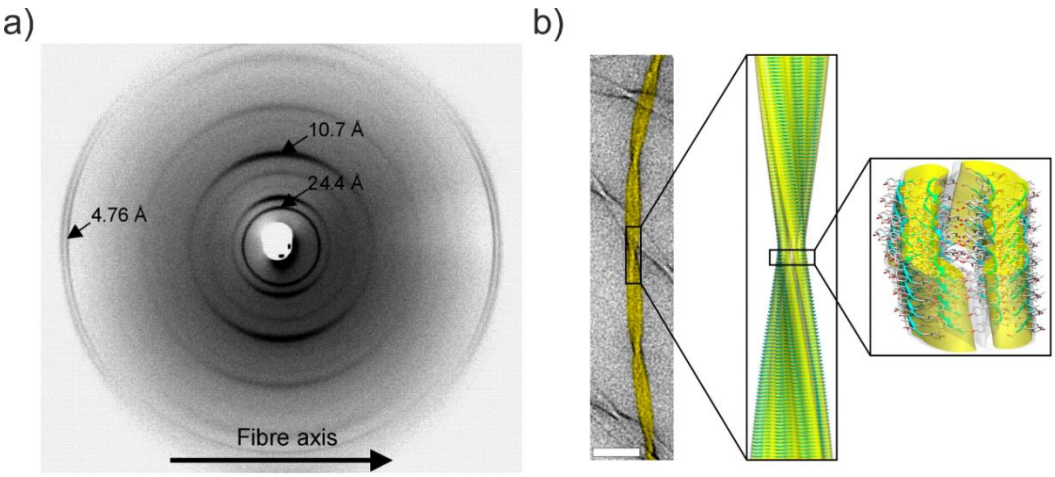


Figure 1.27: Typical characteristics of an amyloid fibril. Amyloid fibrils have a characteristic X-ray diffraction pattern (a) containing a reflection at 4.7 Å corresponding to the inter β-strand spacing in the fibril. The three dimensional structure of a fibril (b) consists of a well ordered core or repeating units. The structure shown here is for a fibril produced by transthyretin residues 105-115. Panel (a) adapted from (Makin *et al.* 2005) and panel (b) adapted from (Fitzpatrick *et al.* 2013).

The common structure of amyloid fibrils also leads to amyloid fibrils binding in a similar manner to certain dye molecules. The interaction of amyloid fibrils with Congo red and the resultant change in birefringence has already been discussed (Bennhold 1922; Divry 1927). Amyloid fibrils can also bind to aromatic dyes such as thioflavin T (ThT) (Vassar and Culling 1959) and Pittsburgh compound B (Hong *et al.* 2014) and lead to enhancement of their intrinsic fluorescence properties - a phenomenon that can be exploited to measure aggregation. It should be noted that the enhancement of ThT fluorescence can be observed in amyloid-protein aggregation prior to the observation of fibrils, indicating that some of the properties of mature amyloid fibrils occur prior to the deposition of mature

fibrillar structures (Ellisdon, Thomas and Bottomley 2006). This indicates that ThT also binds to protofibrillar structures.

Advances in structural biology techniques have recently allowed the observation of the all-atom structure of amyloid fibrils (Colvin *et al.* 2016; Fitzpatrick *et al.* 2013; Fitzpatrick *et al.* 2017; Lu *et al.* 2013b; Paravastu *et al.* 2008; Petkova *et al.* 2005; Petkova, Yau and Tycko 2006; Walti *et al.* 2016; Xiao *et al.* 2015). However, obtaining high resolution structures of these species is not trivial due to a combination of the insolubility of amyloid fibrils and the inherent heterogeneity in many samples. Cryo-EM and solid state NMR have been at the forefront of determining 3D structures of amyloid fibrils. (Colvin *et al.* 2016; Qiang *et al.* 2017; Vugmeyster *et al.* 2016; Walti *et al.* 2016; Xiao *et al.* 2015; Fitzpatrick *et al.* 2013; Fitzpatrick *et al.* 2017) Atomic level structures of transthyretin (TTR) and Tau fibrils have both been solved recently by cyro-EM by Fitzpatrick and co-workers (Fitzpatrick *et al.* 2013; Fitzpatrick *et al.* 2017). The Tau structure was derived from *ex vivo* material from a patient with Alzheimer's disease. Interestingly, the fibril derived from a single patient showed two distinct morphologies indicating that heterogeneity observed *in vitro* is perhaps reflective of the case *in vivo*. Both the TTR and Tau fibril structures revealed the details of the packing interactions in the fibril cores. In the case of Tau, the two morphologies of fibril observed were found to share the same protofibrilillar structural arrangements, with the root of the morphological changes being the ultra-structural arrangement of the protofibrils in relation to one another (Figure 1.28). The variation in fibril structure between species has been linked to differences in the protein sequences and the packing of the side chains.

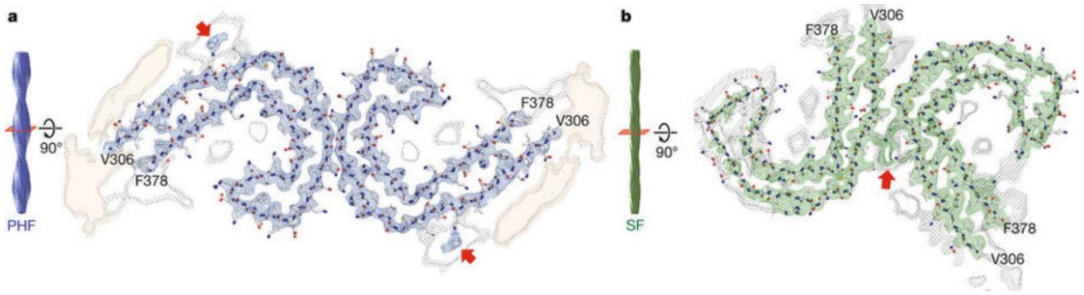


Figure 1.28: Cryo-EM density maps and atomic models of two morphologies of Tau fibrils. (a) A paired helical filament (PHF) structure. (b) A Straight filament (SF) structure. The protofilaments share identical core structure with the variation in structure appearing at the inter-filament packing level. Red arrows indicate additional densities, thought to come from residues 259-274 of Tau, indicating that these residues may play an important role in determining the structure of the SF fibril. Helical reconstructions are shown alongside the atomic structures. Figure take from (Fitzpatrick *et al.* 2017).

There is evidence that the structural variation in fibrils formed from a single protein species may be determined by kinetic factors as opposed to thermodynamic ones (Loveday *et al.* 2012; Morel *et al.* 2010; Pellarin *et al.* 2010). The bottom of the energy well for amyloid fibril may in fact contain many structures with similar stabilities. It has also been suggested that interactions with small molecules, ions or other cellular species (RNA, heparin or proteins) may cause structural heterogeneity (Cremers *et al.* 2016; Jain and Udgaonkar 2010; Okada *et al.* 2008; Selivanova *et al.* 2016; Valle-Delgado *et al.* 2010), possibly by altering the kinetics of aggregation. Indeed, structures have been solved for α-synuclein and Aβ40 in complex with metal ions and glycosaminoglycans respectively (Dearborn *et al.* 2016; Stewart *et al.* 2016).

The occurrence of different morphologies of fibrils in different individuals is one possible cause for the variation in disease progression (Qiang *et al.* 2017). In particular, it has been observed that patients with Alzheimer's disease there are several different fibril morphologies which may correlate with different disease states (Qiang *et al.* 2017).

The structural characterisation of the species which exist in the lag phase and elongation phase of amyloid fibril formation is perhaps even more challenging

that the analysis of the fibrils themselves. The vast array of possible species and the transient nature of the oligomers formed represent a significant challenge for traditional structural biology techniques. It should also be remembered that there is often little indication as to whether a species observed at a given time is on- or off- the pathway to amyloid fibril formation. Research relating to these intermediate species is particularly important in light of the possibility that one or more of them represents the toxic agent in amyloid diseases (Chiti and Dobson 2017; Hoffmann, von Helden and Pagel 2017; Kayed *et al.* 2007; Kayed *et al.* 2003; Kayed and Lasagna-Reeves 2013; Selkoe and Hardy 2016; Sengupta, Nilson and Kayed 2016; Seo *et al.* 2017).

Many amyloid-forming proteins form what are termed “A11-positive oligomers” (Kayed *et al.* 2007; Kayed *et al.* 2003). These oligomeric species were identified by their shared ability to bind the A11 antibody (Kayed *et al.* 2003). Investigation of A11 positive oligomers revealed that they share a spherical morphology and are typically between 10-70 nm in diameter (Kayed *et al.* 2003). A11 positive oligomers have been shown to be more toxic to cells than both the monomer and fibrils of the proteins from which they are formed (Kayed *et al.* 2003).

A second antibody, known as OC, has been shown to bind to fibrilliar oligomers which have a higher β-sheet structure than the A11 positive oligomers (Kayed *et al.* 2007). The OC antibody also recognises mature fibrils suggesting that at least some structural elements are maintained between OC positive oligomers and fibrils (Kayed *et al.* 2007). It has been suggested that the A11 and OC positive oligomeric species may in fact represent elements of two distinct pathways of aggregation (Kayed *et al.* 2007). Indeed there may be many different pathways of aggregation some or all of which may lead to amyloid fibril formation.

A combined MS and gas phase infra-red spectroscopy approach has been developed recently and applied to short synthetic amyloid proteins derived from the insulin B-chain (Hoffmann, von Helden and Pagel 2017; Seo *et al.* 2017). This novel technique has revealed a number of interesting features of amyloid aggregation. The β-sheet transition, a key element of the transition to early

fibriliar oligomers, was observed in oligomers with between 4 and 9 peptide subunits. In addition, the use of IMS allowed the observation of the CCS of the peptides involved. The more expanded conformers were observed to have a higher β-sheet content than the more compact conformations of the same oligomer. These results provide evidence that the transition towards a fibril-like architecture may occur at a very early stage in aggregation, at least in short amyloid-forming peptides.

1.2.4 Amyloid in disease

Amyloid aggregation causes a large number of diseases in humans (Sipe *et al.* 2016). Many of these diseases are neurodegenerative in nature, although there are also systemic diseases and examples affecting other organs (see Table 1.2 for examples). Some of the most well-known amyloid conditions are Alzheimer's disease, Parkinson's disease and type II diabetes.

Some diseases, such as the polyQ aggregation disorders (Fan *et al.* 2014), have a clear and inheritable cause of disease, while many others are described as sporadic (Chiti and Dobson 2006). Alzheimer's disease, normally a sporadic disease, also has an early onset heritable form (Giri *et al.* 2017; Shea *et al.* 2016). Transmissible amyloid diseases are also known, most famously prion diseases (Prusiner 1991), though several recent publications have shown that other amyloid fibrils are capable of acting in an infectious manner (Frontzek *et al.* 2016; Jaunmuktane *et al.* 2015). In the case of dialysis-related amyloidosis, the disease occurs only in patients undergoing dialysis (Miyata *et al.* 1993). However, in a related disease, a mutant form of the same protein, β2m, is observed to aggregate (Valleix *et al.* 2012).

Table 1.2: Examples of amyloid diseases. Figure adapted from (Chiti and Dobson 2017) with additional information from (Fan *et al.* 2014).

Disease	Aggregating protein(s) (precursor)	Organ/tissue(s) affected	Location of Fibrils	Clinical features
Alzheimer's disease	Amyloid-β peptide (Amyloid precursor protein), Tau	Brain	Extracellular (Aβ) Intracellular (Tau)	Dementia
Parkinson's disease	α-synuclein	Brain	Intracellular	Movement disorder, Dementia
Dialysis related amyloidosis	β-2-microglobulin	Systemic (Joints)	Extracellular	Renal failure, Joint tissue destruction
Transthyretin amyloidosis	Transthyretin	Systemic	Extracellular	Peripheral neuropathy
Transmissible spongiform encephalopathies	Prion protein	Brain	Extracellular	Dementia, Ataxia, Psychiatric disturbance
Type II diabetes	Islet amyloid polypeptide	Pancreas (islets of Langerhans)	Extracellular	Insulin resistance, Insulin deficiency,
Huntington's disease	Huntingtin	Brain	Intracellular	Loss of motor control, Dementia
Spinocerebellar ataxia type-3	Ataxin-3	Brain	Intracellular	Ataxia, Cognitive decline
Spinal and bulbar muscular atrophy	Androgen receptor	Brain and spinal cord	Intracellular	Trouble swallowing, Movement disorder, Endocrine dysfunction

The mechanism of the toxicity of amyloid is not well understood (Kayed and Lasagna-Reeves 2013; Selkoe and Hardy 2016; Verma, Vats and Taneja 2015). Both gain of function and loss of function mechanisms have been suggested (Winklhofer, Tatzelt and Haass 2008) (Figure 1.29). It is worth noting that there are known functional amyloids, so the toxicity of amyloid may not be an inherent feature to all amyloid forming proteins (Otzen 2010). However, proteins which do not naturally form amyloid but can be induced to do so (by loss of co-factors (Pertinhez *et al.* 2001) or exposure to acid pH (Guijarro *et al.* 1998) for example) have been shown in some cases to be toxic in their aggregated state (Goldschmidt *et al.* 2010; Auer *et al.* 2008; Chiti *et al.* 1999; Guijarro *et al.* 1998; Urbanc *et al.* 2004; Bucciantini *et al.*). This discovery, along with the toxicity of a wide range of functionally unrelated amyloids, suggests that a toxic gain of function is likely the cause of most amyloid toxicity. However, there may also be contributions from the disruption of the protein original function (Cattaneo *et al.* 2001; Neves-Carvalho *et al.* 2015; Reaume *et al.* 1996).

There is evidence for a number of toxic functions of amyloid fibrils and pre-fibriliar species. Amyloid fibril plaques have long been known to contain species other than the amyloid protein itself (Olzscha *et al.* 2011; Stewart and Radford 2017). Hartl and co-workers have demonstrated that some of the most abundant species in an amyloid fibril deposit are cellular chaperones and unfolded protein response factors (Olzscha *et al.* 2011). The loss of cellular machinery in this manner could lead to significant disruption of cellular function. The reduced ability to respond to protein aggregation (amyloid or otherwise) due to loss of chaperones would be particularly devastating in aging cell where the proteostatic machinery is less efficient (Taylor and Dillin 2011).

The aberrant interactions described in the example above could also lead to a knock-on loss of function in other proteins as opposed to a loss of the function of the aggregating protein. This could occur either through co-aggregation of the interacting species or competition for interaction with their natural ligands. The downstream toxic mechanisms of these interactions is limited only by the

function of the proteins in the cell with which the amyloid species can interact and therefore is potentially extremely complex.

Several amyloid proteins, such as Aβ, form extracellular fibrils. Research has shown that these fibrils, as well as intracellular amyloids, can interact with cellular membranes (Braun *et al.* 2014; Delgado *et al.* 2016; Zhang *et al.* 2017). In addition, monomers (Engel 2009; Engel *et al.* 2008) and soluble oligomeric species (Demuro, Smith and Parker 2011; Lin, Bhatia and Lal 2001; Quist *et al.* 2005) have been also been shown to interact with membranes. Several amyloid protein fibrils have been shown to interact with the cell surface receptor RAGE which causes the activation of cellular stress responses and the recruitment of immune response factors (Yan *et al.* 1996; Yan *et al.* 2000). In addition to activation of existing membrane bound features, it has been suggested that amyloid fibrils and oligomers can disrupt the membrane either directly or via the formation of membrane pores and lead eventually to cell death (Demuro, Smith and Parker 2011; Engel 2009; Lin, Bhatia and Lal 2001; Quist *et al.* 2005; Zhao, Tuominen and Kinnunen 2004).

While toxic gain of function certainly appears to be a strong candidate for the mechanism of amyloid toxicity, there are examples where a loss of function also appears relevant. For example, the protein huntingtin, which aggregates upon expansion of its polyQ domain in the Huntington's disease, is believed to have an anti-apoptotic role in the cell (Luo and Rubinsztein 2009; Nuzzo *et al.* 2016). The sequestering of huntingtin in amyloid fibrils and the attended conformational changes in the protein may prevent this function and thus lead to cell death in cells which normally would not undergo apoptosis.

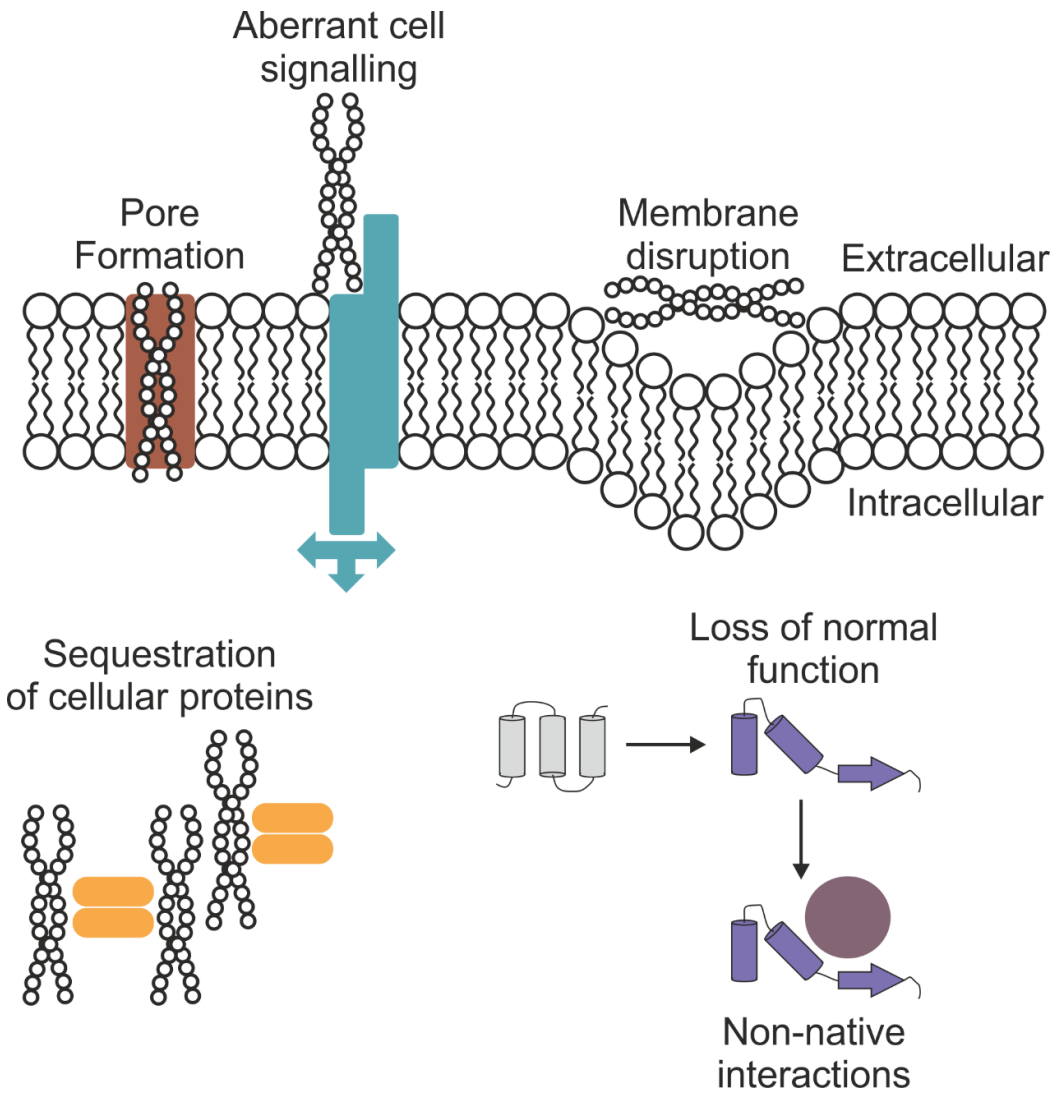


Figure 1.29: Examples of possible amyloid toxic mechanisms. Amyloid proteins may interact with the membrane resulting in pore formation of direct disruption of the membrane. They may also cause aberrant cellular signalling by interacting with cell surface receptors (blue). In the cell aggregated proteins may sequester or co-aggregate with key cellular components such as chaperones (orange). The amyloid protein may lose its original function or develop aberrant non-native interactions with other cellular components (plum). Figure adapted from (Mahood 2015).

Current therapeutic strategies are generally focused towards reducing the effect of the downstream effects of amyloid toxicity. An exception to this is Tafamidis, a small molecule which stabilises the native TTR complex and thus prevents the release of the unfolded monomer required for amyloid formation (Hammarstrom *et al.* 2003). As many amyloid precursors are intrinsically disordered, targeting a native state of the protein is extremely difficult. Development of future therapeutics is focused towards reducing the toxicity of the amyloid species and

reducing the amyloid burden (Folch *et al.* 2016; Sacchettini and Kelly 2002; Yiannopoulou and Papageorgiou 2013).

1.2.5 Examples of amyloid diseases – Alzheimer's disease and Amyloid-β

Alzheimer's disease is a progressive neurodegenerative disease characterised by dementia, cognitive decline and eventually death (Murphy and LeVine 2010). Alzheimer's disease is the most common amyloid disease and the most common dementia and is predicted to become more prevalent as the global population becomes older. The high occurrence, socio-economic costs and the position of Alzheimer's as a quintessential amyloid disease makes AD research a highly active field (Chiti and Dobson 2017; Selkoe and Hardy 2016).

Alzheimer's disease involves the formation of two distinct species of amyloid fibrils. The first consist of the peptide amyloid-β derived from the transmembrane amyloid precursor protein (APP). Cleavage of APP results in the release of soluble peptides 38-43 residues in length, including Aβ40 and Aβ42 (Figure 1.30) (O'Brien and Wong 2011; Ida *et al.* 1996; Qi-Takahara *et al.* 2005; Roher *et al.* 1993a; Roher *et al.* 1993b; Welander *et al.* 2009). These species can then go on to aggregate and form amyloid fibrils which are deposited extracellularly (Glenner and Wong 1984). The second species of fibril is formed from Tau, a microtubule associated protein involved in transport of organelles and large biomolecules as well as stabilisation of apoptotic factors (Lee *et al.* 1991; Wang and Mandelkow 2016). Tau aggregates to form intracellular neurofibrilar tangles (Lee *et al.* 1991). The phosphorylation of Tau is believed to play a key role in its aggregation propensity (Lee *et al.* 1991; Wang and Mandelkow 2016). Both classes of fibril are observed in the brains of Alzheimer's disease patients. The fibrils and the consequential loss of brain tissue are localised primarily to the cerebral cortex (Murphy and LeVine 2010). As with other amyloid diseases, the actually identity of the toxic species in AD is debated and may be produced by Aβ, Tau or both (Bloom 2014; Brier *et al.* 2016; Mondragon-Rodriguez *et al.* 2012).

There are two categories of Alzheimer's disease, sporadic or late onset Alzheimer's disease (LOAD), and early onset Alzheimer's disease (EOAD). The majority of cases (97%) are LOAD (Masters *et al.* 2015). LOAD has no single clear causative feature though many correlatery relationships have been observed which include: age, head trauma (Van Den Heuvel, Thornton and Vink 2007) and genetic features (Carrasquillo *et al.* 2009; Corder *et al.* 1993; Harold *et al.* 2009). EOAD, while less common, has provided many insights into AD and has led to the identification of several heritable genetic features including mutations in PSEN1 (Larner and Doran 2006), PSEN2 (Marcon *et al.* 2004) and APP (Van Broeckhoven 1995). These discoveries have highlighted the importance of APP processing, for example, in the development of AD.

Aβ is generated by the cleavage of APP at the membrane by β- and γ-secretases (Figure 1.30) (O'Brien and Wong 2011). APP is a large transmembrane protein (major isoforms in the brain range from 695 – 770 amino acids residues in length (Matsui *et al.* 2007)) with a complex processing pathway and poorly understood function (Priller *et al.* 2006). Two main pathways are involved in APP processing (Figure 1.30). Cleavage by α-secretase followed by γ-secretase results in a non-amyloidogenic products (Figure 1.30a). Cleavage by β-secretase followed by γ-secretase results in a variety of Aβ peptides from 38-43 residues in length or longer of which Aβ40 and Aβ42 are the two major products (Figure 1.30b-c) (O'Brien and Wong 2011).

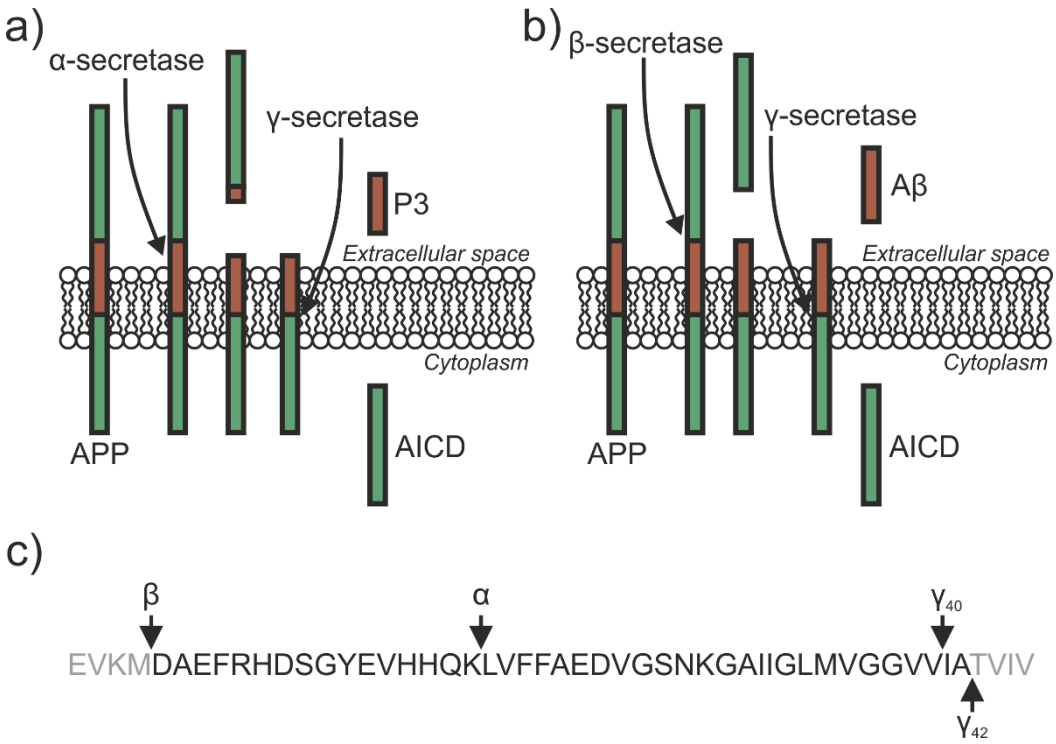


Figure 1.30: Aβ processing at the cell membrane. Aβ is generated by the sequential cleavage of APP by α/β and γ-secretases. This process can proceed via a non-amyloidogenic route (a) involving α-secretase or via an amyloidogenic route (b) involving β-secretase. (c) The site of cleavage can result in a variety of product proteins some of which are shown here. Black shows the sequence of Aβ42 while grey shows the surrounding residues in the context of APP. Figure adapted from (O'Brien and Wong 2011).

Aβ40 and Aβ42 are amphipathic peptides containing a charged N-terminal region followed by a hydrophobic central and C-terminal domain (Figure 1.30c). The central region, residues 14 to 23 (Tjernberg *et al.* 1999), 16-21 (Gorevic *et al.* 1987; Preston *et al.* 2012) or 16 to 22 (Tao *et al.* 2011), is thought to be important for amyloid formation as it has been shown to form amyloid in isolation. The additional two residues on the C-terminus of Aβ42, Ile and Ala, lead increased hydrophobicity of the peptide. This increase in hydrophobicity may drive the increased aggregation of Aβ42 compared to Aβ40.

Aβ fibrils have demonstrable toxicity in cells via a number of mechanisms (Cameron and Landreth 2010; Petkova *et al.* 2005; Qiang *et al.* 2012; Sultana, Perluigi and Butterfield 2009; Tomic *et al.* 2009; Walsh *et al.* 1999). For example fibrils can induce oxidative damage (Butterfield, Swomley and Sultana 2013; Cai,

Zhao and Ratka 2011) and initiate inflammatory responses (Currais *et al.* 2016). Several groups have attempted to determine a structure of Aβ fibrils making use of the techniques discussed previously (Colvin *et al.* 2016; Lu *et al.* 2013b). Aβ fibrils show polymorphism, as observed for fibrils from other proteins (Tycko 2014), although this may be a natural feature of amyloid fibril formation or a consequence of the conditions in which the fibrils formed (these of course are not mutually exclusive conclusions). Certainly there is polymorphism in the fibril morphology between individual disease sufferers, however recent work indicates that a single polymorph dominates Aβ40 and Aβ42 fibrils throughout the brain of a given individual (Qiang *et al.* 2017). Given that the individuals from which the fibrils were derived had differing clinical histories it is possible that the fibril morphology may correlate with a given disease progression.

Probably the most detailed Aβ40 fibril structures are the 2A and 3Q structures published by Tyko and co-workers (Paravastu *et al.* 2008; Petkova *et al.* 2005; Petkova, Yau and Tycko 2006). These two fibril morphologies were produced with agitation and quiescently, respectively, and show strikingly different morphologies (Figure 1.31). While both conformations consist of protofibrils with a strand-turn-strand architecture, 2A fibrils display 2-fold symmetry (Figure 1.28a), with the primary inter-protein contacts mediated by β-strands. The 3Q fibrils, meanwhile, have 3-fold symmetry with the ends of the β-strands and the loops forming inter-protein contacts (Figure 1.31b). It should also be noted that the 2A and 3Q fibrils were produced *in vitro* and both vary from the *ex vivo* structures.

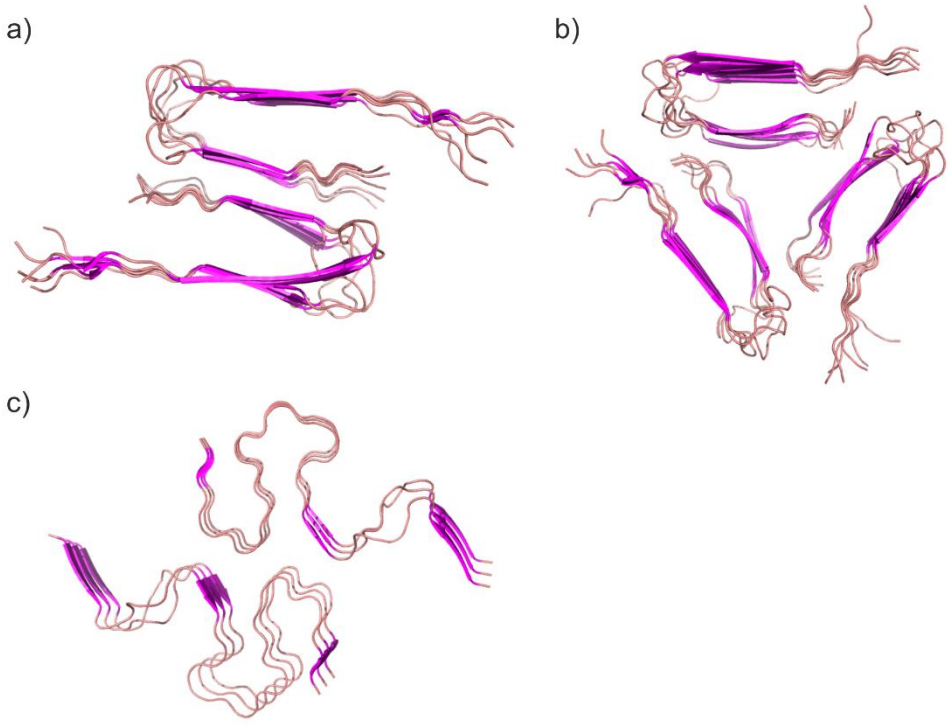


Figure 1.31: Aβ fibril structures. (a) The 2A fibril structure of Aβ40 showing 2-fold symmetry. (b) The 3Q fibril structure of Aβ40 showing 3-fold symmetry. (c) The triple parallel β-strand morphology of the Aβ42 fibril. In all cases the fibril axis is into the page. (a) and (b) adapted from (Paravastu *et al.* 2008) and (c) adapted from (Walti *et al.* 2016).

The Aβ42 fibril structures have historically been less well defined than that of Aβ40 fibrils. Recent ssNMR structures of Aβ40 and Aβ42 fibrils have revealed an unprecedented level of detail, allowing differences between the fibrils to be distinguished for the first time (Qiang *et al.* 2017; Walti *et al.* 2016). Interestingly, recent work has indicated that the Aβ42 protofibril forms three β-strands (Figure 1.31c) rather than the two seen in nearly all Aβ40 fibril models (Walti *et al.* 2016). The triple strand structure is stabilised by interactions from Ala42 with other residues and so is unlikely to be stable for Aβ40. In addition the triple strand Aβ42 fibrils cannot seed Aβ40 fibrils.

Oligomers, either generated from primary nucleation or via secondary nucleation or shedding from fibrils are harder to study by structural techniques than the mature fibrils. The population of prefibrillair species is often transient and heterogeneous rendering investigation by many techniques difficult. Mass

spectrometry has been successfully employed to study oligomers formed by amyloid proteins, for example Aβ40 and Aβ42 (Bernstein *et al.* 2009a). IMS-MS (Figure 1.32) has demonstrated that the two Aβ peptides forms different distributions of oligomers, with Aβ42 forming a tetramer with an "open" conformation which is thought to rapidly form a nucleus for further expansion (Bernstein *et al.* 2009a). The absence of this open-tetramer and the nucleus in Aβ40 have been suggested as possible rationale for the reduced aggregation rate and toxicity of Aβ40 compared to Aβ42 (Bernstein *et al.* 2009a).

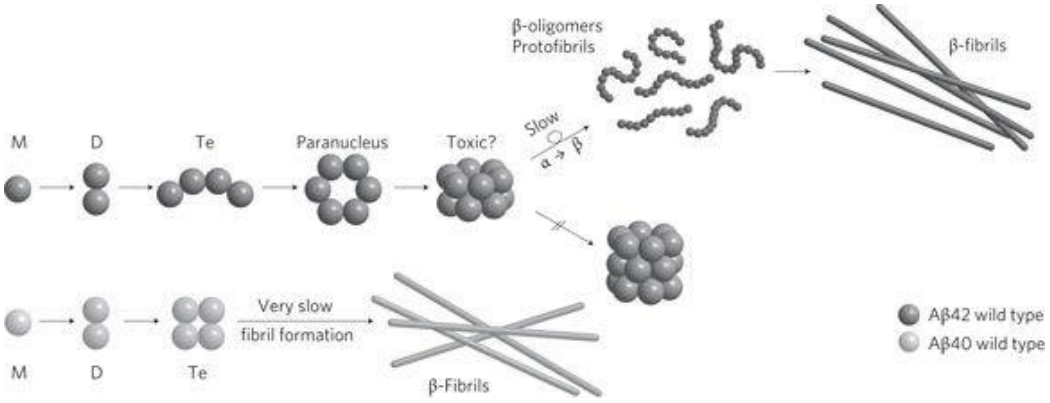


Figure 1.32: An example of possible oligomer types observed in the aggregation of Aβ by IMS-MS. Molecular modelling, based on experimentally determined CCS values, suggest different pathways of aggregation for Aβ40 (light grey) and Aβ42 (dark grey). The progression to fibrils and oligomers was not observed in the experiment. Aβ42 forms an open tetramer (Te) which allows the formation of a hexameric paranueclus which then aggregates rapidly. Aβ40 forms a closed tetramer (Te) which then undergoes slow formation of fibrils. Figure from (Bernstein *et al.* 2009a).

The characterisation of Aβ oligomers has revealed a diverse range of species ranging from dimers (Bernstein *et al.* 2009a; Hung *et al.* 2008) to oligomers of up to 56 kDa (Gong *et al.* 2003; Lesné *et al.* 2006). Mixed oligomers of Aβ40 and Aβ42 have also been studied (Pauwels *et al.* 2012). Many of these species have demonstrated toxicity (measured via inhibition of long term potentiation of neuronal cells) in hippocampal brain slices (Gong *et al.* 2003; Lesné *et al.* 2006) though given the inherent transience of the species, it is difficult to determine which form of the protein the toxicity should be attributed to.

1.2.6 Examples of amyloid diseases – Polyglutamine disease and ataxin-3

Polyglutamine (polyQ) expansion diseases comprise a group of ten inherited protein aggregation disorders characterised by the deposition of proteins containing aberrantly expanded polyQ domains into amyloid fibrils (Fan *et al.* 2014). Unlike Alzheimer's disease, there is a clear and observable genetic element to the polyQ diseases. In each disease, a different protein encoding gene is mutated by the extension of a CAG repeat region (Fan *et al.* 2014).

Expansion of the polyQ domains can occur rapidly for reasons that are not competely understood (McMurray 2010). In a process known as genetic slippage additional regions of CAG repeat can be extended during replication of cellular DNA resulting in expression of a protein with a longer polyQ sequence (McMurray 2010). PolyQ diseases are inherited in an autosomal dominant manner with the exception of Spinal and bulbar muscular atrophy (La Spada 1999), and due to progressive genetic slippage have a tendency to become more severe with each generation (Fan *et al.* 2014).

A different protein is involved in each disease within this class and the length of the polyQ domain required for fibril formation *in vitro* and disease onset varies (Table 1.3), indicating that the sequence and structural context of the polyQ domain play an important role in determining aggregation propensity. Consistent with this, there is no sequence homology between the different proteins involved in polyQ aggregation disorders outside of their polyQ domains, and numerous biochemical and biophysical studies *in vitro* have shown the importance of the sequences flanking the polyQ tract in determining its aggregation propensity (Bhattacharyya *et al.* 2006; Darnell *et al.* 2007b; Darnell *et al.* 2009; Eftekharzadeh *et al.* 2016; Rockabrand *et al.* 2007; Thakur *et al.* 2009). The age of onset of a given disease shows an inverse correlation with the length of the polyQ domain, while disease severity is increased as the expansion lengthens (Fan *et al.* 2014; Wang *et al.* 2016b).

Table 1.3: Lengths of polyQ expansion in the polyQ diseases. Note that the ataxin proteins are unrelated. Figure adapted from (Adegbuyiro *et al.* 2017) and (Fan *et al.* 2014).

Disease	Protein	Normal PolyQ length	PolyQ length in disease
Huntington's disease	Huntingtin	6-35	36-100
Spinobulbar muscular atrophy	Androgen receptor	6-36	38-65
Dentatorubral-pallidoluysian atrophy	Atrophin-1	3-38	49-88
Spinocerebellar ataxia type 1	Ataxin-1	6-39	39-88
Spinocerebellar ataxia type 2	Ataxin-2	13-32	33-77
Spinocerebellar ataxia type 3	Ataxin-3	12-40	55-86
Spinocerebellar ataxia type 6	CACNA1A	4-18	21-33
Spinocerebellar ataxia type 7	Ataxin-7	7-18	38-120
Spinocerebellar ataxia type 12	PPP2R2B	4-32	66-78
Spinocerebellar ataxia type 17	TATA-box binding protein	25-43	47-63

Machado Joseph disease (MJD), also known as Spinocerebellar ataxia type-III (SCA3), is a polyQ disease caused by the expansion of the polyQ domain found in the protein ataxin-3 (Costa Mdo and Paulson 2012; Kawaguchi *et al.* 1994). MJD is characterised by progressive neurological symptoms including ataxia (impaired motor control), tremor, cognitive decline and death (Costa Mdo and Paulson 2012). The age of onset of MJD can range from early childhood in the most severe cases to much later in life for those with shorter polyQ regions (Fan *et al.* 2014). *In vivo* ataxin-3 fibrils are deposited in the nucleus and cytoplasm of neurons (Wang *et al.* 2016b). The deposition of ataxin-3 amyloid fibrils is centred in the hind brain and the pons (Wang *et al.* 2016b). Several mechanisms of toxicity have been suggested for ataxin-3, including impaired protein degradation, mitochondrial dysfunction and transcriptional dysregulation (Evers, Toonen and van Roon-Mom 2014).

Ataxin-3 is a soluble protein found in the cytoplasm (Paulson *et al.* 1997) and nucleus (Tait *et al.* 1998). It is expressed ubiquitously throughout the body in several isoforms (Paulson *et al.* 1997). Ataxin-3 has a deubiquitinase activity with a preference for branched ubiquitin chains with K63 linkages (Winborn *et al.*

2008). This deubiquitinase activity plays a role in protein homeostasis (Wang, Li and Ye 2006). Other suggested functions for ataxin-3 include DNA break repair and apoptosis (Liu *et al.* 2016; Pfeiffer *et al.* 2017).

Ataxin-3 consists of an N-terminal 21 kDa globular Josephin domain (JD) followed by two ubiquitin interacting motifs (UIMs), a polyQ domain and, depending on isoform, a third UIM (Figure 1.33a) (Burnett, Li and Pittman 2003; Masino *et al.* 2003). These domains are separated by disordered linker regions of varying length.

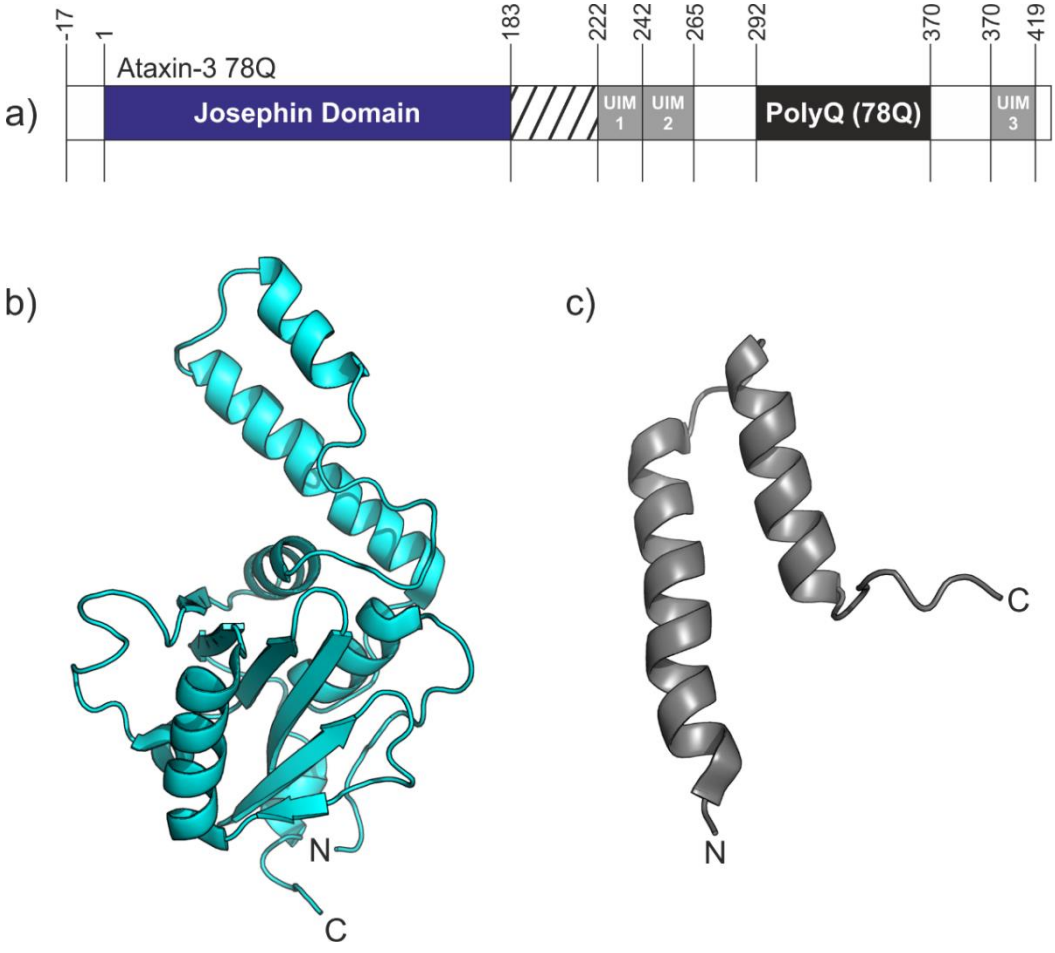


Figure 1.33: Ataxin-3 structure. (a) The domain organisation of ataxin-3 78Q. From N- to C-terminus the protein consists of the global Josephin domain, two UIMs, a polyQ domain and a third UIM each separated by an unstructured region of varying length. The third UIM does not occur in all splice variants. The NMR structures of (b) the JD (Nicastro *et al.* 2005) and (c) the UIM12 region (Song *et al.* 2010) are shown (not to scale).

The JD contains the site of deubquitinase activity, as well as two ubiquitin binding motifs situated on either side of the helical hairpin (top of Figure 1.33b) (Nicastro *et al.* 2009). The structure of JD is well defined with several structures available (Figure 1.33b) (Nicastro *et al.* 2005). UIM1 and UIM2 also have known structures (Figure 1.33c) (Song *et al.* 2010) while UIM3 is thought to share a similar structure. The UIMs are predominantly helical (Song *et al.* 2010). The UIMs are thought to play a role in substrate recognition and ataxin-3's preference for K63 linkages in substrates (Winborn *et al.* 2008). In healthy individuals the length of the polyQ domain varies from 12 to 44 residues (NCBI 2016). Some progress has been made in solving the structure(s) of the polyQ domains for short polyQ variants of ataxin-3 (Zhemkov *et al.* 2016), however, a structure in the context of the full length protein or of an expanded polyQ domain has not been produced to date.

Ataxin-3 is unusual amongst polyQ proteins in that it undergoes aggregation to form amyloid-like structures without the expansion of its polyQ domain (Figure 1.34) (Ellisdon, Thomas and Bottomley 2006). The JD alone can undergo aggregation to form ThT positive amyloid protofibrils *in vitro* (Ellisdon, Thomas and Bottomley 2006). The protofibrils produced in this polyQ independent/JD dependent step are 10-12 nm in diameter and grow in length with time up to several hundred nm and have a curly morphology when viewed by TEM (Ellisdon, Thomas and Bottomley 2006). Expansion of the polyQ domain over 50 residues results in ataxin-3 undergoing a second aggregation step to form longer straight fibrils (50-100 nm in diameter and up to several μm in length) with a morphology more reminiscent of typical amyloid fibres (Ellisdon, Thomas and Bottomley 2006). These fibrils are resistant to degradation by SDS (Ellisdon, Thomas and Bottomley 2006). Despite detailed kinetic analysis of ataxin-3 aggregation (Apicella *et al.* 2012; Ellisdon, Pearce and Bottomley 2007; Ellisdon, Thomas and Bottomley 2006; Lupton *et al.* 2015; Saunders *et al.* 2011) it remains unclear why ataxin-3 undergoes this two-step process. Interestingly, ataxin-3 aggregates formed by a truncation lacking the polyQ domain have been shown to be toxic in yeast (Bonanomi *et al.* 2015), providing evidence that fragments of ataxin-3 may

retain toxicity and that the polyQ domain is not required for some toxic mechanisms.

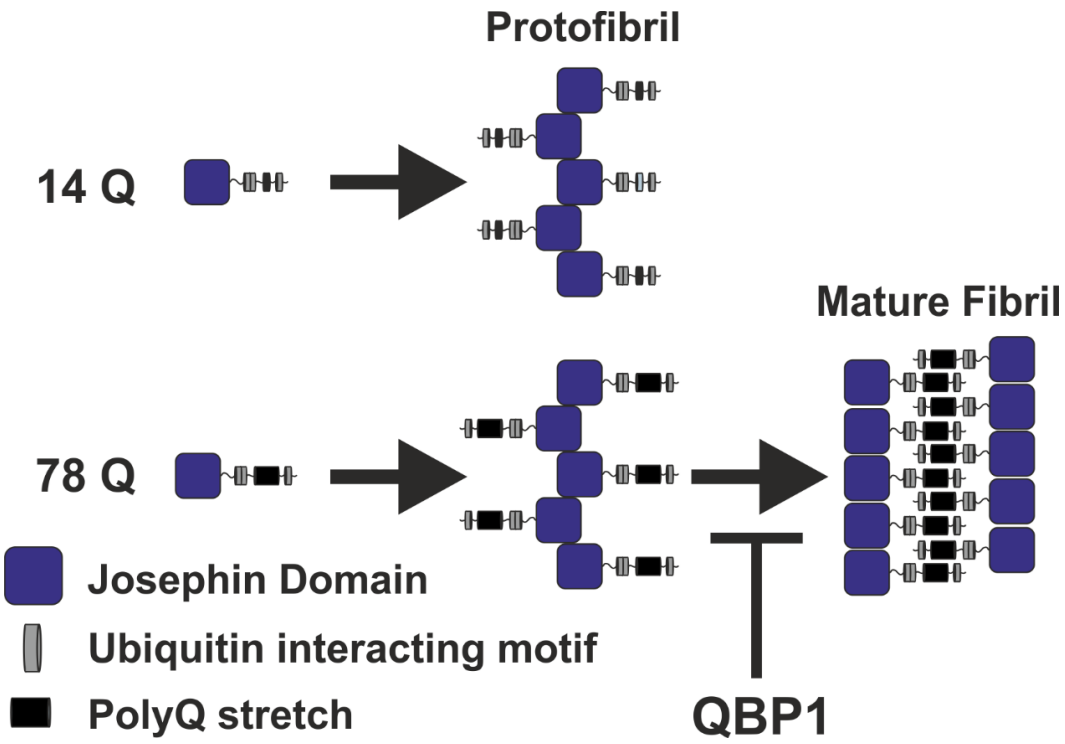


Figure 1.34: Ataxin-3 aggregation occurs via a two-step mechanism. Ataxin-3 containing a non-pathogenic length polyQ domain undergoes a polyQ independent aggregation step to form protofibrils. This step is thought to be mediated by the JD. Ataxin-3 containing an expanded polyQ domain then undergoes a second polyQ length dependent aggregation step to from mature fibrils. The peptide QBP1 has been shown to inhibit this second step and is relevant for the work presented in Chapter 4. Figure adapted from (Ellisdon, Thomas and Bottomley 2006).

Ataxin-3 can be processed by enzymes from the calpain and caspase families (Hubener *et al.* 2013; Jung *et al.* 2009; Liman *et al.* 2014; Simoes *et al.* 2014; Wellington *et al.* 1998). For calpain in particular there is evidence that the cleavage of ataxin-3 (by caspases at residues 241, 244 and 248 (Berke *et al.* 2004) and calpains at residues 60, 221 and 260 (Goti *et al.* 2004; Haacke, Hartl and Breuer 2007; Hubener *et al.* 2013; Simoes *et al.* 2012)) can modulate its aggregation and toxicity in animal models (Hubener *et al.* 2013; Simoes *et al.* 2014). However, ataxin-3 alone also appears to form toxic amyloid fibrils (Bonanomi *et al.* 2015; Evers, Toonen and van Roon-Mom 2014).

The lack of detailed understanding of the processes involved in ataxin-3 aggregation and toxicity has hindered the development of therapeutics for MJD. This is also the case for other polyQ diseases (Fan *et al.* 2014). The characterisation of polyQ proteins aggregation is challenging due to the high frequency of intrinsically disordered regions and the inherent aggregation propensity of the polyQ itself (Fan *et al.* 2014). Therefore, the study of ataxin-3 aggregation is important both in the understanding of MJD and as a model system in the wider context of polyQ aggregation.

1.3 Selected other methods for studying amyloid proteins

1.3.1 Thioflavin T fluorescence assay

As noted in Chapter 1.2, several dyes have been shown to selectively bind to amyloid fibres. Upon interacting with amyloid fibres these dyes show characteristic changes in their physical behaviour; for example, the apple-green birefringence exhibited by Congo-red upon binding to an amyloid fibre (Bennhold 1922; Divry 1927).

The molecule thioflavin T, a benzothiazole salt, exhibits fluorescence at 482 nm when excited at 440 nm. Upon ThT interacting with the β-sheet rich amyloid fibres, this fluorescence increases in intensity (LeVine 1999; Naiki *et al.* 1989). ThT can be used to follow the formation of amyloid fibres in real time by use of a continuous fluorescence detector (hereafter referred to as the ThT assay), allowing for detailed kinetic analysis of the aggregation process. The interpretation of such data is a cornerstone of the study of amyloid disease processes.

The use of ThT as an indicator of the presence of amyloid fibres is not without problems. Unless the conditions are very carefully controlled, the relationship between fibre concentration and ThT fluorescence increase are not necessarily linear (Gade Malmos *et al.* 2017). In addition, not all amyloids cause the same degree of increase in ThT fluorescence (Cloe *et al.* 2011). This raises the potential of under or over estimating the degree of amyloid formation or a false negative result. The species that the ThT fluorescence reports on are not exclusively the final amyloid fibril (Goldsbury *et al.* 2000; Habicht *et al.* 2007; Walsh *et al.* 1999). For example, during the aggregation of the polyglutamine protein ataxin-3, the formation of polyQ independent protofibirils which do not have a classical amyloid morphology results in an increase in the ThT fluorescence (Ellisdon, Thomas and Bottomley 2006). The subsequent maturation process, mediated by the polyQ domain, does not result in further increase in the ThT fluorescence (Ellisdon, Thomas and Bottomley 2006). This renders the interpretation of ThT

fluorescence data more complicated for ataxin-3 aggregation than for other amyloid proteins, as detailed in Chapter 4.

Despite these issues the ability of the ThT assay to provide kinetic information relating to the formation of amyloid fibres and the relative ease (and low cost) of the assay mean that the ThT assay remains a valuable tool for the study of amyloid formation.

1.3.2 Transmission electron microscopy

The macromolecular structures formed by amyloid proteins can vary considerably (see Chapter 1.2.3 for examples). Imaging amyloid fibrils can both give insight into their structure and also serve to ensure that any fibrils formed but not observed by other means (such as the ThT assay) are detected. Comparison of amyloid fibrils produced under different conditions or in the presence of an inhibitor can provide useful information.

In this thesis negative stain transmission electron microscopy (TEM) is used to observe the fibrils and other species formed during amyloid aggregation. In negative stain TEM, a thin carbon grid is prepared onto which a sample is deposited. Molecules in the sample stick to the surface of the grid. After washing away any debris, an electron dense material (known as the stain), such as uranyl acetate is added. The electron dense material adsorbs onto the surface of the biological molecules. The prepared grid can then be dried and stored.

The sample grid is placed in an electron microscope in such a way that the electron beam of the instrument will pass through the sample on its way to a detector (traditionally photographic film or more recently digital detectors). Electrons in the beam which encounter electron dense regions (such as the stain) will be scattered and prevented from reaching the detector, resulting in regions of low signal. Conversely, electrons which encounter relatively electron poor regions (such as unstained regions of the grid) will undergo little scattering and thus reach the detector resulting in high signal. The resulting data can be combined to form a two dimensional image of the material deposited on the grid.

The relatively low resolution TEM performed here should not be confused with the high resolution cryo-electron microscopy described in Chapter 1.2.3, which can be used to produce high resolution structural reconstructions of proteins and other biological molecules.

1.3.3 1H-T1ρ NMR experiments:

1H-T1ρ NMR experiments allow the observation of complex formation in solution while requiring only low concentrations of protein, an advantage when investigating amyloid protein systems in which aggregation may occurs rapidly at higher concentrations.

1H-T1ρ NMR measures the line broadening which occurs in the proton channel of the ligand NMR spectrum when a low molecular weight ligand binds to a high molecular weight partner such as a protein (Lepre, Moore and Peng 2004). This line broadening occurs as a result of the increased tumbling time of a large complex compared to the free ligand. T1ρ experiments are particularly sensitive when measuring weak interactions. Where other NMR techniques rely on peaks shifts, which are not always observed in weak interactions, T1ρ experiments measure only the loss of signal resulting from the line broadening.

In a T1ρ NMR experiment an excess of ligand is exposed to a low concentration of the protein. The protein itself will contribute little to the overall signal in this situation due to its low concentration and slow tumbling time. Any losses in the signal upon ligand binding are indicative of interaction. The magnitude of the signal reduction is dependent upon several factors including the strength of the interaction and the size of the complex formed (Lepre, Moore and Peng 2004). The relationship between these factors is complicated. As such the data presented in Chapter 4 should not be treated as indicative of the relative strengths of interaction between QBP1 and the various proteins, but instead were used qualitatively to assess binding of QBP1 to the different ataxin-3 constructs in solution (Figure 4.17).

1.3.4 Far UV-circular dichroism

Far UV-circular dichroism (CD) analyses differences in the absorbance of left hand and right hand circularly polarised light by chemical groups in amino acids caused by the secondary structure of proteins (Greenfield 2006). Various methods exist for calculating the relative abundance of each secondary structure element in a sample (Andrade *et al.* 1993; Compton and Johnson 1986; Provencher and Glockner 1981; Sreerama and Woody 1993). This can be extremely useful for examining or confirming the secondary structure of an unknown protein or peptide or to confirm that an isolated peptide retains its secondary structure. Comparison of the secondary structure determined by CD with those predicted from NMR of x-ray crystallography can also be undertaken.

Given that amyloid proteins often undergo an increase in β-sheet structure during or upon aggregation (as described in Chapter 1.2), CD provides an excellent opportunity for the examination of amyloid protein aggregation.

1.4 Aims of this thesis

Amyloid diseases represent a growing issue in the developed and developing world. There are few effective therapeutics to treat these diseases. In order to develop novel therapeutics our understanding of the processes involved in aggregation must be improved. The aim of this study is to investigate the aggregation of two amyloid systems in relation to molecules which modulate amyloid formation. The primary tools used for this investigation were mass spectrometry and ThT aggregation kinetics supported by other biophysical and biochemical techniques.

The work in Chapter 3 describes the investigation of the effect of YDL085CA on the aggregation of amyloid proteins. YDL085CA is an ortholog of the human protein SERF1 which has been shown to selectively accelerate the formation of amyloid fibrils (Falsone *et al.* 2012). The effect of the addition of YDL085CA on the kinetics of the aggregation of $A\beta_{40}$ are investigated. In addition, crosslinking mass spectrometry is used to define the binding site of YDL085CA on $A\beta_{40}$ and to examine the effects this has on protein conformation.

In Chapter 4, a known inhibitor of polyQ aggregation, the peptide QBP1, is investigated. The interaction of QBP1 with a range of ataxin-3 constructs is examined by use of native mass spectrometry and compared with NMR data. A novel site of interaction is observed and the potential relevance of this site to the literature is discussed.

Finally, in Chapter 5 the results are summarised and their relevance to the field of amyloid aggregation is discussed.

2 Materials and methods

2.1 Equipment

Mass spectrometers:

Synapt High Definition Mass Spectrometer (HDMS) quadrupole ion-mobility time-of-flight mass spectrometer (Waters Corp., Manchester, UK).

Synapt G2-Si HDMS quadrupole ion-mobility time-of-flight mass spectrometer (Waters Corp., Manchester, UK).

Synapt G2S HDMS quadrupole ion-mobility time-of-flight mass spectrometer (Waters UK, Manchester, UK)

Xevo G2-XS quadrupole time-of-flight mass spectrometer (Waters Corp., Manchester, UK).

Liquid chromatography equipment:

Acquity UPLC M-Class (Waters Corp., Manchester, UK).

MassPREP protein desalting column (Waters Corp., Manchester, UK)

Symmetry C18 trap column (Waters Corp., Manchester, UK)

HSS T3 C18, 75 µm i.d. x 150 mm analytical column (Corp., Manchester)

Nuclear magnetic resonance instruments:

600 MHz NMR magnet (Oxford Instruments, Abingdon, UK)

QCI-P-cryoprobe and an Avance III HD console (Bruker Corpn., Coventry, UK)

Micro plate readers and plates:

FLUOstar Omega (BMG Labtech, Ortenburg, Germany)

FLUOstar Optima (BMG Labtech, Ortenburg, Germany)

CLARIOstar (BMG Labtech, Ortenburg, Germany) with monochromator

Corning® 96 Well Half Area Black with Clear Flat Bottom Polystyrene NBS™ Microplate, non-sterile (Corning GmbH, Wiesbaden, Germany)

Corning Costar® 96-Well Black Polystyrene Plate (Corning GmbH, Wiesbaden, Germany)

Electron microscope:

JEOL JEM-1400 transmission electron microscope (JEOL USA Inc., Peabody, USA) with Gatan US1000XP 2k x 2k CCD camera (Gatan Inc., Pleasanton, USA)

Circular dichroism:

Jasco J715 spectropolarimeter (JASCO, Easton, USA) with a 6 cell changer and peltier temperature control

Spectrophotometers:

Nanodrop 2000 (Thermo Scientific, Surrey, UK)

VersaWave microvolume spectrophotometer (Expedeon Ltd., Cambridgeshire, UK)

Ultrospec 1200 pro (Amersham Biosciences/GE healthcare, Little Chalfont, UK)

UV 1800 UV-VIS Spectrophotometer (Shimadzu, Kyoto, Japan)

Centrifuges:

Avanti J-26 XP Centrifuge (Backman Coulter, Brea, CA, USA)

GenFuge 24D Centrifuge (Progen Scientific, London, UK)

Eppendorf 5810R Centrifuge (Fisher Scientific, Loughborough, UK Beckman Coulter)

Eppendorf 5430R Centrifuge (Fisher Scientific, Loughborough, UK Beckman Coulter)

Eppendorf 5418 Centrifuge (Fisher Scientific, Loughborough, UK Beckman Coulter)

Mini Fuge microcentrifuge (Starlab (UK), Milton Keynes, UK)

Incubators, shakers and related equipment:

Innova 43 Shaker Incubator (New Brunswick Scientific, USA)

Innova 44 Shaker Incubator (New Brunswick Scientific, USA)

SI600 orbital incubator (Stuart, Staffordshire, UK)

Economy incubator size 1, with fan (Gallenkamp)

T18 digital ULTRA TURRAX (IKA, North Carolina, USA)

Constant cell disruption systems cabinet (Constant systems ltd., Nottinghamshire, UK)

Protein purification:

AKTAprime plus (GE healthcare, Little Chalfont, UK)

Peristaltic Pump P-1 (GE healthcare, Little Chalfont, UK)

HisTrap FF 5 mL Ni Sepharose (GE healthcare, Little Chalfont, UK)

Hi Load 26/60 Superdex Peptide S75 (GE healthcare, Little Chalfont, UK)

Hi Prep Superdex Sephacryl HR S200 (GE healthcare, Little Chalfont, UK)

GL Superdex 10/300 Peptide 200 (GE healthcare, Little Chalfont, UK)

Gel electrophoresis:

Slab Gel Electrophoresis Chamber AE-6200 (ATTO, Tokyo, Japan)

Powerpac 3000 (Bio-Rad Lab., Hercules, CA, USA)

Powerpac Basic (Bio-Rad Lab., Hercules, CA, USA)

Other equipment:

Zeba™ Spin Desalting Columns, 7K MWCO, 0.5 mL (Thermo Scientific, Surrey, UK)

Bio-Rad Micro Bio-Spin^{TM} 6 columns (Bio-Rad Lab., Hercules, CA, USA)

Bio-Rad T1000 Thermocycler (Bio-Rad Lab., Hercules, CA, USA)

Grant JB1 Unstirred Waterbath (Grant Instruments, Shepreth, UK)

Jenway 3020 Bench pH Meter (Bibby Scientific, Stone, UK)

SnakeSkin Pleated Dialysis Tubing – 3,500 MWCO (Thermo Scientific, Surrey, UK)

Techne Dri-block Heater (Bibby Scientific, Stone, UK)

Thriller shaker incubator (VWR International Ltd., Leicestershire, UK)

Software:

Masslynx 4.1 (Waters Corp., Manchester, UK).

Driftscope 2.5 (Waters Corp., Manchester, UK).

Peaks 8.0 (Bioinformatics Solutions Inc., Waterloo, Canada)

StavroX 3.6.0.1 (University of Halle-Wittenburg, Halle, Germany)

Bruker TopSpin (Bruker Corpn., Coventry, UK), NMRPipe (Institute for Bioscience and Biotechnology Research, University of Maryland, Maryland, USA) and CcpNMR analysis software (CCPN, UK)

2.2 Chemicals

Table 2.1: Chemicals used in this investigation

Item	Company
(Bacto)Tryptone	Fisher Scientific, Loughborough, UK
Acetic acid, glacial	Fisher Scientific, Loughborough, UK
Acetonitrile	Fisher Scientific, Loughborough, UK
Agar	Fisher Scientific, Loughborough, UK
Agarose	Melford Laboratories, Suffolk, UK
Ammonium Acetate, 7.5 M stock	Sigma Life Sciences, St. Louis, USA
Ammonium Bicarbonate	Sigma Life Sciences, St. Louis, USA
APS	Thermo Scientific, Surrey, UK
BS^3 (bis(sulfosuccinimidyl)suberate)	Thermo Scientific, Surrey, UK
Caesium iodide	Sigma Life Sciences, St. Louis, USA
DTT (1,2-dithioltheritol)	Formedium, Norfolk, UK
Ethanol	Honeywell research chemicals, Romania
Ethidium Bromide, EtBr	Nippon Genetics Europe, Germany
Ethylenediamine tetra acetic acid, EDTA	Thermo Scientific, Surrey, UK
Formic acid	Fisher Scientific, Loughborough, UK
Glucose	Fisher Scientific, Loughborough, UK
Glycerol	Fisher Scientific, Loughborough, UK
Imidazole	Sigma Life Sciences, St. Louis, USA
Potasium phosphate, monobasic (KH_2PO_4)	Thermo Scientific, Surrey, UK
LB broth, granulated	Melford Laboratories, Suffolk, UK
MgCl2	Sigma Life Sciences, St. Louis, USA
$MgSO_4$	Fisher Scientific, Loughborough, UK
NH_4Cl	Honeywell research chemicals, Romania
PMSF	Sigma Life Sciences, St. Louis, USA
Sodium chloride, NaCl	Fisher Scientific, Loughborough, UK
Sodium dodecyl sulphate, SDS	Fisher Scientific, Loughborough, UK
Sodium hydroxide, NaOH	Honeywell research chemicals, Romania
Sodium Iodide (Waters ToF Instruments Service Sample kit, 2 µg/µL NaI in 50/50 2-propanol/water)	(Waters Corp., Manchester, UK)
Sodium phosphate dibasic, Na2HPO4	Thermo Scientific, Surrey, UK
Sodium phosphate monobasic, NaH2PO4	Sigma Life Sciences, St. Louis, USA
sodium sulphate, NaSO4	Fisher Scientific, Loughborough, UK
Tetramethylethylenediamine (TEMED)	Sigma Life Sciences, St. Louis, USA

Tris-(hydroxymethyl)-aminomethane, Tris base	Fisher Scientific, Loughborough, UK
Yeast extract	Fisher Scientific, Loughborough, UK
α-lactose	Sigma Life Sciences, St. Louis, USA
IPTG	Formedium, Norfolk, UK
Sodium azide	Fisher Scientific, Loughborough, UK
HPLC grade water	Fisher Scientific, Loughborough, UK

2.3 Enzymes for molecular biology and protein digest

Table 2.2: Enzymes used in this investigation

Item	Company
BAMHI	New England Biolabs (UK) Ltd., Herts., UK
DNase I	Sigma Life Sciences, St. Louis, USA
ECORI	New England Biolabs (UK) Ltd., Herts., UK
Lysozyme	Sigma Life Sciences, St. Louis, USA

2.4 Antibiotics

Table 2.3: Antibiotics used in this investigation

Item	Company
Carbenicillin	Formedium, Norfolk, UK
Chloramphenicol	Sigma Life Sciences, St. Louis, USA
Kanamycin	Formedium, Norfolk, UK

2.5 Markers and dyes

Table 2.4: Gel standards and dyes used in this investigation

Item	Company
Expedeon Instant Blue Coomassie stain	Expedeon, Harston, UK
Precision Plus Protein™ Dual Xtra Standard	Bio-rad, Hemel Hempstead, UK

2.6 Kits

Table 2.5: Kits used in this investigation

Item	Company
Peirce micro BCA kit	Thermo Scientific, Surrey, UK
Q5 Quikchange mutagenesis kit	New England Biolabs (UK) Ltd., Herts., UK
Wizard ® Plus SV Minipreps DNA purification system	Promega, Southampton, UK

2.7 LB-agar plates

LB-Agar plates containing various antibiotics were produced using the recipe in table 2.6. The mixture, minus antibiotics, was autoclaved and allowed to cool. Once the solution had cooled to around 50 °C the antibiotics were added and mixed. The LB-agar solution was then poured into plastic petri dishes and allowed to cool.

Table 2.6: LB-Agar plate recipe

LB-Agar
1.2 g Agar
2.5 g LB
Variable amounts of antibiotic
100 mL water

2.8 Buffers

The buffers used in this investigation are described in the experimental sections. Where a buffer is particularly complex or unusual it is described below.

Auto induction media

The auto induction media used in the expression of ataxin-3 and ataxin-3 truncations is described in table 2.7. The individual components are described in table 2.8. All buffers were autoclaved before use with antibiotics and α-lactose/glucose 0.22 µm filtered and added after autoclaving.

Table 2.7: Details of the ataxin-3 auto induction media

Component	Volume
2ZY	463 mL
$MgSO_4$	1 mL
50x Lac	10 mL
20x NPSC	25 mL
Carbenicillin and Chlorampehicol	500 µL each

Table 2.8: Components of the ataxin-3 auto induction media

Solution	Components
2ZY	5 g Yeast extract 10 g (Bacto)tryptone 463 mL water
$MgSO_4$	12.3 g $MgSO_4$ 50 mL water
50x Lac	25 g Glycerol 2.5 g Glucose 10 g α-lactose 100 mL water (Heat to dissolve the α-lactose)
20x NPSC	26.75 g NH_4Cl 16.1 g $NaSO_4$ 34 g KH_2PO_4 35.5 g Na_2HPO_4 pH 6.75 500 mL water
Carbenicillin	100 mg mL^{-1} in water
Chloramphenicol	25 mg mL^{-1} in 100% EtOH

Ataxin-3 purification buffers

The ataxin-3 purification buffers are described below in table 2.9. All buffers were filtered (0.22 µm filter) and degassed prior to use and stored at 4 °C.

Table 2.9: Buffers used in the Nickel affinity chromatography

Buffer	Components
200 mM sodium phosphate, pH 7.5	Set the pH of solution 1 to 7.5 by addition of solution 2.
Solution 1	$Na_2HPO_4.2H_2O$ 35.598 g 1 L water
Solution 2	$NaH_2PO_4.2H_2O$ 7.120 g 100 mL water
Buffer A (20 mM sodium phosphate,0.5 M NaCl, 5% glycerol v/v, pH 7.5)	100 mL of sodium phosphate buffer 200 mM 29.22 g NaCl 5 % glycerol (v/v) 1 L water pH 7.5.
Buffer B (20 mM sodium phosphate,0.5 M NaCl, 5% glycerol v/v, 500 mM imidazole, pH 7.5	50 mL of sodium phosphate buffer 200 mM 14.6 g NaCl 17.0 g imidazole 5 % glycerol (v/v) pH 7.5.

2.9 Methods

2.9.1 Molecular biology

In all cases the sequence of the final constructs was confirmed through DNA sequencing.

2.9.1.1 Ataxin-3

The ataxin-3 encoding plasmids for ataxin-3 78Q, ataxin-3 14Q and JD were provided by Prof. Sandra Macedo-Ribeiro (Institute for Molecular and Cell Biology, Porto, Portugal). They consist of an ataxin-3 gene in a pET11a plasmid as described in Appendix I.

The truncation variants of atatxin-3 were generated by inserting stop codons at the appropriate site in ataxin-3 using a Q5 Quikchange mutagenesis kit. The primers used are shown in Table 2.10. The primers were mixed with ataxin-3 14Q plasmid DNA and the manufacturer's protocol was followed.

Table 2.10: Primers used in the generation of novel truncated ataxin-3 constructs

Construct	Primer
JDU1	CGCCAAGAAATTGACTGATAAGATGAGGAAGCAGATCTCCGCAGGG and CCCTGCGGAGATCTGCTTCCTCATCTTATCAGTCAATTTCTTGGCG
JD+	TAATTAGACGAAGATGAGCAG and CATTCCTGAGCCATCATTTG

2.9.1.2 MBP-183-221

Dr David Brockwell (University of Leeds) provided a modified pMAL-c5X plasmid containing the BamA POTRA domains with the addition of an N-terminal 6x His-tag (HT) and replacement of the thrombin cleavage site with a TEV cleavage site. The sequence encoding residues 183-221 of ataxin-3 were excised from pET11a ataxin-3 14Q by PCR with the primers shown in table 2.11. The pMAL-c5X POTRA plasmid was digested with BAMHI and ECORI and the digested plasmid ligated with the ataxin-3 sequence excised above. The resulting plasmid contained a construct that contained (from N- to C-terminus of the expressed protein) a hexa-His tag, MBP, TEV cleavage site and residues 183-221 of ataxin-3.

Table 2.11: Primers used in the generation of MBP-183-221

Construct	Primer
MBP+	ATTACTTGCGGATCCGTCCAACAGATGCATCGACCAAAAC and AGGACTAGAGAATTCCTTACATTCCTGAGCCATCATTTGCTTC

2.9.1.3 Aβ40

The pETSac plasmid encoding WT Aβ40 was provided by Dr. Sara Linse (Lund University, Sweden) and Prof. Dominic Walsh (Harvard Institute of Medicine, USA).

2.9.1.4 α-synuclein

The pET23a plasmid encoding α-synuclein was provided by Prof. Jean Baum (Department of Chemistry and Chemical Biology, Rutgers University, NJ, USA).

2.9.1.5 YDL085CA

YDL085CA protein was provided by Prof. James Bardwell (University of Michigan). Two sequences were used in the expression of YDL085CA (see experimental section for details). In both cases the sequence for YDL085CA was inserted into pET28b with an N-terminal His-SUMO tag. The first contained the cDNA sequence from *S. cerevisiae* while the second was codon optimised for *E. coli*. Both sequences are shown in Table 2.12 with the rare codons encoding arginine highlighted in red.

Table 2.12: The sequence of the YDL085CA coding region of the plasmids. Rare codons encoding arginine highlighted in red.

Sequence name	Sequence
YDL085CA	ATG GCT AGA GGT AAT CAA AGA GAC TTG GCA AGA CAA AAA AAC TTG AAA AAA CAA AAG GAC ATG GCT AAG AAC CAG AAA AAG AGT GGT GAT CCT AAG AAA AGA ATG GAG TCA GAC GCC GAA ATC TTG AGA CAA AAG CAG GCC GCC GCA GAT GCT AGA AGA GAG GCT GAA AAG CTT GAG AAG TTA AAA GCC GAA AAG ACG AGA AGA TAA
YDL085CA codon optimised	ATG GCG CGC GGT AAC CAG CGT GAC CTG GCC CGT CAG AAA AAT CTG AAA AAA CAA AAA GAT ATG GCG AAG AAT CAA AAA AAA TCT GGT GAT CCA AAA AAA CGC ATG GAG TCC GAC GCA GAA ATT CTG CGC CAG AAA CAG GCC GCC GCC GAT GCA CGT CGC GAA GCA GAA AAG CTG GAA AAA CTG AAA GCG GAG AAA ACC CGC CGT TAA

The ULP1 cleavage sequence in codon optimised plasmid was subsequently also modified to delete the c-terminal serine and allow expression and purification of a YDL085CA without the non-natural serine residue at its N-terminus.

2.9.2 Expression and purification of proteins

2.9.2.1 Expression and purification of ataxin-3 constructs

All ataxin-3 constructs (ataxin-3 78Q, ataxin-3 14Q, JDU1, JD+ and JD) were expressed and purified following the same protocol, previously described (Scarff *et al.* 2013) and altered as indicated below.

BL21(DE3)pLysS cells were transformed with pDEST17 expression vector containing the appropriate ataxin-3 sequence. 2 µL of plasmid DNA (100-200 ng μL^{-1}) was added to 50 µL of cells and incubated on ice for 30 minutes. The cells were then heat shocked by incubation in a water bath at 42 °C for 45 seconds followed by 5 minutes of further incubation on ice. The transformation mixture was then plated on LB-Agar plates containing 100 µg mL^{-1} carbenecillin/25 µg mL^{-1} chloramphenicol. The plates were then incubated at 37 °C overnight.

The following day transformed colonies were selected and used to inoculate 50 mL of LB media containing 100 µg mL^{-1} carbenecillin/25 µg mL^{-1} chloramphenicol. The cultures were incubated overnight at 37 °C with shaking.

The overnight suspension cultures were used to inoculate 500 mL of autoinduction media containing 100 µg mL^{-1} carbenecillin/25 µg mL^{-1} chloramphenicol in 2.5 mL conical flasks (without baffles). 2 mL of the overnight culture was added to each 500 mL of autoinduction media. The cultures were incubated at 25 °C with 200 rpm shaking for 20-30 hours (ideally 22 hours).

After the incubation period the OD_{600} of the culture was determined (approximately 10). The culture was centrifuged at 6500 rpm in a JLA 8.1 rotor for 30 mins at 4°C. The supernatant was discarded and the pellet for each litre of culture resuspended in 40 mL of buffer A (20 mM sodium phosphate, 0.5 M NaCl, 5% glycerol v/v, pH 7.5) with 1 mL of 50 µg mL lysozyme added. The resuspended pellets were frozen and stored at -20 °C.

For the purification typically the pellets of two litres of culture were combined and purified together. Higher volumes of culture may have resulted in the overloading of the nickel affinity columns.

The cell suspension was thawed under running warm water. DNase, $MgCl_2$ and PMSF were added to final concentrations of 5 µg mL^{-1}, 10 mM and 1 mM respectively. The solution was incubated at 4 °C with gentle stirring until the solution was homogenous. If the solution remained “stringy” after 20 minutes the concentration of DNase was doubled. Once the solution was homogenous it

was poured into a syringe with a blunt needle and gently syringed to break up any remaining large particles. The resulting solution was then centrifuged at 16000 rpm in a JA 25.50 rotor at 4 °C for 40 minutes. The supernatant was collected and syringe filtered through a 0.22 µm filter.

Two NiNTA columns were arranged in series and equilibrated with 5 column volumes of water followed by 5 column volumes of buffer A. The filtered supernatant was applied to the column using a peristaltic pump. After the supernantent was applied the columns were moved to an AKTA prime plus. A stepped gradient of buffer A and buffer B (buffer A with the addition of 500 mM imidazole) was applied as shown in Table 2.13. The flow rate was 3 mL min with a pressure limit of 0.15 mPa. Each step of the gradient was applied for 5 column volumes which was enough to return the A_{280} signal to baseline post elution.

Table 2.13: Details of the AKTA program used in the NiNTA purification of ataxin-3

Buffer A	Buffer B	Imidazole concentration	Fractions
100%	0%	-	5 mL
95%	5%	25 mM	2.5 mL
90%	10%	50 mM	2.5 mL
85%	15%	75 mM	2.5 mL
80%	20%	100 mM	10 mL
0%	100%	500 mM	10 mL

The fractions collected from the peaks at 50 mM and 75 mM imidazole were combined to form a single solution and buffer exchanged into 250 mM ammonium bicarbonate/1 mM DTT/5% glycerol (v/v), pH 8.0 and concentrated to a volume of 5-10 mL using a vivaspin 20 kDa molecular weight cut off column. The concentrated solution was syringe filtered through a 0.22 µm filter.

A HiPrep 26/60 sephacryl S200 column with a bed volume of 230 mL was attached to an AKTA prime plus system in a cold room at 4 °C. The column was equilibrated with 1 column volume of water followed by 1 column volume of 250 mM ammonium bicarbonate/1 mM DTT/5% glycerol (v/v), pH 8.0. The protein solution was loaded into a loop (either a single 5 mL loop or two 5 mL loops in

series depending on the volume of solution). The AKTA was run programmed as indicated in Table 2.14.

Table 2.14: Details of the AKTA program used to purify ataxin-3

Volume	Flow rate	Fractions	Notes
0 mL	1 mL min^{-1}	-	
5 mL	0.5 mL min^{-1}	-	Switch to loop
25 mL	0.5 mL min^{-1}	-	Switch to Load
50 mL	0.5 mL min^{-1}	2.5 mL	
220 mL	0.5 mL min^{-1}	5.0 mL	
320 mL	0.5 mL min^{-1}	-	
321 mL	0.1 mL min^{-1}	-	

The peak corresponding to the monomeric protein was collected and combined. Only the second half of the peak should be collected in order to avoid contamination with dimer and larger species. The protein was concentrated as described above to 10 mg mL^{-1} and snap frozen in 50 µL or 100 µL aliquots. The aliquoted proteins were stored at -80 °C.

An example purification for ataxin-3 78Q is shown in Figure 2.1. The 50 mM and 75 mM imidazole elutions from the NiNTA purification were combined and then purified by SEC as described above. The gel sample shows that only ataxin-3 78Q is present in the final sample. The SEC elution profiles for the other ataxin-3 constructs (ataxin-3 14Q, JDU1, JD+ and JD) are shown in Appendix I.

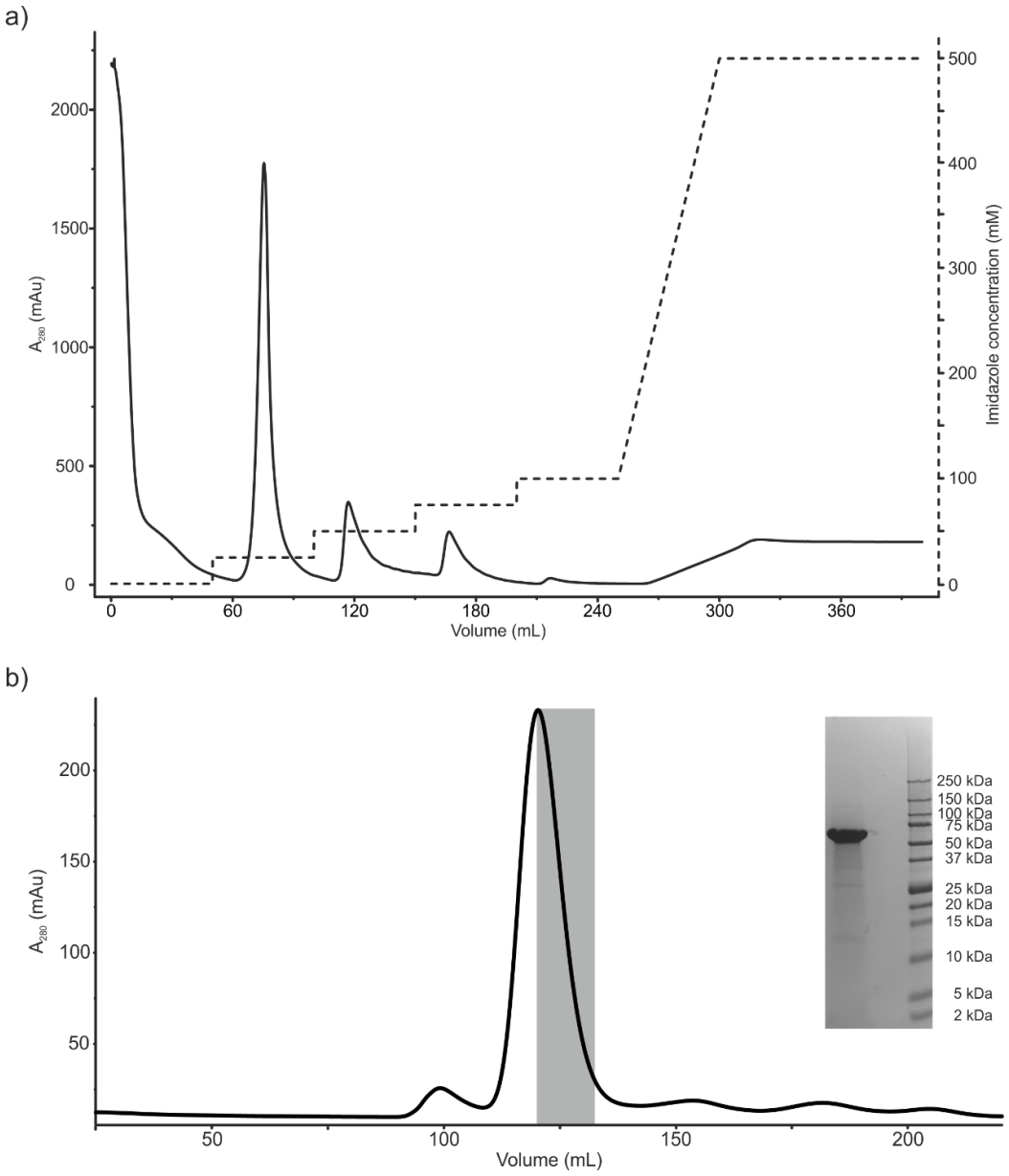


Figure 2.1: The purification of ataxin-3 78Q. (a) A_{280} trace of the NiNTA purification of ataxin-3 78Q. The dashed line shows the concentration of imidazole. Note that on this occasion the 100 mM to 500 mM change was a gradient rather than a step. (b) The A_{280} trace for the SEC of ataxin-3 78Q. The grey box indicates the fractions collected as monomeric protein. Inset: SDS-PAGE of the resulting protein at 20 µM showing a single intense band at approximately the correct molecular weight (53 kDa).

2.9.2.2 Expression and purification of maltose binding protein constructs

The pMAl-c5X plasmid containing the MBP or MBP-183-222 sequence was transformed into BL21(DE3) cells as described for ataxin-3 and grown overnight on LB-agar plates containing 100 µg mL^{-1} cabenicillin at 37 °C. Single colonies were used in to inoculate 10 mL of LB media containing 100 µg mL^{-1} cabenicillin. The cultures were then grown overnight at 37 °C with shaking. The overnight cultures were used to inoculate flasks containing 1 L LB media and 100 µg mL^{-1} cabenicillin. The cultures were grown at 37 C with 200 rpm shaking until the OD_{600} reached 0.6. At this time the expression of protein was induced by the addition of 1 mM IPTG. The cells were grown for another 4 hours and then harvested by centrifugation at 5000 *xg* for 15 min at 4 °C in a JLA 8.1 rotor.

The subsequent steps of the MBP purification we carried out as described for the ataxin-3 constructs.

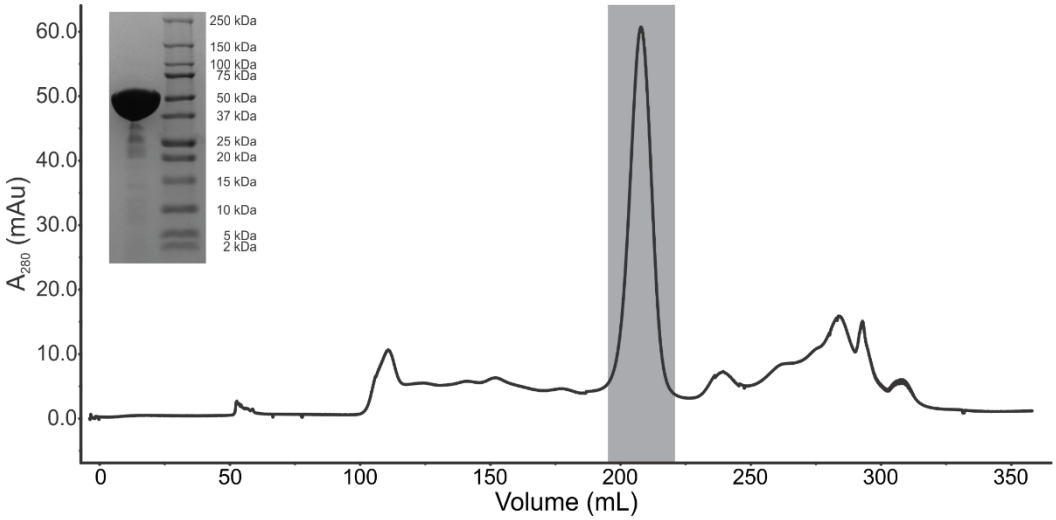


Figure 2.2: The purification of the maltose binding protein – 183-222 construct. The A_{280} trace for the SEC of MBP-183-222. The grey box indicates the fractions collected. Inset: SDS-PAGE of the resulting protein at 100 µM post SEC showing a single intense band at approximately the correct molecular weight (49.2 kDa).

2.9.2.3 Expression of TEV protease

TEV protease was expressed and purified by Ellen Kendirk (University of Leeds). In short pMHTDelta238 containing the his-tagged TEV protease sequence was transformed into BL21-CodonPlus(DE3)-RIPL cells and grown in LB media in the

presence of 50 µg/mL kanamycin at 37 °C with shaking (200 rpm) until the culture reached an OD600 of ~0.6. The temperature was then lowered to 30 °C and expression induced with 0.5 mM IPTG. The cells were harvested by centrifugation and the pellet was resuspened in in 25 mM sodium phosphate buffer, pH 8.0, 200 mM NaCl, 10% (v/v) glycerol, 25 mM imidazole, 1 mM PMSF, 2 mM benzamidine, ~0.02 mg/ml DNase (Sigma, UK), and lysed by sonication (6 x 30 s bursts with 1 min cooling on ice between each sonication).

The lysed cells were centrifuged (20 mins, 4 °C, 39000 *xg*) and the lysate was applied to Ni2+ Sepharose beads (GE Healthcare) and washed twice with 25 mM sodium phosphate buffer, pH 8.0, 200 mM NaCl, 10% (v/v) glycerol, 25 mM imidazole. His-tagged TEV was eluted with 25 mM sodium phosphate buffer, pH 8.0, 200 mM NaCl, 10% (v/v) glycerol, 500 mM imidazole.

The protein solution was filtered (0.2 µM syringe filter, Sartorius, UK) and gel filtered on a HiLoad Superdex 75 26/60 column (GE Healthcare) equilibrated with 25 mM sodium phosphate buffer, pH 8.0, 200 mM NaCl, 25 mM imidazole, 10% (v/v) glycerol, 5 mM β-mercaptoethanol. Peak fractions were concentrated to ~1 mg/mL using Vivaspin 20 (5 kDa MWCO) concentrators (Sartorius, UK), aliquoted, snap-frozen in liquid nitrogen and stored at 80 °C.

2.9.2.4 Expression and purification of Aβ40

The purification of Aβ40 follows the protocol first reported by Walsh *et al.* (Walsh *et al.* 2009) and optimised in our laboratory by Dr Katie Stewart.

BL21(DE3) cells were transformed with pETSac vector containing the Aβ40 sequence by adding 1 µL of DNA to 50 µL of cells. The cells were incubated on ice for 30 mins before being heat shocked by incubation in a water bath at 42 °C for 45 seconds. The cells were left to recover on ice for 10 mins. 500 µL of LB media was added to the cells and the solution incubated for 30 mins at 37 °C for 30 min. 200 µL of the culture was spread on an LB-agar plate containing 100 µg mL^{-1} carbenicillin and incubated at 37 °C overnight.

A single colony was used to inoculate 100 mL of LB media containing 100 µg mL^{-1} carbenicillin and incubated at 37 °C with 200 rpm shaking overnight. After incubation the OD_{600} of these culture should be 1.0 or greater.

500 mL of LB media containing 100 µg mL^{-1} carbenicillin was prepared in baffled conical flasks to which was added 5 mL of the overnight culture. This solution was incubated at 37 °C with 200 rpm shaking. Once the OD_{600} reached 0.5 (generally 2 to 3 hours) exspression of Aβ40 was induced by the addition of isopropyl-β-D-thiogalacto-pyranoside (IPTG) to a final concentration of 1 mM and then grown for a further 4 hours. Two 500 mL cultures were combined in to a single 1 L centrifuge bottle and centrifuged for 15 min at 6000 rpm at 4 °C in a JLA 8.1 rotor. The pellets were collected and frozen after which they were stored at -20 °C.

The cell pellets were thawed on ice and 12.5 mL of 10 mM Tris, 1 mM EDTA, pH 8.5 added to each pellet. 0.5 mg of DNase and 300 µL of PMSF (10 µg/mL) were added to the solution. The pellets were then homogenised by mechanical stirring overnight at 4 °C. The following morning the solution was pulled through a 50 mL syringe with a 1 1/2" mixing needle and incubated at 4 °C for 30 minutes to further homogenise the solution. The solution was then sonicated for 30 seconds at 4 W. The sonicated solution was poured in to centrifuge tubes which had been preincubated on ice. These were then centrifuged for 15 min at 18000 rpm at 4 °C in a JA 25.5 rotor. The supernatants were removed and the pellets resuspended in 10 mL of 10 mM Tris, 1 mM EDTA, pH 8.5. The resuspended pellets were incubated at 4 °C with stirring for 1 hour followed by homogenisation with a syringe as described above. The centrifugation step was repeated and the pellets resuspended in 8 M urea. The resuspended pellet wad homogenised by mechanical stirring at 4 °C and then stirred with a benchtop vortexer. The homogenised sample was then sonicated for 30 seconds as described above and centrifuged at 18000 rpm at 4 °C in a JA 25.5 rotor. The supernatant, which was yellow/brown was collected.

25 mL of Q-sepharose resin was washed in 10 mM Tris, 1 mM EDTA, pH 8.5 by shaking the solution to resuspend the resin and then leaving it for one hour to

settle. The buffer was discarded and the process repeated three times. To the washed resin was added the supernatant diluted 1:4 with 10 mM Tris, 1 mM EDTA, pH 8.5. The mixture was placed on a rocker at 4 °C for 30 min.

The resin solution was next poured into a Buchner funnel with a 70 mm number 1 Whatman filter paper circle. The mixture was stirred with a glass rod and allowed to settle for 5 min. After 5 min the resin and buffer had separated and a vacuum was applied. The flow through was collected. This process was repeated as shown in table 2.15.

Table 2.15: Solutions conditions used in the purification of Aβ40

	Solution
Wash 1	10 mM Tris, 1 mM EDTA, pH 8.5
Wash 2	10 mM Tris, 25 mM NaCl, 1 mM EDTA, pH 8.5
Elution 1-4	10 mM Tris, 125 mM NaCl, 1 mM EDTA, pH 8.5
Elution 5	10 mM Tris, 250 mM NaCl, 1 mM EDTA, pH 8.5
High salt wash 1	10 mM Tris, 1 M NaCl, 1 mM EDTA, pH 8.5
High salt wash 2	10 mM Tris, 1 M NaCl, 8 M Urea 1 mM EDTA, pH 8.5

The elution profile of Aβ40 was confirmed by SDS-PAGE and the fractions containing large amounts of Aβ40 were combined (generally elutions 1-5). The Q-sepharose resin was reequilibrated by washing with water twice and 10 mM Tris, 25 mM NaCl, 1 mM EDTA, pH 8.5 six times. The flow through and wash 1 were added to the resin and the steps detailed above repeated up to elution 3. Again the elution profile of Aβ40 was confirmed by SDS-PAGE and the relevant fractions combined with those from the first round of purification.

The Aβ40 containing fractions were dialysed into 50 mM ammonium bicarbonate. The solution was poured into snake skin dialysis tubing (3500 MWCO) and the sealed dialysis bags added to 5 L of 50 mM ammonium bicarbonate. Over the course of two days the 5L of 50 mM ammonium bicarbonate was changed four times. The resulting solution was poured into 50 mL flacon tubes (15 mL in each) and lyophilised by freezing in a dry ice/ethanol slurry and freeze drying for 2-3 days.

A Superdex 75 16/60 SEC column with a column volume of 330 mL was equilibrated into 50 mM ammonium bicarbonate in a cold box at 4 °C. A tube (or tubes) of lyophilised Aβ40 was resuspended in 5.1 mL of 50 mM Tris, 7 M guanidine HCL. The solution was centrifuged for 10 min at 14000 rpm at 4 °C on a benchtop centrifuge to remove aggregates. The solution was added to a 5 mL loop on the AKTA prime plus and injected onto the column. The column run is detailed in table 2.16.

Table 2.16: Details of the AKTA run for the purification of Aβ40

Time	Flow rate	Fractions	Valve position
0 min	2.0 mL min^{-1}	-	Load
2 min	2.0 mL min^{-1}	-	Inject
10 min	2.0 mL min^{-1}	-	Load
25 min	2.0 mL min^{-1}	4.0 mL	Load
180 min	2.0 mL min^{-1}	-	Load
181 min	0.1 mL min^{-1}	-	Load

Monomeric Aβ40 eluted in fractions 33-39. The whole peak was collected as a second SEC step is performed later. The solution is lyophilised as described above with 2 mL of solution in a 15 mL falcon tube. 500 µL is lyophilised in a 1 mL Eppendorf for concentration determination.

The concentration of Aβ40 was determined by resuspending the 500 µL Aβ40 sample in 200 µL of 50 mM Tris, 7 M guanidine HCL and recording the A260, A280 and A340 on a dual beam spectrophotometer. The extinction coefficient of Aβ40 under these conditions is 1490 M^{-1} cm^{-1}.

In order to achieve a higher level of purity and to remove any dimer or higher order species from the sample a second round of SEC was performed. This was performed on an analytical Superdex S75 column with a column volume of 24 mL. The column was equilibrated into 50 mM ammonium bicarbonate. 2.5-3.0 mg of Aβ40 was resuspended as before and injected into the loop of the AKTA prime plus. The column was run at 0.5 mL min^{-1} collecting 1 mL fractions. The second half of the elution peak (fractions 11-13) was collected to avoid contamination

with dimer. The Aβ40 was freeze dried as described previously in 0.5 mL or 1.0 mL aliquots with 200 µL saved for determination of the final concentration of the samples.

An example of an Aβ40 purification is shown in figure 2.3. As described above elutions 1-5 were combined and the flow through and wash 1 were combined and repurified using the Q-sepharose resin.

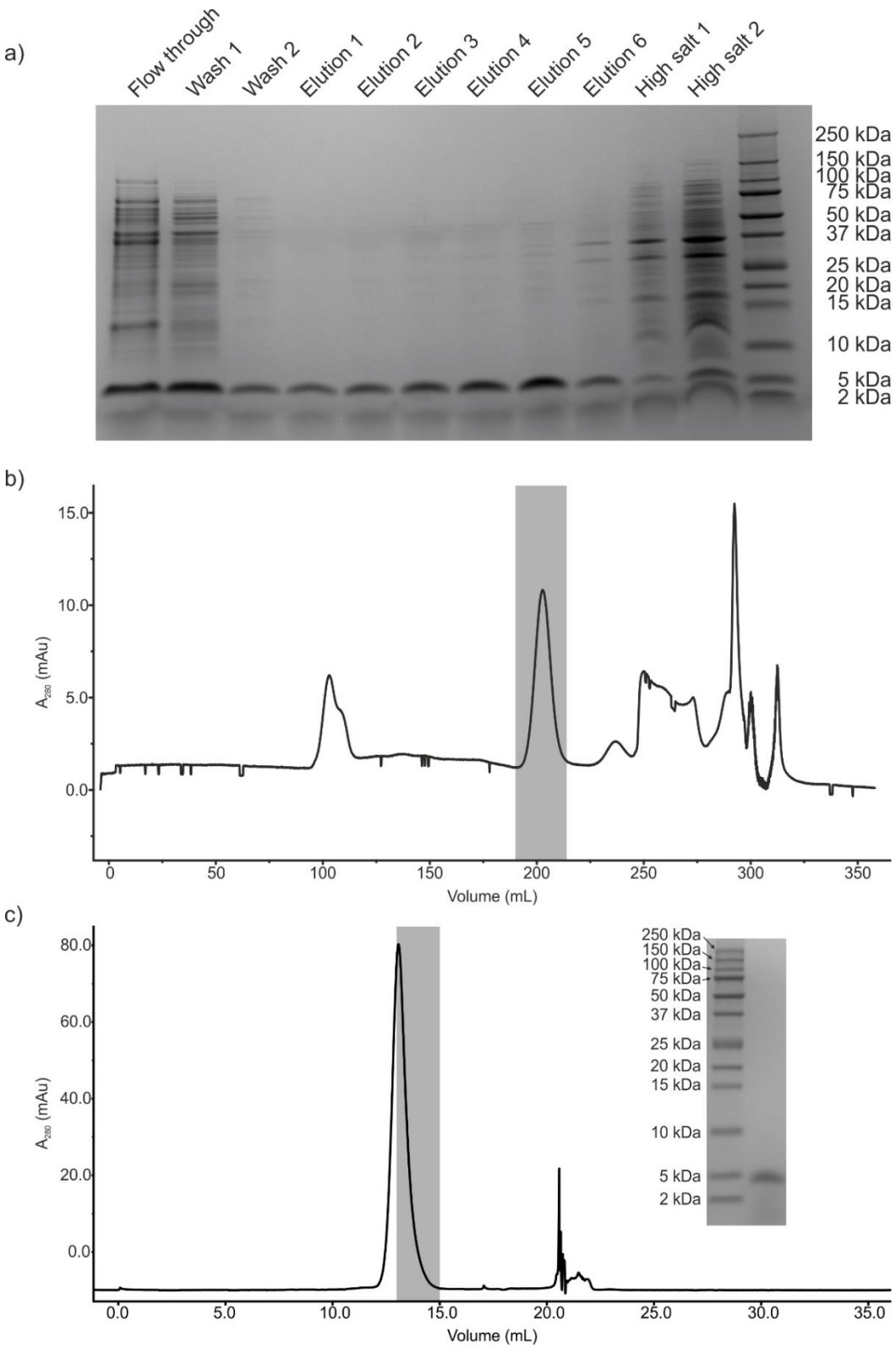


Figure 2.3: The purification of Aβ40. (a) The SDS-PAGE of the fractions recovered from the first round of the Q-sepharose purification. (b) The A_{280} trace of the first round of SEC. The grey box indicates the factions collected and taken forward. (c) The A_{280} trace of the second round of SEC. The grey box indicates the factions collected and taken forward. The peak at 21 mL is from guanidine HCl. Inset: An SDS-Page of the purified Aβ40 at 25 µM.

2.9.2.5 Expression and purification of α-synuclein

α-synuclein was expressed from a pET23a vector. The pET23a-α-syn was transfected into BL21(DE3) cells which were then grown overnight on LB-agar plates containing 100 µg mL^{-1} carbenicillin. A 100 mL of LB was inoculated with a single colony and grown overnight in the presence of 100 µg mL^{-1} carbenicillin.

The following day 10 1 L LB flask containing 100 µg mL^{-1} carbenicillin were inoculated with 10 mL each of the overnight culture. The cultures were grown for four hours at 37 °C with 200 rpm (by which time the OD_{600} was approximately 0.6) and then expression was induced by the addition of 0.5 mM IPTG. Post induction the cells were grown for a further 4 hours with the same conditions described above. The cells were then harvested by centrifugation at 6000 *xg* for 15 min at 4 °C in a JLA 8.1 rotor. The resulting cell pellets were resuspended in 25 mM Tris-HCl, pH 8.0 and stored at -20 °C.

The resuspended pellets were thawed and 100 µg mL^{-1} lysozyme, 50 µg mL^{-1} PMSF and 20 µg mL^{-1} DNase added to the solution. The cell pellets were homogenised in a cell disruptor (30000 psi) and then heated to 80 °C in a water bath for 10 mins. The heated homogenate was then centrifuged in a JA 25.5 rotor at 30000 *xg* for 30 min at 4 °C and the pellet discarded. The proteins contained in the supernatant were precipitated with 50% (w/v) ammonium sulfate for 30 min at 4 °C. The precipitated protein was collected by centrifugation in JA 25.5 rotor at 30000 *xg* for 30 min at 4 °C and the precipitation step repeated. The precipitated protein was again centrifuged and the pellet resuspended, this time in 20 mM Tris-HCl, pH 8.0.

The resuspended protein was loaded onto a 500 mL Q-sepharose anion exchange column preequilibrated in 20 mM Tris-HCl, pH 8.0 attached to an AKTA prime plus. A linear gradient from 0-500 mM NaCl was used to elute the proteins and the absorbance at 280 nm monitored to observe elution. The presence of α-syn was confirmed by SDS-PAGE and the relevant fractions were dlaiyed into 50 mM ammonium bicarbonate using snakeskin dialysis tubing with a 2500 MWCO. The dialysed protein was lyophilised by freeze drying for 2-3 days.

The lyophilised protein was resuspended in 20 mM sodium phosphate, pH 7.5, and loaded onto a preequilibrated HiLoadTM 26/60 Superdex 75 prep grade gel filtrations column attached to an AKTA prime plus. The flow rate of the AKTA was set to 2 mL min-1 with 2.5 mL fractions collected. The fractions containing monomeric α-syn were collected. The central 50% of the peak was collected in order to avoid contamination with higher order species or with truncations of α-syn which have been observed in some preparations. The presence and purity of α-syn were confirmed by SDS-PAGE and the relevant fractions combined and dialysed into 50 mM ammonium bicarbonate. The dialysed protein solution was the lyophilised by freeze drying for 2-3 days and stored at -20 °C.

An example of an α-synuclein preparation SEC is shown in Figure 2.4.

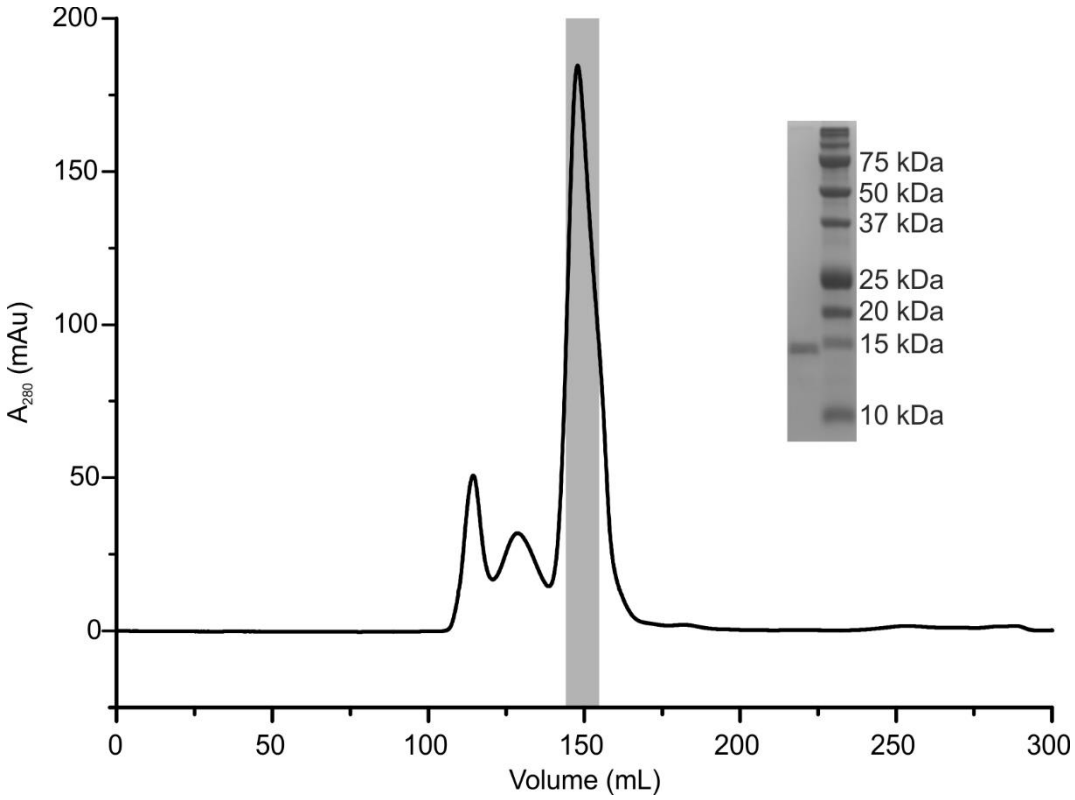


Figure 2.4: An example α-synuclein SEC A_{280} profile. Several peaks are observed corresponding to higher order species. The grey box shows the fractions collected as monomeric protein. Inset: SDS-PAGE of purified α-synuclein.

2.9.2.6 Expression and purification of YDL085CA

YDL085CA protein was expressed and purified by Prof. James Bardwell (University of Michigan). The protein was expressed as a His-SUMO-YDL085CA fusion protein.

The pET28b-YDL085CA plasmid was transfected into BL21(DE3) cells and grown on LB-Agar plates containing 50 µg mL^{-1} kanamycin. Single colonies were selected and used to inoculate 25 mL LB media containing 50 µg mL^{-1} kanamycin and grown overnight. These cultures were used to inoculate terrific broth media and grown at 37 °C with 200 rpm shaking until they reached an OD_{600} of 0.6. YDL085CA expression was then induced by the addition of 1 mM IPTG and grown overnight at 22 °C. The cells were collected by centrifugation followed by disruption of the cells by means of a French press (1300 psi). The lysate was purified by nickel affinity chromatography the subsequently digested overnight with ULP1 to cleave the SUMO tag. The digested samples was again purified by nickel affinity chromatography. The fractions containing YDL085CA were identified by SDS-PAGE. These fractions were then concentrated, snap frozen in liquid nitrogen and stored at -80 °C.

Note that the ULP1 cleavage sequence was modified so that the C-terminal serine was replaced with the N-terminal methionine of YDL085CA. Efficient cleavage was still observed and the purified protein lacked the additional N-terminal serine residue.

2.9.2.7 TEV cleavage of MBP-183-221

MBP-183-221 was digested with TEV in order to separate the fusion protein domain from the ataxin-3 derived sequence. His-tagged TEV protease was added at a 1:10 TEV:MBP-183-221 molar ratio and incubated for 1 hour 45 minutes at 20 °C. These conditions were found to be optimal for the digestion, longer digest times and higher concentrations resulted in poor protein recovery. MS analysis of the digested products observed the presence of MBP but the ataxin-3 derived sequence could not be observed.

2.9.3 Native IMS-MS

All native IMS-MS experiments were performed using a Synapt G1 HDMS. Samples were prepared as indicated and introduced to the MS by nanoESI via palladium coated borosilicate capillaries produced in house. All samples were analysed in positive ion mode. The m/z was calibrated using 2 mg/mL sodium iodide (NaI) or 10 mg/mL aqueous caesium iodide (CsI) across the acquisition range required for the experiment. Data acquisition and processing was performed using the MassLynx and driftscope software supplied with the MS.

Samples were thawed and buffer exchanged into appreciate buffer by means of zebaspin desalting columns where possible (for proteins and peptides under 7.5 kDa drop dialysis or simple dilution was used). The samples were then diluted as indicated in the experimental sections. Buffers were ammonium bicarbonate or ammonium acetate based for native samples.

Typical instrumental conditions varied by protein and are described in Table 2.17. Theoretical masses for each protein are provided in Table AI.1.

Table 2.17: Instrumental conditions for native IMS-MS of proteins in this study. All proteins were investigated using positive polarity.

Protein	Instrumental conditions
Ataxin-3	capillary voltage 1.5 – 1.7 kV sampling cone 60 V trap collision energy 15 V transfer collision energy 25 V IMS wave velocity 250 m/s IMS wave height 8 V
UIMs	capillary voltage 1.2 – 1.5 kV sampling cone 10 V trap collision energy 4 V transfer collision energy 6 V
Aβ40/YDL085CA	capillary voltage 1.5 kV sampling cone 170 V trap collision energy 5 V transfer collision energy 5 V IMS wave velocity 300 m/s IMS wave height 6 V
α-synuclein	capillary voltage 1.5 kV sampling cone 60 V trap collision energy 5 V transfer collision energy 5 V IMS wave velocity 300 m/s IMS wave height 8 V

The TWIMS cell was calibrated using native protein calibrants from the Bush database (Bush *et al.* 2010). The typical calibrants used were: cyctochrome C, concanavalin A, β-lactoglobulin, alcohol dehydrogenase, and avidin. Calibrant drift times were recorded under the same conditions as the sample spectra. The quoted CCS values were calculated from the logarithmic fit produced using the method described by Bush *et al.* and Ruotolo *et al.* (Bush *et al.* 2010; Ruotolo *et al.* 2008).

2.9.4 LC-MS

Intact mass analysis was performed on an online M-class ACQUITY UPLC (Waters UK, Manchester, UK) interfaced to a Synapt G2S Q-IMT-TOF mass spectrometer (Waters UK, Manchester, UK). 1 µL of 5 µM sample was injected onto a MassPREP protein desalting column (Waters UK, Manchester, UK) washed with 10% (v/v)

acetonitrile in water, 0.1% formic acid for 5 min at 25 µL min^{-1}. The bound protein was eluted by a gradient of 2-40% (v/v) acetonitrile in water, 0.1% formic acid over 1 min at 25 µL min^{-1}. Typical instrument conditions are given in Table 2.18. Mass calibration was performed via MS/MS of a lockmass solution consisting of 250 fmol $µl^{-1}$ [Glu]-fibrinopeptide b.

Table 2.18: Typical LC-MS conditions for intact mass analysis

Condition	Setting
Polarity	Positive
Capillary voltage	3.0 kV
Sample cone voltage	20 V
Source offset voltage	80 V
Backing pressure	7.9 mbar
Trap bias	4.1 V
Source temperature	100 °C
Desolvation temperature	250 °C
Buffer gas	Argon ($9.1x10^{-3}$ mbar)

Sequencing of peptides digested in solution was performed on a Xevo G2-XS quadrupole time-of-flight mass spectrometer (Waters Corp., Manchester, UK) equipped with an ACQUITY M-Class UPLC (Waters UK, Manchester) with a Symmetry C18 trap column (180 µM i.d. 20 mm) and a HSS T3 C18 (75 µm i.d. 150 mm) analytical column (Waters UK, Manchester). Samples were injected onto the trap column and washed 1% acetonitrile/0.1% formic acid for 5 min at 5 µL min^{-1} and eluted with 1-60% acetonitrile, 0.1% formic acid in water, 0.1% formic acid over 30 min. at 0.3 µL min^{-1}.

Cross linking samples from in-gel digests were analysed using a Synapt G2-Si HDMS quadrupole ion-mobility time-of-flight mass spectrometer (Waters Corp., Manchester, UK). The peptides were separated as described for sequencing of in-solution digests. The typical instrumental settings for the Synapt G2-Si HDMS are shown in Table 2.19.

Table 2.19: Typical LC-MS/MS conditions for sequencing of digested peptides

Condition	Setting
Polarity	Positive
Capillary voltage	3.0 kV
Sample cone voltage	40 V
Source offset voltage	80 V
Backing pressure	3.58 mbar
Trap bias	2.0 V
Source temperature	80 °C
Desolvation temperature	250 °C
Buffer gas	Argon ($8.6x10^{-3}$ mbar)

Mass calibration and lock mass calibration were performed via MS/MS of a lock mass solution consisting of 250 fmol μl^{-1} [Glu]-fibrinopeptide b. The lock mass calibration was checked every 30 seconds. Data was acquired using a data independent acquisition method. A 0.2 s scan window was acquired across a range of 350-2000 m/z followed by a 0.5 second MS/MS acquisition of the five most intense ions with the collision energy applied dependent on the mass and charge of the ion. Only 2^+ and 3^+ ions were selected for MS/MS.

2.9.5 Crosslinking

Crosslinking of Aβ40 and YDL085CA was undertaken using the crosslinking reagent bis(sulfosuccinimidyl)suberate (BS^3). A variety of concentrations of BS^3 were used during optimisation of the experiment as detailed in the experimental section. In addition an isotopically labelled BS^3 was also utilised.

The protein-crosslinker mixture was incubated at room temperature for 1 hour and the reaction quenched by the addition of 50 mM Tris-HCl. The samples were then examined by SDS-PAGE and the relevant bands extracted by cutting them out of the gel using a clean scalpel. The extraction and digestion (below) were performed in a laminar flow hood to prevent contamination with environmental proteins.

The extracted bands were cut into small pieces (approximately 1 mm^3) with a scalpel. The gel pieces were then destained by incubation in 30% ethanol at 60 °C

for 30 min with shaking. The ethanol solution was discarded and the wash repeated until the gel pieces were completely destained. The gel pieces were then washed in 50% acetonitrile/25 mM ammonium bicarbonate for 10 mins with shaking. The solution was removed and the gel pieces were dehydrated by addition of 100% acetonitrile for 10 min with shaking. The dehydrated gel pieces turned opaque upon dehydration. The excess acetonitrile was removed and the residue evaporated by leaving the tubes open in the laminar flow hood for 1 hour.

The gel pieces were incubated on ice and precooled 0.02 µg μL^{-1} trypsin in 25 mM ammonium bicarbonate added. The gel pieces were rehydrated on ice for 30 min. Once the gel pieces were transparent the excess trypsin was removed and the gel pieces covered with 25 mM ammonium bicarbonate. The solution was then incubated at 37 °C with shaking for 18 hours.

Subsequent to the trypsin digest the solution was centrifuged using a benchtop microcentrifuge and 50 µL of water added. After a 10 min incubation with shaking the solution was removed and stored in an Eppendorf tube containing 5 µL of 60 % acetonitrile/5 % formic acid (v/v). 50 µL of 60 % acetonitrile/5 % formic acid (v/v) was added to the gel pieces and incubated for 10 min with shaking. The solution was removed and combined with the previously extracted solution. This process was repeated twice more and the solutions recovered pooled. The combined solution was evaporated using a speed vac until only 5-10 µL of sample remained (usually 2-3 hours). The sampele was then diluted to 20 µL with 5% acetonitrile/0.1 % formic acid (v/v) for LC-MS/MS analysis.

Note that neither Aβ40 nor YDL085CA contain disulphide bonds and as such there was no requirement for reduction and alkylation so this step was omitted.

Data analysis was performed using the PEAKS software (version 8.0) and StavroX (version 3.0.6.1). Data files were refined using the default settings in PEAKS. The refined data files were exported as .MGF files and imported into Stavrox. In StavroX the isotopically labelled BS^3 was manually added to the crosslinker library. Mass comparison settings were: precursor precision 3.0 ppm, fragment

ion precision 0.8 Da, b and y ions. The score settings were: score cut off 10.0, score settings slow, precise scoring and decoy analysis was applied using a shuffled sequence that maintained protease sites. The FASTA file used for the sequence search contained the sequences of Aβ40 and YDL085CA only. The crosslinks were manually validated in the StavroX software and in MassLynx.

Figures were prepared with the xVis server. The StavroX data was exported as csv files and any N-terminal crosslinks (listed as position 0) were manually altered to position 1 (the first residue) in order to facilitate the generation of diagrams. Only crosslinks with StavroX scores that had a FDR of less than 0.05 were included in the data set.

2.9.6 ThT fluorescence assay

All ThT fluorescence assays were performed on BMG labtech plate readers (FLUOstar Omega, FLUOstar Optima and CLARIOstar). In the case of the Omega and Optima the excitation and emission wavelengths were selected via filters while for the clariostar they were set by means of a monochromator. In all cases the top optic was used. The excitation wave length was set to 440 nm and emission was read at 475 nm for the Omega and Optima. For the Clariostar the emission was 440 nm ± 10 nm and emission was read at 475 ± 10 nm.

Data was exported and analysed in Microsoft excel and Origin pro 8.6. All data was corrected to base line by the subtraction of the lowest data point within a data set for each well subtracted from all values. Where normalised data are shown the values are normalise between 0 and 1 for each well individually. Quoted lag times were determined by applying a linear fit to the region from 0.2 to 0.6 of the relative signal for each well and extrapolating the x-intercept. Quoted gradients of the linear growth phase were determined form the same linear fit described for lag times. All fitting was performed in Origin 8.6.

The conditions for the assays varied and are described for each protein below.

2.9.6.1 Aβ40

Aβ40 aggregation assays were performed in 50 mM ammonium bicarbonate, pH 7.4 and 20 mM sodium phosphate, 0.2 mM EDTA, 0.02% sodium azide (w/v), pH 7.4. The assays were performed at 37 °C without shaking in Corning® 96 well half area plates.

2.9.6.2 Ataxin-3

The ataxin-3 aggregation assays were performed in 250 mM ammonium bicarbonate, 1 mM DTT, pH 7.8 at 37 °C without shaking. The protein was 10 µM and the ThT was also added at 10 µM. The assays were performed in Corning Costar® 96-Well black polystyrene plates.

2.9.6.3 α-synuclein

The aggregation of α-syn was investigated in at 80 µM in 20 mM Tris-HCl, 200 mM NaCl, pH 7.5 with 10 µM ThT. The samples were incubated at 37 °C in Corning® 96 well half area plates with orbital shaking at 600 rpm. The top optic was used to read the fluorescence and orbital averaging was applied.

2.9.7 Transmission electron microscopy

The formation and morphology of fibrils formed by the amyloid proteins ataxin-3, Aβ40 and α-syn were examined by negative stain transmission electron microscopy. 15 µL samples were pipetted onto the surface of carbon coated formvar grids provided by the University of Leeds. The sample was left for 15 seconds and then removed by blotting with filter paper. The grid was washed by the addition of 15 µL of water which was then removed as described previously. This wash was then repeated. 15 µL of 4% (w/v) uranyl acetate was added for 15 seconds and blotted. A final wash of water was added and the water allowed to evaporate.

The images were captured using a Jeol JEM 1400 transmission electron microscope operating at 120 kV. Some images were edited to increase contrast in order to increase clarity upon printing. This was performed in Coreldraw.

2.9.8 Nuclear Magnetic Resonance spectroscopy

Nuclear Magnetic Resonance spectroscopy was performed by Dr Theo Karamanos and Dr Arnout Kalverda (University of Leeds) with samples prepared by the author. T1ρ experiments were performed on ataxin-3 and truncations of ataxin-3.

Unlabelled proteins were prepared in buffer containing 10% v/v deuterated water. The data were acquired on a 600 MHz NMR magnet (Oxford Instruments, Abingdon, UK) using a QCI-P-cryoprobe and an Avance III HD console (Bruker Corpn., Coventry, UK) using a 100 ms spin-lock. The data were acquired and processed with Bruker TopSpin 3.5, NMRPipe 8.2 and CcpNMR analysis software 2.4.1. Errors we estimated based on the signal to noise ratio comparing the signal in the region 0.55 to 0.5 ppm to noise in the region 0 to -2 ppm. Errors were then propagated to account for the data processing.

2.9.9 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Tris-tricine buffered SDS-PAGE was used to investigate various aspects of protein purity and to separate corsslinked species in the crosslinking experiments. The gels used are two layered gels consisting of a stacking gel and a resolving gel, 1.5 mm thick. The details of the components of the 30% acrylamide gels used are given in table 2.20.

Table 2.20: Components of a 30% Tris-trycine gel for SDS-PAGE.

Component	Resolving gel	Stacking gel
30% (w/v) Acrylamide:0.8% (w/v) bisacrylamide	7.5	0.83
3 M Tris.HCl, 0.3% (w/v) SDS pH 8.45	5.0	1.55
H2O	0.44	3.72
Glycerol	2.0	-
10% (w/v) ammonium persulphate	0.05	0.10
TEMED	0.005	0.005

APS and TEMED were added immediately before pouring the gel. The resolving gel was poured first to fill 4/5th of the gel volume. This was allowed to settle for several seconds and then the stacking gel gently added directly on top of the

resolving gel. A comb was added to create wells for samples. The gels were allowed to set before use (usually 40 min to 1 hour).

Samples for SDS-PAGE were diluted with 2x or 6x SDS reducing loading buffer and boiled for 5 mins. The samples were then centrifuged at 16000 x g for 15 seconds to recover the solution. 15 µL of sample was loaded into each lane of the gel with any empty lanes (accounting for molecular weight marker lanes) filled with a mixture of water and loading buffer. Molecular weight markers (precision plus protein standards) were added 5 µL to a lane.

The gels were placed in a gel tank with cathode buffer in the inner reservoir and anode buffer in the out reservoir. The gels were attached to a power pack and run at a constant current of 30 mA until the samples entered the resolving gel and at 60 mA thereafter.

Gels were stained in Expedeon Instant Blue Coomassie stain. Gels for images were stained overnight while gels for cross linking were stained for the minimum amount of time required to see bands of interest. Where destaining was required it was performed by incubating the gels in water for 1 hour.

Gels were imaged using an InGenius Gel documentation system (Syngene, Cambridge, UK). Images were altered in order to increase contrast for printing using the CorelDraw software.

2.9.10 Circular dichroism

Far-UV CD was performed using a Chirascan spectrophotometer (Applied Photophysics, Leatherhead, Surrey, UK) with 4 nm bandwidth, in a 1.0 mm cuvette. The sample temperature was controlled by a Peltier temperature controller. The spectra were recorded over the range 180-260 nm. Each spectrum was recorded three times and the average of the three scans is shown.

The secondary structure content of each of the peptides was calculated from these data using the CDSSTR method available on Dichroweb (Compton and Johnson 1986; Whitmore and Wallace 2004).

The peptides UIM1 and UIM12 were each suspended in 50 mM sodium phosphate buffer, pH 8.0 at 0.25 mg/mL and 0.11 mg/mL, respectively. The concentration of the peptides was determined by the Scopes method (Scopes 1974). The spectra were acquired at 20 °C.

The proteins Aβ40 and YDL085CA were each suspended in 25 mM sodium phosphate, pH 7.4 at 22 µM. EDTA was excluded from the buffer to prevent interference from the carboxylate groups of EDTA which absorb below 200 nm.

2.9.11 Protein concentration determination

For the proteins Ataxin-3 (all variants), MBP, α-syn and Aβ40 the concentration of the protein was determined by using the Beer-Lambert law (Grimsley and Pace 2004). The absorbance at 280 nm of the solution was recorded on an Ultrospec 1200 pro spectrophotometer. A sample of buffer alone was used as a blank to correct for background absorbance. The extinction coefficients used are given in Table 2.21.

Table 2.21: Extinction coefficients for the protein used in this study. These values were calculated using ProtParam (*The Proteomics Protocols Handbook* 2005).

Protein	Extinction coefficient
Ataxin-3 78Q	31400
Ataxin-3 14Q	31400
JDU1	29110
JD+	29110
JD	29110
MBP	67840
MBP+	67840
α-syn	5960
Aβ40	1490

YDL085CA lacks lack tryptophan residues and as such does not have a high extinction coefficient at 280 nm. The concentration of YDL085CA was instead determined by means of the Peirce BCA assay.

For UIM1 and UIM12, which lack tryptophan residues and are small peptides which are not recommended for use with the BCA assay, the concentrations of

the peptides were determined from the absorbance at 205 nm on a UV-1800 dual beam spectrophotometer (Shimadzu, Tokyo, Japan) following the Stokes method (Scopes 1974). The absorbance at 205 nm was recorded and an extinction coefficient of 31 was used to determine the concentration. The variation between the extinction coefficients of proteins at 205 is very low, allowing for an accurate concentration to be determined.

2.9.12 Data analysis and statistical methods for Thioflavin T data

The statistical analysis of the data acquired through the ThT assay is described below. The number repeat experiments for each value are described in the figure legends and in the data table provided in the appendices. Data analysis was performed in Microsoft Excel 2013 and Origin Pro 8.6.

The ThT data were first background subtracted by subtracting the mean of the black wells (i.e. those containing only ThT and buffer solution) from all fluorescence values. At this stage the maximal ThT fluorescence signal values were recorded foe each sample.

The data were next normalised between 0 and 1 by first subtracting the minimum value in a given well data set from all other values within that well data set followed by dividing all values in a well data set by the maximum value in that well data set. This normalisation served the dual purpose of making data sets with different maximal fluorescence single more easily comparable and facilitating further analysis. From these data the T_{50} values were determined for each well based on the time value at which the normalised fluorescence first reached 0.5 arbitrary units.

To determine the lag times and elongation rates the linear region of the normalised data was isolated (the range chosen was 0.2 to 0.6 arbitrary units as this was observed to prevent the inclusion of non-linear regions with minimal reduction of the data set). These data were then plotted in Origin Pro 8.6 and independently fitted with linear fits. The gradient of these linear fits was taken as the elongation rate in arbitrary units $hour^{-1}$. Using the find *x* from *y* function in

Origin Pro 8.6, with *y* set to 0, the projected lag time was determined for each sample. At this stage the T_{50} value could also be calculated from *y* = 0.5, however this method of determining the T_{50} was not used.

As the aggregation kinetics of amyloid proteins can vary between experiments the values described above (T_{50}, lag time and maximal ThT fluorescence) were normalised across experiments by dividing all vales obtained by the average of the relevant amyloid protein alone control. These ratios were then compared between experiments.

Data are presented as both ThT fluorescence traces and as box plots. Box plots were generated in Origin Pro 8.6 with 25th and 75th percentile boxes and whiskers representing outlier values within 1.5 inter quartile ranges of the mean. Data points outside of the whiskers were not shown. This method of presentation follows the guidance presented by Nature Methods (Krzywinski and Altman 2014).

Significance p values were calculated in Microsoft Excel 2013 using the t-test. The data were treated as describing a population with a two-tailed distribution containing two samples with unequal variance. P values were calculated for each treated population compared to the untreated control. In the case of YDL085CA, each concentration used was compared to the untreated control and to each other concentration. T-test p values of below 0.05 were taken as cause to reject the hypothesis that the two groups were sampling the same population.

3 YDL085CA – a modulator of amyloid protein aggregation

3.1 Introduction

3.1.1 Introduction to MOAG-4

As discussed in Chapter 1.2, many diseases are caused by the aggregation of amyloid proteins. As such, it could be considered surprising that proteins which increase the rate of amyloid formation occur in nature. However, several instances of proteins that accelerate amyloid protein aggregation have been described and may be associated with disease states (Goehler *et al.* 2004; Muchowski *et al.* 2002; Sittler *et al.* 1998).

The protein Modifier of aggregation 4 (MOAG-4) was identified as a modulator of age-related phenotypes and death in *C. elegans* expressing polyglutamine proteins (van Ham *et al.* 2010). Initial screening identified the deletion of MOAG-4 as reducing the severity of the polyQ phenotype. Further investigation demonstrated that deletion of MOAG-4 reduced the rate of aggregate formation by 75%. The activity of MOAG-4 was shown to be autonomous (i.e. mediated by MOAG-4 alone) suggesting that MOAG-4 may represent an easily manipulated mechanism of aggregation enhancement (van Ham *et al.* 2010).

MOAG-4 is an 82 residue protein with a theoretical molecular weight of 8.9 kDa. MOAG-4 is part of a highly conserved family of orthologous proteins which occur in both fungi and animals (van Ham *et al.* 2010). In all cases, the MOAG-4 family are characterised by having large numbers of charged residues leading to an overall basic charge (example sequences are shown in Figure 3.1a). Lysine in particular is over-represented amongst the residues of these proteins. Two conserved regions are found: one at the N-terminus (residues 1 to 24 of MOAG-4) and one in the central region of MOAG-4 (residues 38 to 63). NMR data (Figures 3.1b and c) for MOAG-4 suggest that parts of these regions form α-helices (residues 8-22 and 45-70) (Yoshimura *et al.* 2017). In the case of the N-terminal helix this secondary structure is likely to be transient (Yoshimura *et al.* 2017). Secondary structure prediction algorithms suggest that this secondary structure

is maintained in several MOAG-4 orthologs (Figure 3.1d) (Cuff and Barton 1999; Drozdetskiy *et al.* 2015).

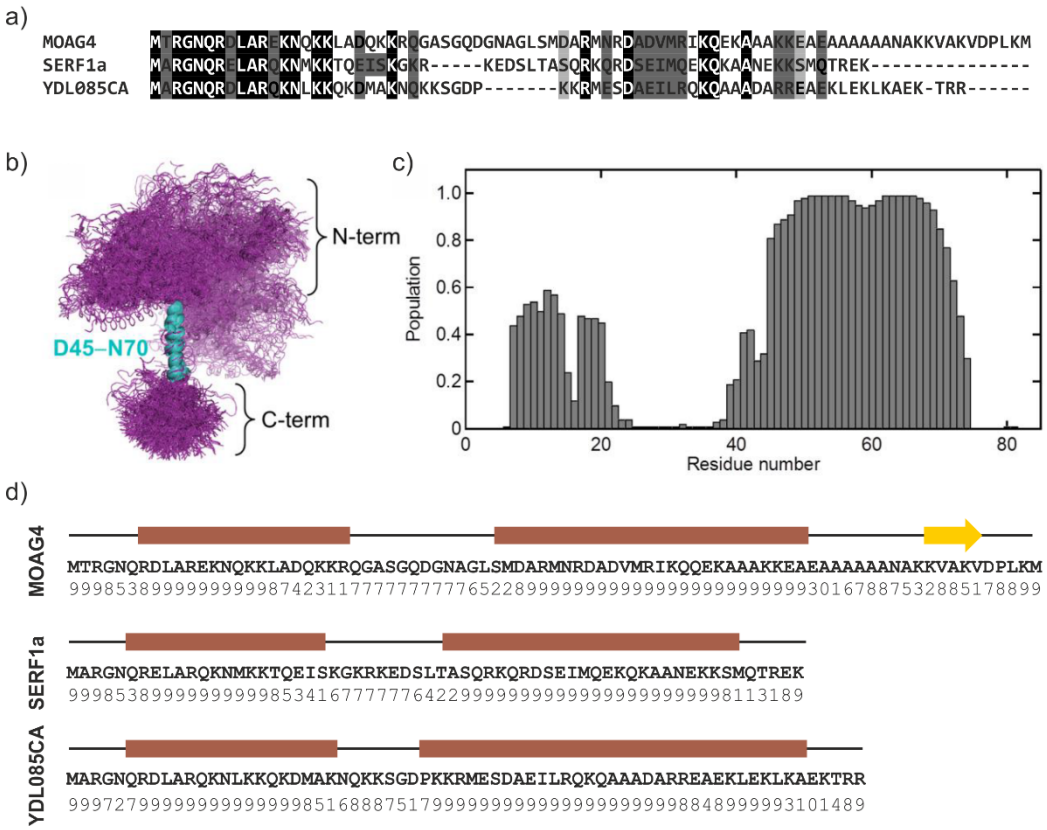


Figure 3.1: Structural relationships of the MOAG-4 protein family. (a) The sequences of MOAG-4, SERF1a and YDL085CA, representatives of the MOAG-4 family from *C. elegans*, *H. sapiens* and *S. cerevisiae* respectively. The sequence alignment shown was performed using BlastP (Altschul *et al.* 1990) with the black boxes indicating preservation of an amino acid and the grey boxes indicating conservation of characteristics. (b) The solution NMR ensemble structure for MOAG-4 depicting a helix from residues D45 to N70. The remaining sequence of MOAG-4 appears to be disordered. (c) The population of helical structure per residues determined from the ensemble structures depicted in b. (d) Secondary structure prediction for the MOAG-4, SERF1a and YDL085CA. α-helices are shown as brown boxes and β-strands as orange arrows as determined by jnetpred. The JENTCONF value is given below each residue, with nine indicating a high degree of agreement between simulations and lower values indicating increasing uncertainty.(Cuff and Barton 1999; Drozdetskiy *et al.* 2015) Panels (b) and (c) reproduced from Mulder and co-workers (Yoshimura *et al.* 2017).

Both MOAG-4 and its human orthologs, SERF1 and SERF2, have been shown to directly alter the aggregation of amyloid proteins (Falsone *et al.* 2012; van Ham *et al.* 2010). Interestingly, the SERF1 isoform SERF1a has been identified as a

modifier of spinal muscular atrophy, a disease in which protein misfolding has been suggested to play a role (Scharf *et al.* 1998). However, the effect of SERF1 on aggregation appears to be selective as the aggregation kinetics of non-amyloid forming aggregating proteins (insulin, actin and citrate synthase) are unaltered upon the addition of SERF1 (Falsone *et al.* 2012). The aggregation process of amyloid-forming proteins, as measured by the ThT-assay, appears to be accelerated in the presence of SERF1 even at substoichiometric levels (Falsone *et al.* 2012). The reduction in the length of the lag phase has been demonstrated for a variety of amyloid forming proteins including Aβ40, α-synuclein, huntingtin exon 1 and the prion protein, PrPc (Falsone *et al.* 2012).

The interaction between SERF1 and α-synuclein leads to a rapid increase in the amount of soluble oligomers of high molecular weight formed by α-synuclein, which were not normally detected during aggregation (Falsone *et al.* 2012). These species were found to be unstable, disappearing over time (3-4 days), although they were ThT positive in their own right, suggesting that they already had at least some of the characteristics of mature amyloid fibrils. When examined by TEM, the morphologies of the ThT positive species were found to be primarily short rods and spherical aggregates, as opposed to the long fibrils observed in the absence of SERF1 (Falsone *et al.* 2012). Over time these short fibrils matured into structures that could not be differentiated from those formed in the absence of SERF1.

Both SERF1 and MOAG-4 appear not to interact with the mature aggregates and do not co-localise with them (van Ham *et al.* 2010). This suggests that monomeric or early stage aggregate species are the primary substrate for SERF interaction. NMR studies of the interaction between SERF1 and α-synuclein demonstrate that C-terminus of α-synuclein is perturbed upon interaction with SERF1 suggesting that binding may occur in this region (Falsone *et al.* 2012). Further NMR studies with MOAG-4 support this proposed binding site, once again showing evidence for the involvement of the C-terminus in this interaction (Yoshimura *et al.* 2017). This second study also showed that the secondary structure of α-synuclein upon binding was largely unchanged from random coil, however, α-synuclein samples

a less compact conformation upon interaction with MOAG-4 (Yoshimura *et al.* 2017).

The effect of MOAG-4/SERF on the aggregation of α-synuclein is theorised to be mediated by the exposure of the central amyloidogenic region of α-synuclein due to the interaction between MOAG-4/SERF and the negatively charged C-terminus of α-synuclein which normally shields the core of α-synuclein (Yoshimura *et al.* 2017). The exposure of normally shielded amyloidogenic regions upon binding to MOAG-4 family proteins may represent a general mechanism by which they accelerate aggregation. Perturbation of the C-terminus of α-synuclein has been shown previously to drive amyloid formation by exposing the hydrophobic core (Fernandez *et al.* 2004; Hoyer *et al.* 2004; Levitan *et al.* 2011; McClendon, Rospigliosi and Eliezer 2009; Ranjan and Kumar 2017; Rasia *et al.* 2005).

While the interaction between MOAG-4/SERF1 and α-synuclein is now well defined (Falsone *et al.* 2012; Yoshimura *et al.* 2017), there are no published data investigating the interaction between MOAG-4 family proteins and other amyloid proteins in the same molecular detail. This chapter aims to characterise the interaction of the MOAG-4 ortholog YDL085CA and the amyloid protein Aβ40 by means of MS based techniques. In addition, a more detailed kinetic analysis of the effects of YDL085CA on the aggregation of amyloid proteins is presented.

3.1.2 Amyloid-β

The majority of the experiments described in this Chapter were performed on Aβ40. As described in the Chapter 1.2, Aβ40 is a 40 residue peptide involved in the pathology of Alzheimer's disease. The sequence of Aβ40 (and the related Aβ42) is shown in Figure 3.2.

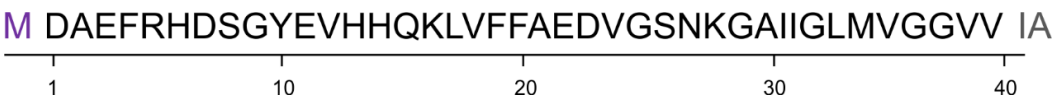


Figure 3.2: The sequence of Aβ40 and Aβ42. The primary sequence of Aβ40/42 is shown. The additional residues present in Aβ42 are shown in grey. The N-terminal methionine residue present in the recombinant Aβ40 used in many of the experiments in this investigation is shown in blue.

Two forms of Aβ40 were used in this investigation. The recombinantly expressed Aβ40 was expressed and purified as described in Chapter 2.9.2.4. As a result of the expression mechanisms employed, this Aβ40 contains an additional methionine residue at the N-terminus which is not present in the endogenously occurring Aβ40 found in disease (Figure 3.2). Many publications have made use of MAβ40 (Assarsson *et al.* 2014; Cukalevski *et al.* 2015; Meisl *et al.* 2014b) and it has been shown to behave in a similar way to Aβ40 (Walsh *et al.* 2009). It was recently shown that MAβ42 and Aβ42 form the same fibrillar structures (Silvers *et al.* 2017).

The second form of Aβ40 used in this investigation is chemically synthesised Aβ40. The synthetic Aβ40 was purchased from BAchem (Bachem AG, Bubendorf, Switzerland). Synthetic Aβ40 does not contain the N-terminal methionine residue and is identical in sequence to the Aβ40 found expressed endogenously in cells. Where there was concern that the N-terminal methionine might play a role in an effect observed for the MAβ40, synthetic Aβ40 was used to repeat the experiment.

Throughout this thesis the recombinantly expressed MAβ40 will be referred to as Aβ40 with notice given when synthetic Aβ40 is used.

3.1.3 Aims of this chapter:

In this Chapter the effect of the yeast homolog of MOAG-4/SERF1, known as YDL085CA, upon the aggregation of the amyloid protein Aβ40 is investigated to determine whether the acceleration of amyloid aggregation previously observed in the presence of MOAG-4/SERF1 (Falsone *et al.* 2012; van Ham *et al.* 2010) is specific to those proteins. The relationship between the concentration of YDL085CA and the effect observed is explored. In addition chemical crosslinking MS is utilised to examine the interaction between YDL085CA and Aβ40.

3.2 YDL08CA modulates the aggregation of amyloid proteins

The effect of the addition of YDL085CA on the aggregation kinetics of Aβ40 were investigated by means of the ThT assay. Falsone *et al.* demonstrated that the effect of SERF1 on α-synuclein aggregation shows a concentration dependence, though the discontinuous nature of the data and the wide errors associated with the measurements make further analysis of the relationship difficult (Falsone *et al.* 2012). In the same publication the addition of equimolar SERF1 to Aβ40 resulted in a reduction in the lag time. Again, the discontinuous nature and sparsity of the data make further analysis problematic however, unlike α-synuclein, the total signal intensity for the Aβ40 treated with SERF1 does not appear significantly different from that observed for Aβ40 alone.

The aggregation kinetics described here were undertaken in the Aβ40 aggregation buffer described by Knowles and co-workers (20 mM sodium phosphate, 0.2 mM EDTA, 0.02% sodium azide, 25 µM ThT, pH 7.4) in order to match previous literature for Aβ40 (Meisl *et al.* 2014a). This varies from the buffer used in previous investigations of SERF1 (138 mM NaCl, 2.7 mM KCl, 0.02% sodium azide, pH 7.4) (Falsone *et al.* 2012; van Ham *et al.* 2010; Yoshimura *et al.* 2017). For instance, the buffer used here buffer lacks sodium chloride.

YDL085CA was expressed and purified by the laboratory of Prof. James Bardwell (University of Michigan). YDL085CA was expressed as sumo-tagged fusion protein and purified by nickel-affinity chromatography. Subsequently, the sumo-tag was removed by digestion with ubiquitin-like specific protease 1 (ULP1). The full details of the expression and purification are given in Chapter 2.9.2.6.

Initial investigation of the effect of YDL085CA on the kinetics of the formation of amyloid by Aβ40 showed that no significant effect ($p=0.40$) was observed in the lag phase upon addition of an equimolar ratio of YLD085CA (Figure 3.3). It does appear that the elongation rate of Aβ40 in the presence of YDL085CA is increased, although not to significance (Aβ40 alone: 0.20 ± 0.06 arbitrary units hour^{-1},

Aβ40+YDL085CA: 0.31 ± 0.06 arbitrary units hour^{-1}, p=0.053) along with the maximal relative fluorescence (Aβ40 alone: 5001 ± 543 arbitrary units, Aβ40+YDL085CA: 7245 ± 834 arbitrary units, p=0.007). Previous work on the ortholog SERF1 had shown a significant decrease in the lag time of Aβ40 fibril formation in the presence of SERF1 (Falsone *et al.* 2012); this finding was therefore somewhat unexpected, although it does represent a single concentration under one set of buffer conditions.

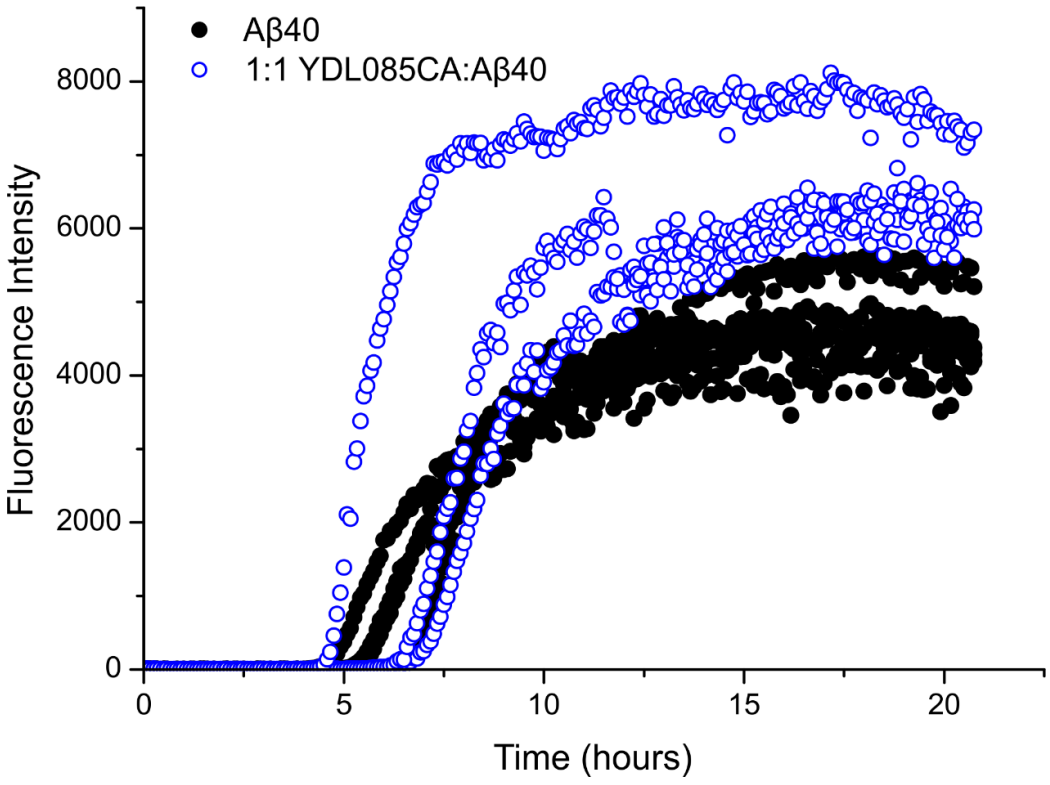
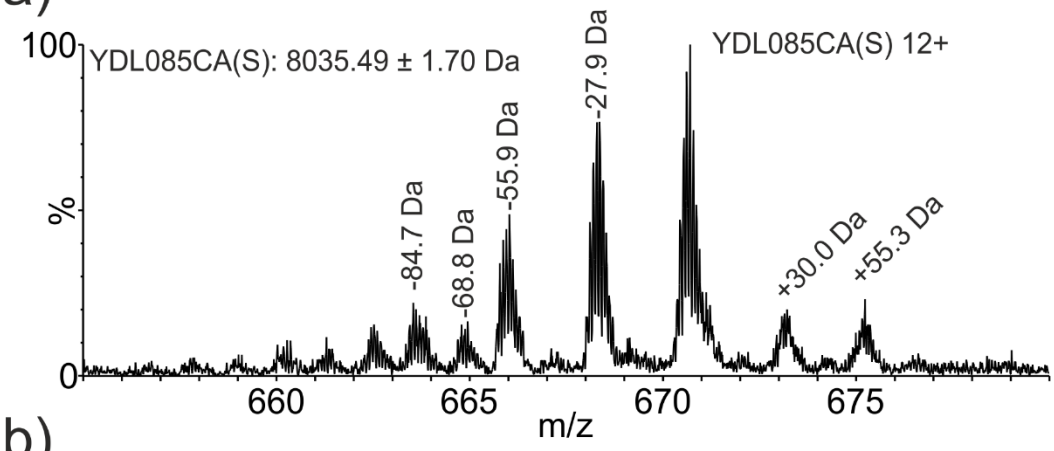


Figure 3.3: Initial experiments suggest that YDL085CA does not modify the lag time or T_{50} of Aβ40 at equimolar concentrations. The aggregation of Aβ40 was measured by means of the ThT fluorescence assay. The samples were 25 µM Aβ40 +/- 25 µM YDL085CA in 20 mM sodium phosphate, 0.2 mM EDTA, 0.02% sodium azide, 25 µM ThT, pH 7.4 incubated at 37 °C without shaking.

LC-MS analyses of intact YDL085CA reveal a mixed population consisting of several species of different masses (Figure 3.4). Several of these mass differences correspond to amino acid substitutions; however, sequencing of the plasmid revealed that no mutations had occurred.

a)



b)

Mass Change	Substitution(s)
+55.3Da	M→W, T→R
+30.0 Da	G→S, A→T, R→W, T→M, V→E
-27.9Da	R→Q, R→K, M→C, D→P, V→A
-55.9 Da	L→G, I→G, double substitution from the row above
-68.8Da	R→S
-84.7Da	Triple substitution from the -27.9 Da row

Figure 3.4: Intact LC-MS of YDL085CA indicates various amino acid substitutions. The primary peak corresponds to the expected mass for YDL085CA including a serine from the sumo tag cleavage however several other major peaks are also observed. The peaks correspond to various single or double amino acid substitutions shown in the table (b). Substitutions indicated in grey are for residues which do not occur in the wild type YDL085CA sequence. The presence of several different peaks indicates the population is a mixture of wild type protein and protein containing different amino acid substitutions. The theoretical mass of YDL085CA(S) is 8036.28 Da.

Digestion of the YDL085CA with the typical enzymes used in the laboratory was deemed unsuitable due to the amino acid content of the protein (Figure 3.5). The large numbers of lysine residues mean that digestion with trypsin or LysC would generate large numbers of very short peptides and so make sequencing difficult. The absence of aromatic residues in YLD08CA results in chymotrypsin being unsuitable for digestion of YDL085CA. Thus, the endopeptidase AspN was selected as digest prediction indicates that digestion with this enzyme would generate peptides ranging from 7 to 19 residues in length with an average length of 11.3 residues.

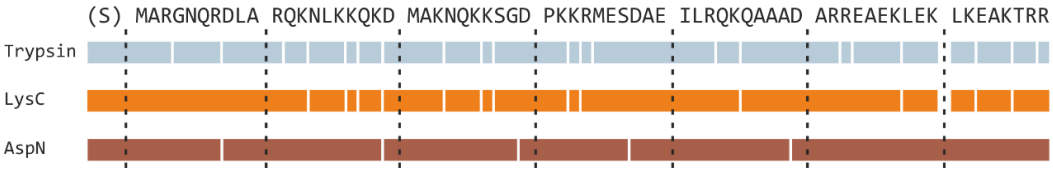


Figure 3.5: Digestion sites of common enzymes on YDL085CA. The predicted peptides after the digestion of YDL085CA by trypsin (blue), LysC (orange) and AspN (red). Chymotrypsin (not shown) resulted in the generation of zero peptides.

Post-digestion LC-MS/MS detected a number of amino acid substitutions in YDL085CA (Figure 3.6). These include substitutions at several arginine residues. The codons encoding these residues were found to be rare in *E. coli* (see Table 2.12). There are instances in published papers describing the incorporation of incorrect amino acid residues at sites of rare codons with other residues by the ribosomes of bacteria (Liu *et al.* 2017; Forman *et al.* 1998). It seems likely that the YDL085CA DNA sequence, which was taken from yeast and poorly optimised for *E. coli*, resulted in a similar process occurring here. Analysis of the DNA sequence encoding YDL085CA in relation to rare codons is presented in the Chapter 2 (Table 2.12). The observation that certain arginine residues are substituted while others are not is consistent with the literature which suggests that the position of a rare codon in a sequence (especially, but not exclusively in relation to other rare codons) can alter the rate of substitution (Liu *et al.* 2017). The amino acid substitutions observed in the post-digest LC-MS/MS data account for the -27.9 Da (R→K and R→Q) and -68.8 Da (R→S) peaks as well as the -55.9 Da and -84.7 Da peaks (double and triple R→K/Q substitutions, respectively). However, the +30.0 Da and +55.3 Da peaks are not explained by the LC-MS/MS data. The R→G and R→N substitutions may be present in the intact LC-MS (Figure 3.4a) at low concentration.

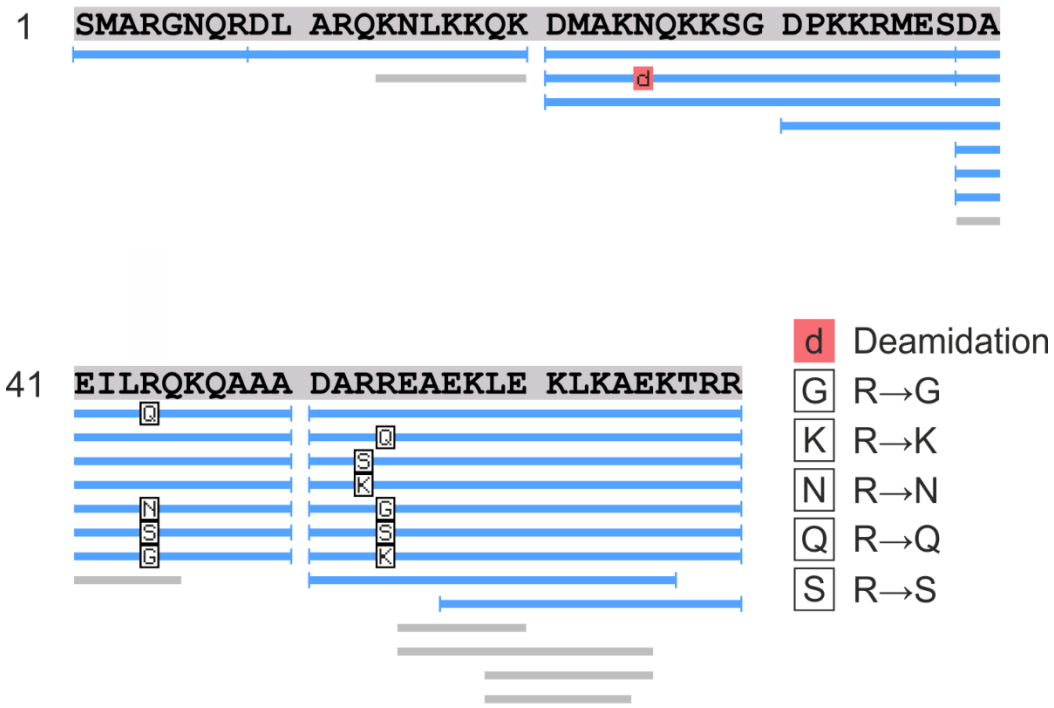


Figure 3.6: Digestion and sequencing of YDL085CA reveals the presence of several amino acid residue substitutions. AspN digest followed by LC-MS/MS resulted in good coverage of the YDL085CA sequence. AspN digestion product peptides are shown in blue beneath the sequence while de novo assembled peptides (corresponding to cleavages at sites not predicted to be cleaved) are shown in grey. Modifications are shown where they were observed and detailed in the key. Several amino acid substitutions were detected at the site of arginine residues.

Given the potential importance of charged residues in the function of YDL085CA, the substitutions observed rendered the samples unsuitable for further analysis. It was also considered that the sequence variation observed may have altered the effect of YDL085CA on the aggregation of amyloid proteins. Codon optimisation for *E. coli* expression of YDL085CA was undertaken and the protein expressed and purified by the laboratory of Prof. James Bardwell (University of Michigan), as described in Chapter 2.9.2.6. The plasmid was also modified to remove the final serine in the ULP1 cleavage site resulting in the purified YDL085CA lacking the N-terminal serine present in the previously expressed protein. This deletion was found not to alter the efficiency of the purification. Intact MS analysis of the *E. coli* optimised YDL085CA revealed a single charge state distribution of the correct mass for the sequence of YDL085CA (Figure 3.7).

In all further experiments protein obtained from the codon optimised plasmid was employed.

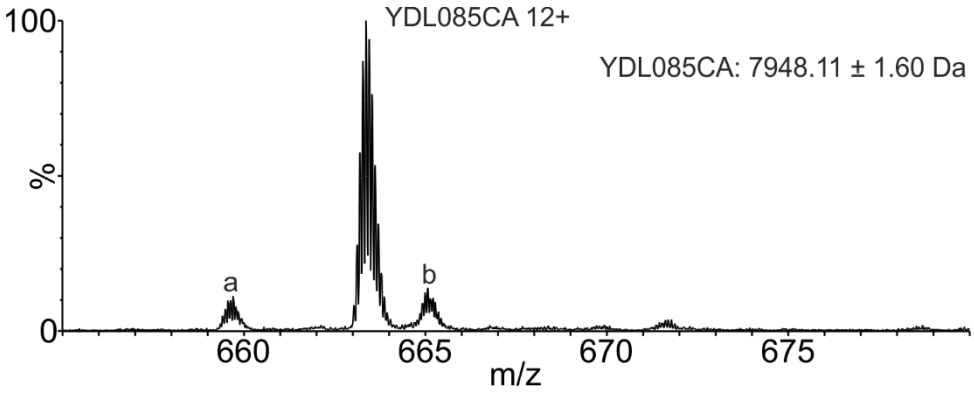


Figure 3.7: Intact LC-MS of the codon optimised YDL085CA reveals a single major species of the correct mass (note that the codon optimised YDL085CA also lacks a serine left over from sumo-tag cleavage). The theoretical mass of YDL085CA is 7949.20 Da. The peaks centred at 659 m/z (labelled a) and 665 m/z (labelled b) are $M+14H^{14+}$ ions from low intensity contaminant ions with masses of 9222 Da and 9298 Da respectively.

3.3 Kinetics of Aβ40 aggregation in the presence of YDL085CA

The effect of the addition of the newly prepared YDL085CA upon the aggregation was again measured by means of the ThT assay. The kinetics of Aβ40 varied considerably from that reported for SERF1a (Falsone *et al.* 2012) where the addition of an equimolar concentration of SERF1a reduced the lag time of fibril formation by around 50%. At equimolar ratios the addition of YDL085CA was observed either to retard the generation of ThT positive species or to cause no change in different experiments (Figure 3.8).

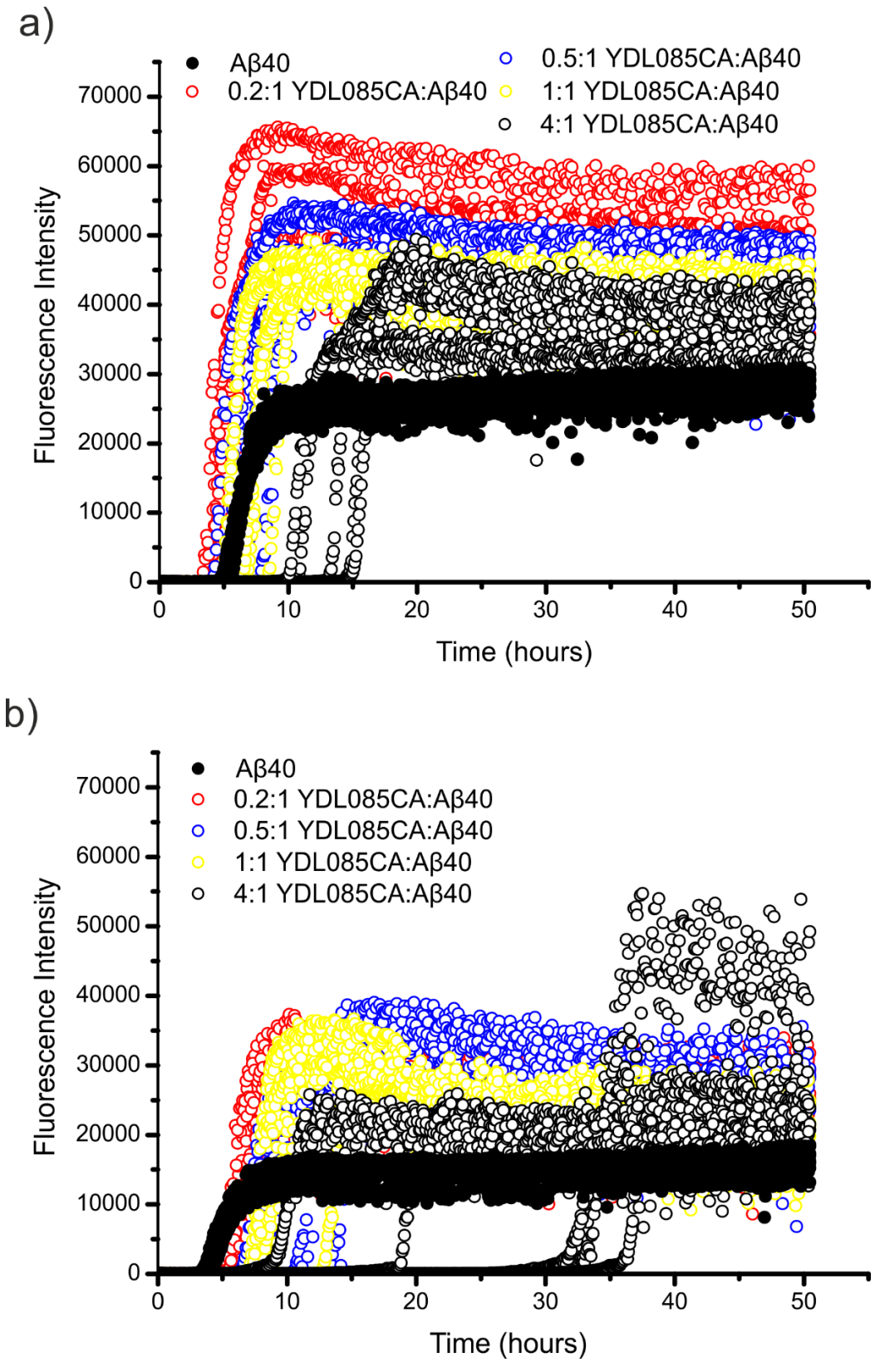


Figure 3.8: Example Aβ40 aggregation kinetics in the presence of YDL085CA. Two repeats of the same experiment are shown, (a) and (b). In both cases the Aβ40 was 25 µM in 20 mM sodium phosphate, 0.2 mM EDTA, 0.02% sodium azide, pH 7.4 incubated at 37 °C without shaking. Note that one sample for 4:1 YDL085CA:Aβ40 in panel (b) has much higher signal than the others but was included in all analysis.

A concentration dependence is observed, with excess YDL085CA causing a significant increase in the lag time of fibril formation (Figure 3.8). Substoichiometric amounts of YDL085CA result in minimal effect, in some cases resulting in a small reduction in the lag time. An analysis of all of the acquired data is presented in Figure 3.9. The combined data set reveals that, while substoichiometric amounts of YDL085CA cause no change in lag time they do result in a small reduction in the T_{50} at a 0.2:1 ratio (an average relative T_{50} of 0.83). When YDL085CA is present in equimolar ratio or at fourfold excess both the lag time and T_{50} of Aβ40 aggregation increase significantly.

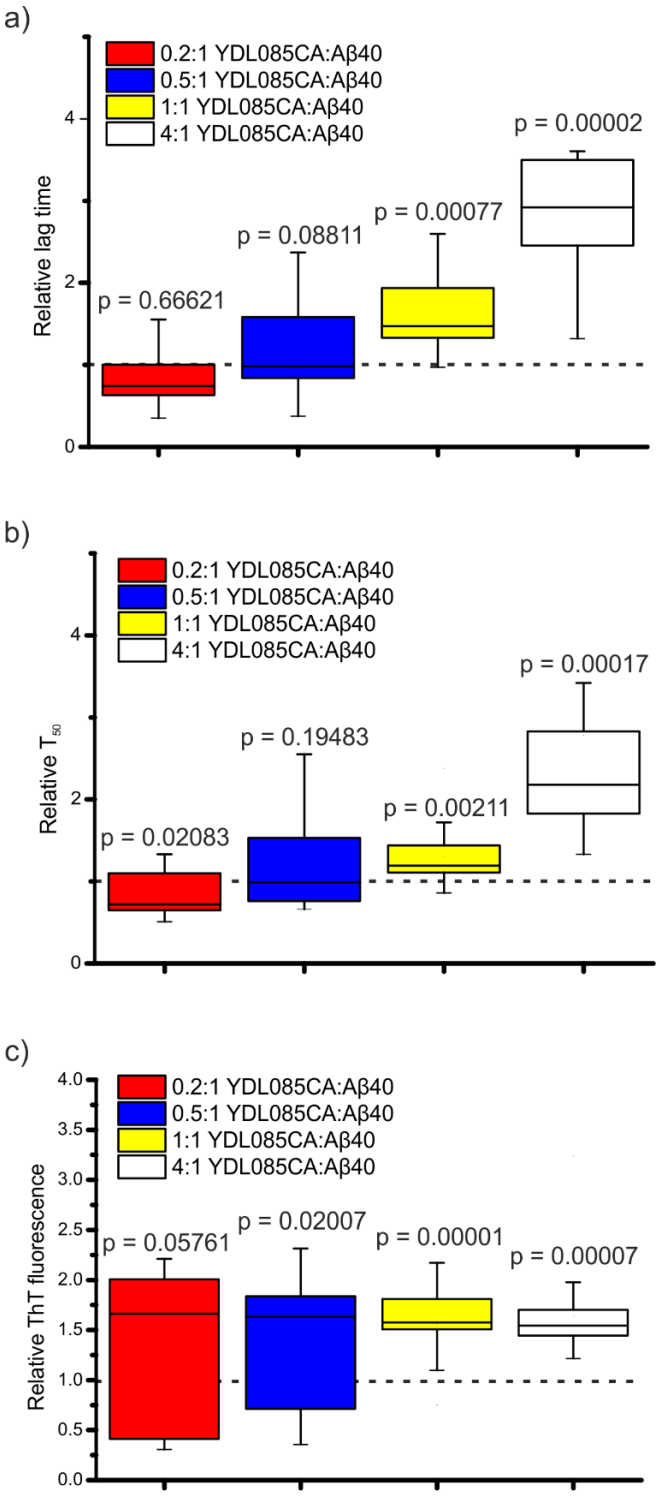


Figure 3.9: Analysis of Aβ40 aggregation in the presence of YDL085CA. (a) The lag time of Aβ40 aggregation, relative to Aβ40 alone, in the presence of increasing concentrations of YDL085CA. (b) The T_{50} of Aβ40 aggregation, relative to Aβ40 alone, in the same experiments. (c) The relative total maximal ThT fluorescence in the same experiments. In all cases the whiskers indicate outliers within the 1.5x IQR. All experiments performed with 25 µM Aβ40 in 20 mM sodium phosphate, 0.2 mM EDTA and 0.02% sodium azide, pH 7.4 at 37 °C without shaking. In (a) n=24, in (b) and (c) n=18. P values are given for a T-test for each data set compared to Aβ40 alone.

In addition to the changes observed in the lag time and T_{50} values of the YDL085CA treated Aβ40, the total fluorescence signal observed also varied when compared to Aβ40 alone (Figure 3.9c). The addition of YDL085CA at any of the investigated concentrations resulted in approximately a 1.5 fold increase in the maximal ThT fluorescence signal on average. However, at lower concentrations of YDL085CA there was greater variation in the effect observed. Maximal ThT fluorescence can, under specific circumstances (Meisl *et al.* 2014a), be used as an indicator of the fibril yield; however, there are many instances of molecules interfering with ThT fluorescence (Coelho-Cerqueira, Pinheiro and Follmer 2014; Gade Malmos *et al.* 2017). As such any conclusions drawn from maximal ThT signal should be considered carefully.

The YDL085CA effect observed is maintained in the reverse experiment wherein the concentration of YDL085CA was kept constant and the concentration of Aβ40 varied (Figure 3.10). As the lag time and kinetics of Aβ40 aggregation are concentration dependent, each concentration of Aβ40 should be compared separately. In all cases, with the exception of the 5 μM Aβ40 experiment, the lag time of Aβ40 aggregation was unchanged. The addition of YDL085CA to 5 μM Aβ40, a fivefold excess of YDL085CA, resulted in an increase in the lag time from around 10 hours to above 22 hours.

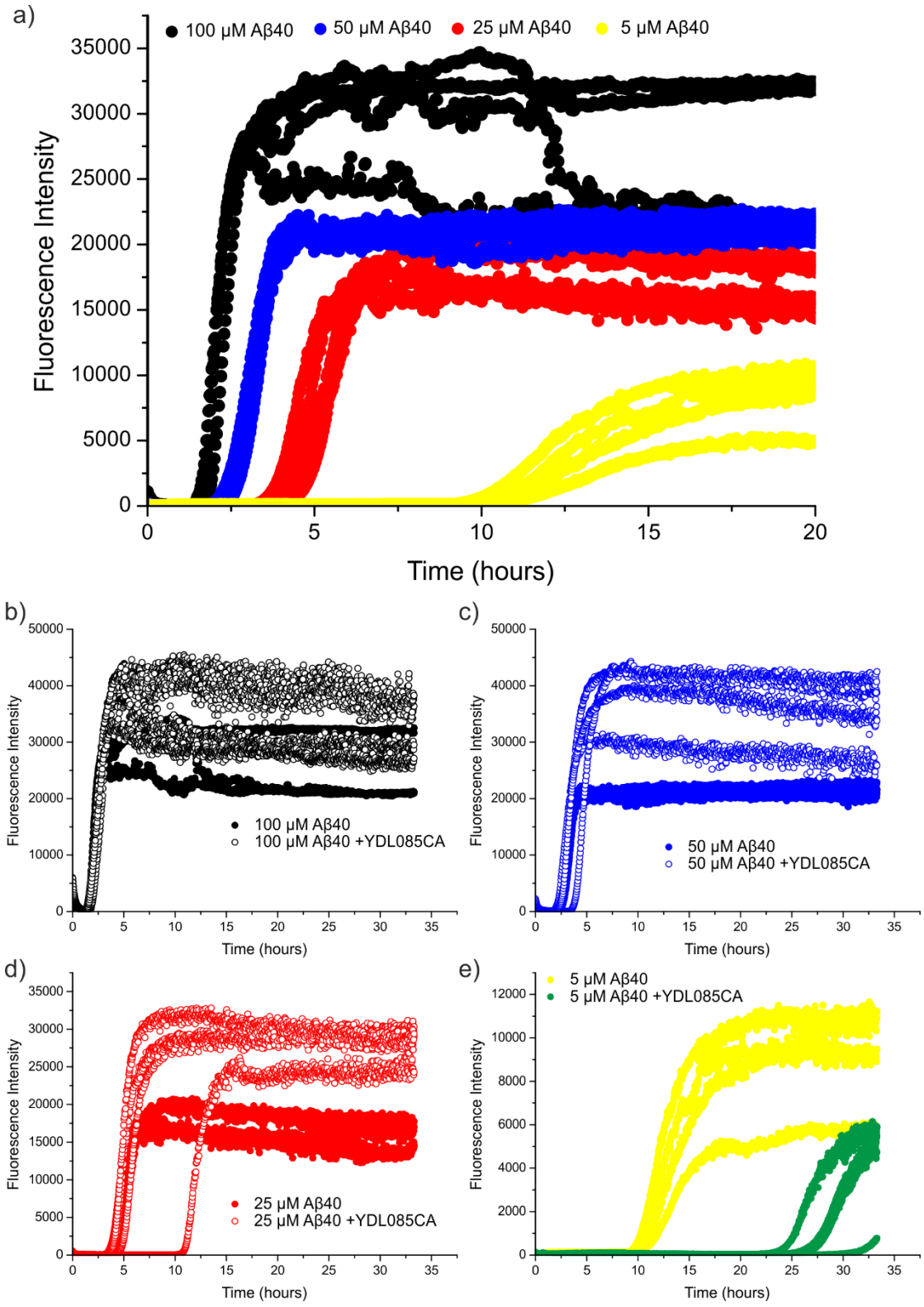


Figure 3.10: Aβ40 aggregation kinetics at different concentrations in the presence of YDL085CA. (a) The kinetics of Aβ40 aggregation vary with concentration. The addition of 25 µM YDL085CA to (b) 100 µM Aβ40 (c) 50 µM Aβ40 and (d) 25 µM Aβ40 cause no change in lag time while the addition of 25 µM YDL085CA to 5 µM Aβ40 (e) results in an increase in lag time. The relative lag times, T50 values and maximum fluorescence are compared in Figure AII.I.

The recombinantly expressed Aβ40 used in the experiments described above includes an N-terminal methionine which does not occur in the natural protein (Figure 3.2). To investigate any possible effect of this residue, a synthetic Aβ40 peptide was purchased. The ThT assay described above was repeated with this Aβ40. The lag time of the Aβ40 aggregation was observed to be slightly shorter (2.7 ± 0.2 hours) than that observed for the same concentration of recombinant Aβ40 (4.3 ± 0.3 hours) (Figure 3.11). Once again, a concentration dependence was observed in the effect of YDL085CA. Similar to the recombinant Aβ40, high concentrations of YDL085CA increase the lag time of the formation of amyloid fibrils. However, in this data substoichiometric concentrations of YDL085CA also caused retardation of aggregation (for the change in lag time for 0.2:1 YDL085CA:Aβ40 $p = 0.00004$ and for 0.5:1 YDL085CA:Aβ40 $p = 0.003$).

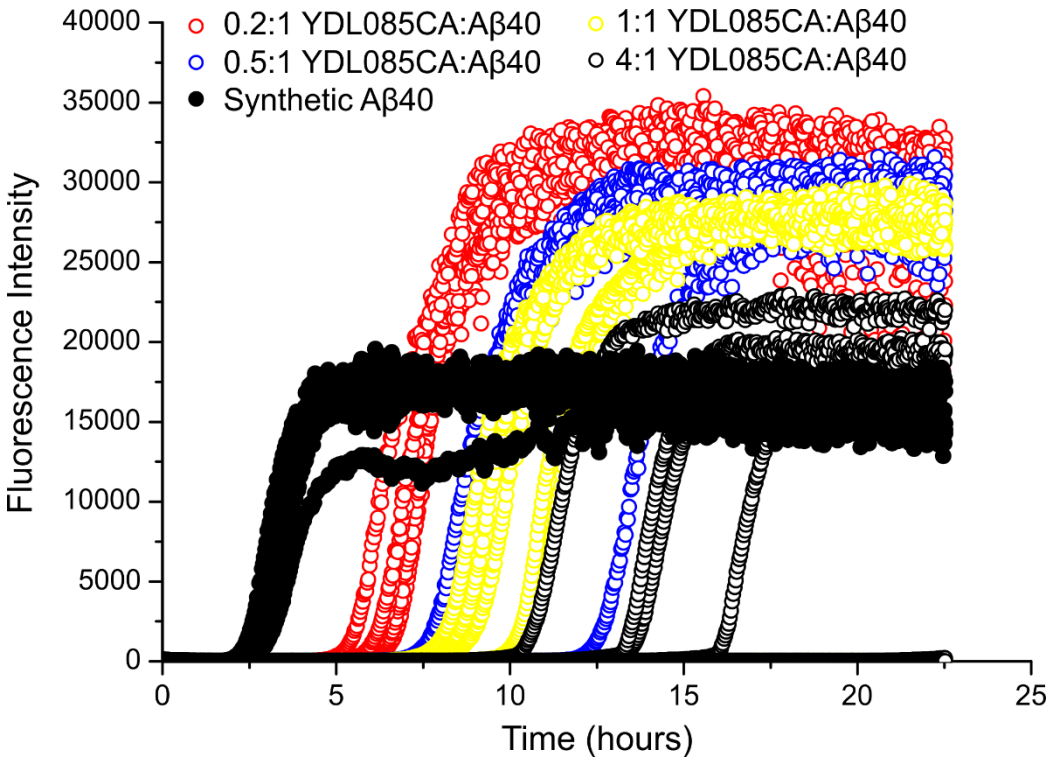


Figure 3.11: Aggregation kinetics for synthetic Aβ40 in the presence of YDL085CA. The Aβ40 was 25 µM in 20 mM sodium phosphate, 0.2 mM EDTA, 0.02% sodium azide, pH 7.4 incubated at 37 °C without shaking. The final concentration of DMSO in each sample was 1% (v/v). The relative lag times, T50 values and maximum fluorescence are compared in Figure AII.II.

Variation in protein concentration, due to poor quantification of the protein, could explain the variation in YDL085CA's activity between experiments.

YDL085CA lacks tryptophan residues and consequentially has an extremely low extinction coefficient at 280 nm making spectrophotometric quantification at this wavelength impossible. The concentration of the YLD085CA stocks were estimated based on the results of a BCA assay (Chapter 2.9.11) (Smith *et al.* 1985). Likewise the extinction coefficient of Aβ40 is very low, decreasing the accuracy of protein concentration determinations by use of this method. Therefore the concentration of the Aβ40 was determined for a highly concentrated stock prior to dilution. A further possibility is that the presence of DMSO in the buffer used in the ThT assay for the synthetic Aβ40, which was not present in the recombinant Aβ40 samples, may have caused the changes seen in the effect of YDL085CA at low concentration.

3.4 The effect of buffer on YDL085CA activity

The sodium phosphate buffer used in the experiments described above is not suitable for native MS analysis. Interestingly, the effects of YDL085CA in ammonium bicarbonate are very different from those seen in sodium phosphate (Figure 3.12). Under these conditions all concentrations of YDL085CA tested resulted in the same reduction in lag time. In addition, a degree of concentration dependence is observed in the final total fluorescence with higher concentrations of YDL085CA resulting in a higher overall fluorescence.

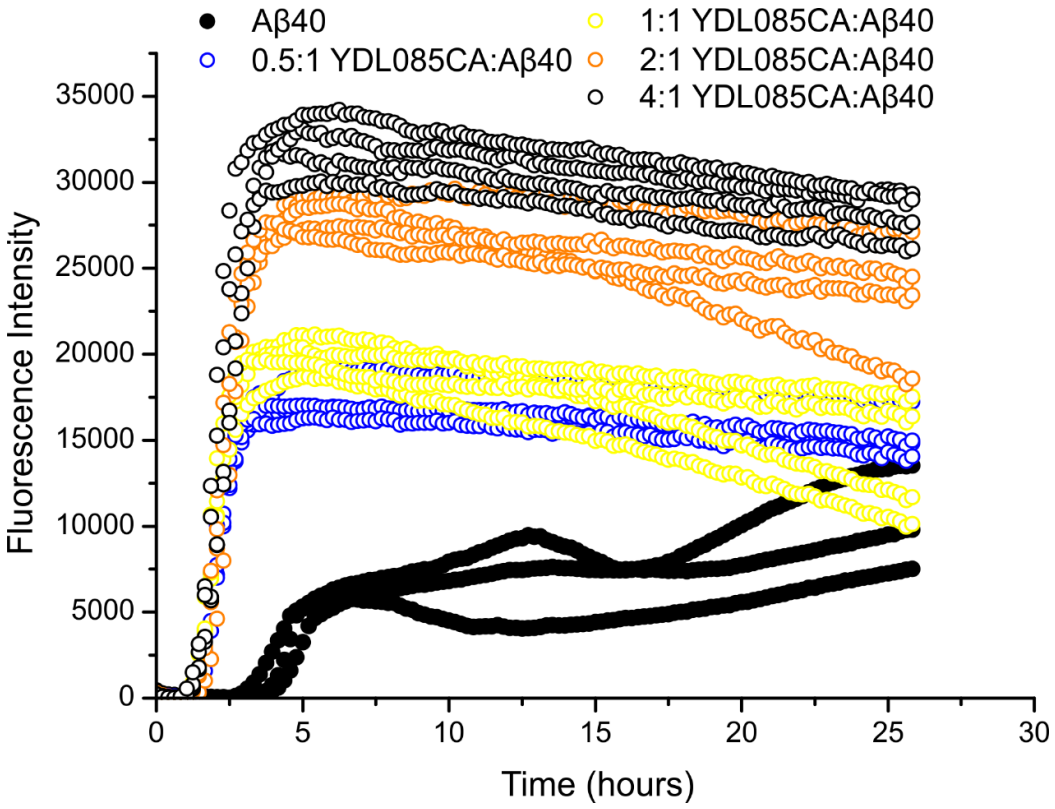


Figure 3.12: Typical kinetics for Aβ40 in the presence of YDL085CA in ammonium bicarbonate. The Aβ40 was 25 µM in 50 mM ammonium bicarbonate pH 7.4 incubated at 37 °C without shaking.

Analysis of multiple data sets for Aβ40 aggregation in the presence of YDL085CA in ammonium bicarbonate shows very different trends from those observed in sodium phosphate (Figure 3.13). While there is little variation in the reduction in the lag time or T_{50} between 0.2:1 and 1:1 ratios, a fourfold excess of YDL085CA does result in a lag time slightly longer than at lower concentrations (T test p value = 0.04 for T_{50} and $p > 0.00001$ for lag time), but still significantly faster than for Aβ40 alone. There does appear to be a concentration effect on the maximal fluorescence, with a fourfold excess of YDL085CA resulting in a much larger increase in maximal signal than lower ratios (Table A.II.5).

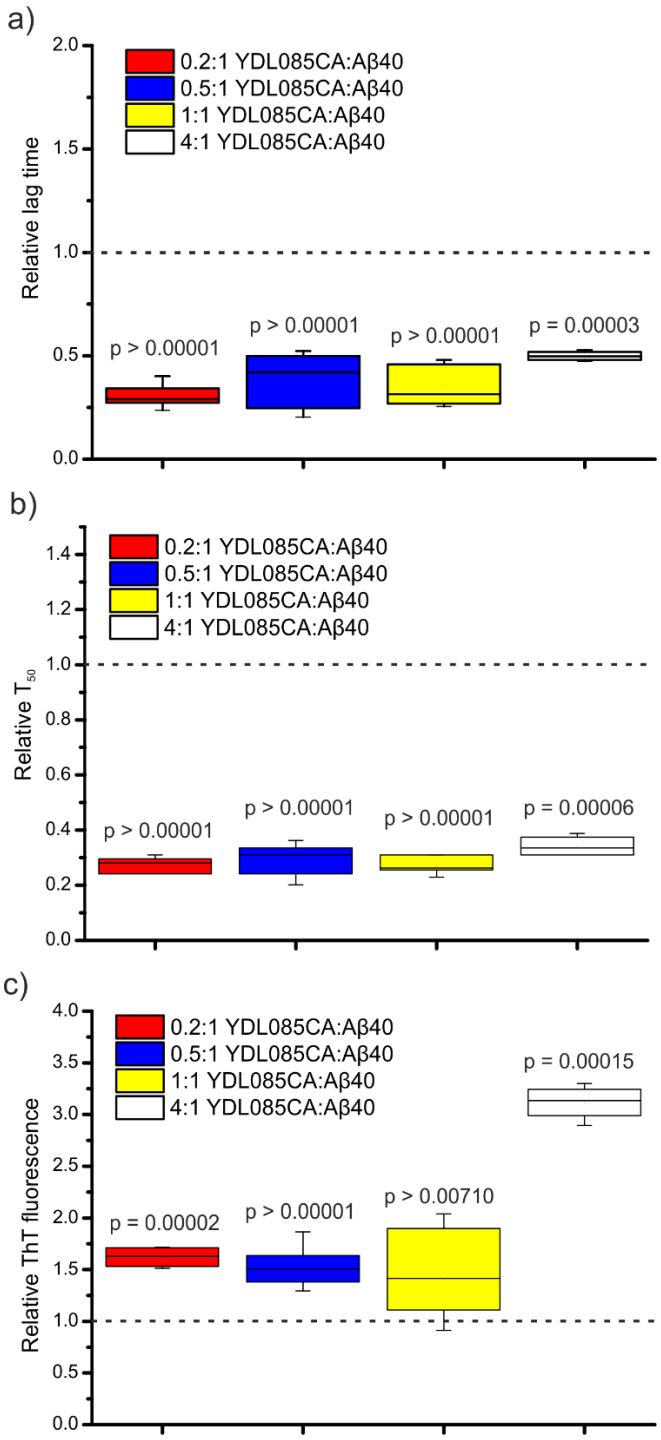


Figure 3.13: Summary of Aβ40 aggregation kinetics in the presence of YDL085CA in ammonium bicarbonate. (a) The lag time of Aβ40 aggregation, relative to Aβ40 alone, in the presence of increasing concentrations of YDL085CA. (b) The T_{50} of Aβ40 aggregation, relative to Aβ40 alone, in the same experiments. (c) The relative total maximal ThT fluorescence in the same experiments. In all cases the whiskers indicate outliers within the 1.5x IQR. All experiments performed with 25 µM Aβ40 in 50 mM ammonium bicarbonate pH 7.8, at 37 °C without shaking. In each case for 0.2:1 n=6, for 0.5:1 and 1:1 n=10 and for 4:1 n=5. P values are given for a T-test for each data set compared to Aβ40 alone.

These data suggest that the effect of YDL085CA is far more complicated than initially thought. It seems that the composition of the buffer in which the experiment is performed can markedly change the effect of YDL085CA with regard to aggregation, changing it from an inhibitor to an accelerator of aggregation. The high number of charged residues present in YDL085CA suggest that the ions present in solution may have a significant effect on its function and structure. In addition, intrinsically disordered proteins such as Aβ40 and YDL085CA have been shown to be especially sensitive to the presence of charged ions in the solution (Liao *et al.* 2013; Wicky, Shammas and Clarke 2017).

Transmission electron microscopy was utilised to examine the morphology of the species in solution at the end of the aggregation assay (26 hours). The TEM images, shown in Figure 3.14, reveal that a significant morphological change occurs when Aβ40 is incubated with YDL085CA. Aβ40 alone forms a dense mat of unbranched linear fibrils several microns in length (Figure 3.14a,b). In the presence of YDL085CA, short crystal-like aggregates are observed which in some cases have the appearance of stacked sheets (Figure 3.14 c-f). This morphology was variable, with longer ribbon like structures also observed (Figure 3.14e) and some fibril-like aggregates also seen (Figure 3.14f). Given the presence of some fibril-like structures it seems possible that the short crystal-like aggregates may go on to form fibrils if left for a sufficient period of time. TEMs were only collected at a single time point therefore it is not possible to comment further on this possibility. Whether YDL085CA is directly remodelling Aβ40 aggregates is unclear. The crystal-like morphology may be a result of the change in the kinetics of aggregation upon the addition of YDL085CA. It can be imagined that increasing the rate of aggregation by reducing the lag phase could be facilitated by increasing the number of critical nuclei and as such there may be less free monomer for extension of the early fibril structures, thus trapping them as shorter structures.

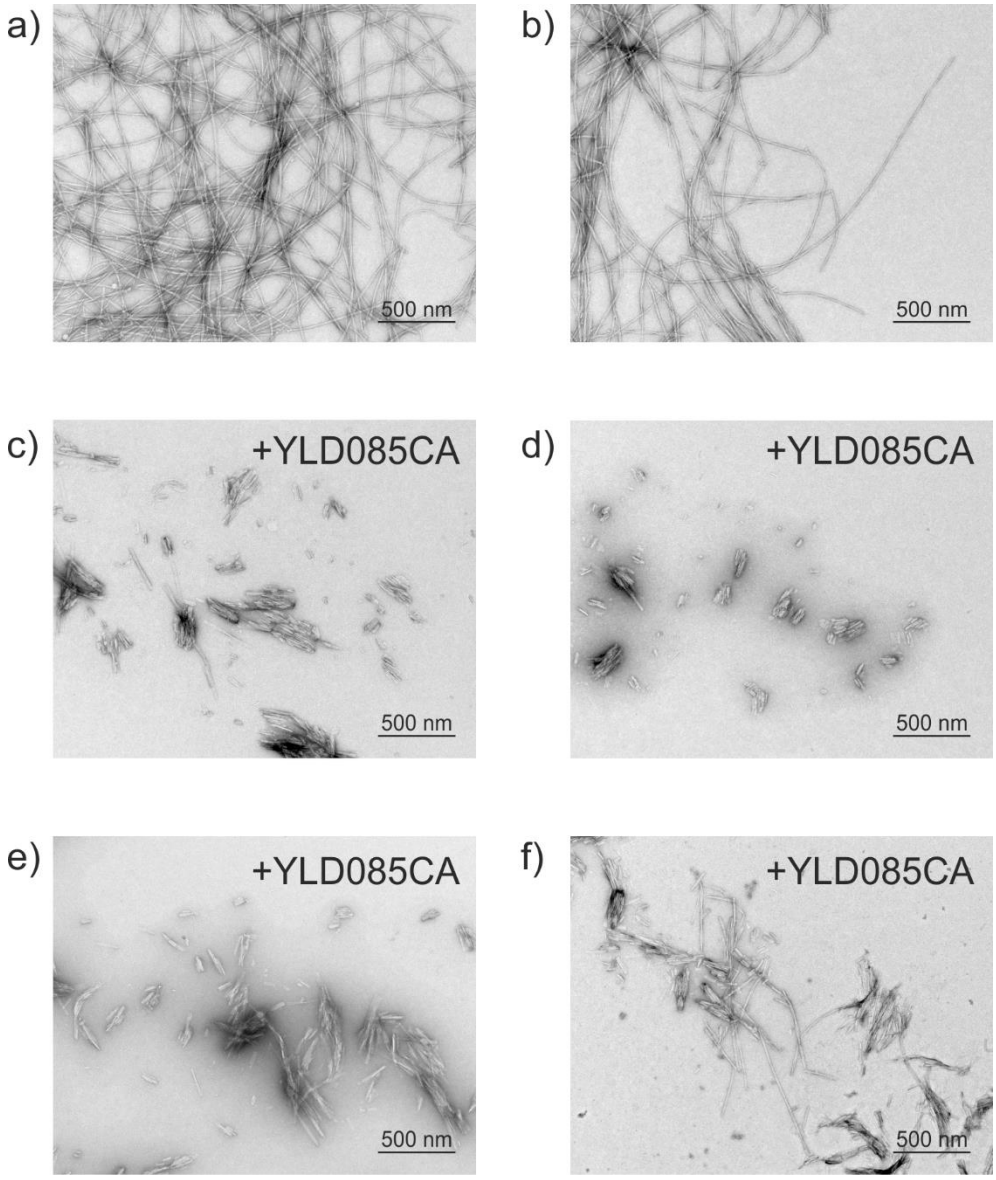


Figure 3.14: Transmission electron microscopy reveals a morphological difference between fibrils formed by Aβ40 in the absence (a,b) and presence (c-f) of YDL085CA. While Aβ40 alone forms long, straight fibrils, in the presence of YDL085CA short crystal-like structures are formed. Some fibrils are still observed. Samples were taken from kinetics assays where Aβ40 was 25 µM in 50 mM ammonium bicarbonate pH 7.4 and was incubated at 37 °C without shaking. YDL085CA was added at equimolar ratio where indicated.

3.5 Native MS investigation of the interaction between YDL085CA and amyloid proteins

Native nESI-MS of YDL085CA was undertaken with the amyloid proteins Aβ40 and α-synuclein, the latter of which has been demonstrated to interact with MOAG-4/SERF1 (Falsone *et al.* 2012; Yoshimura *et al.* 2017). 1:1 ratios of Aβ40/α-synuclein and YDL085CA were prepared in 50 mM ammonium bicarbonate, pH 7.4 and immediately analysed by nESI-IMS-MS. In the case of Aβ40, the use of nESI-IMS-MS revealed the presence of higher order complexes consisting of YDL085CA:Aβ40 complexes with stoichiometries of 1:1, 1:2 and 2:1 as well as a YDL085CA dimer (Figure 3.15). Given the distribution of species it is possible that the observed complexes were the result of the dissociation of a 2:2 complex via subunit ejection in the gas phase due to the energies applied in the instrument. However, no direct evidence for the existence of a 2:2 complex was observed.

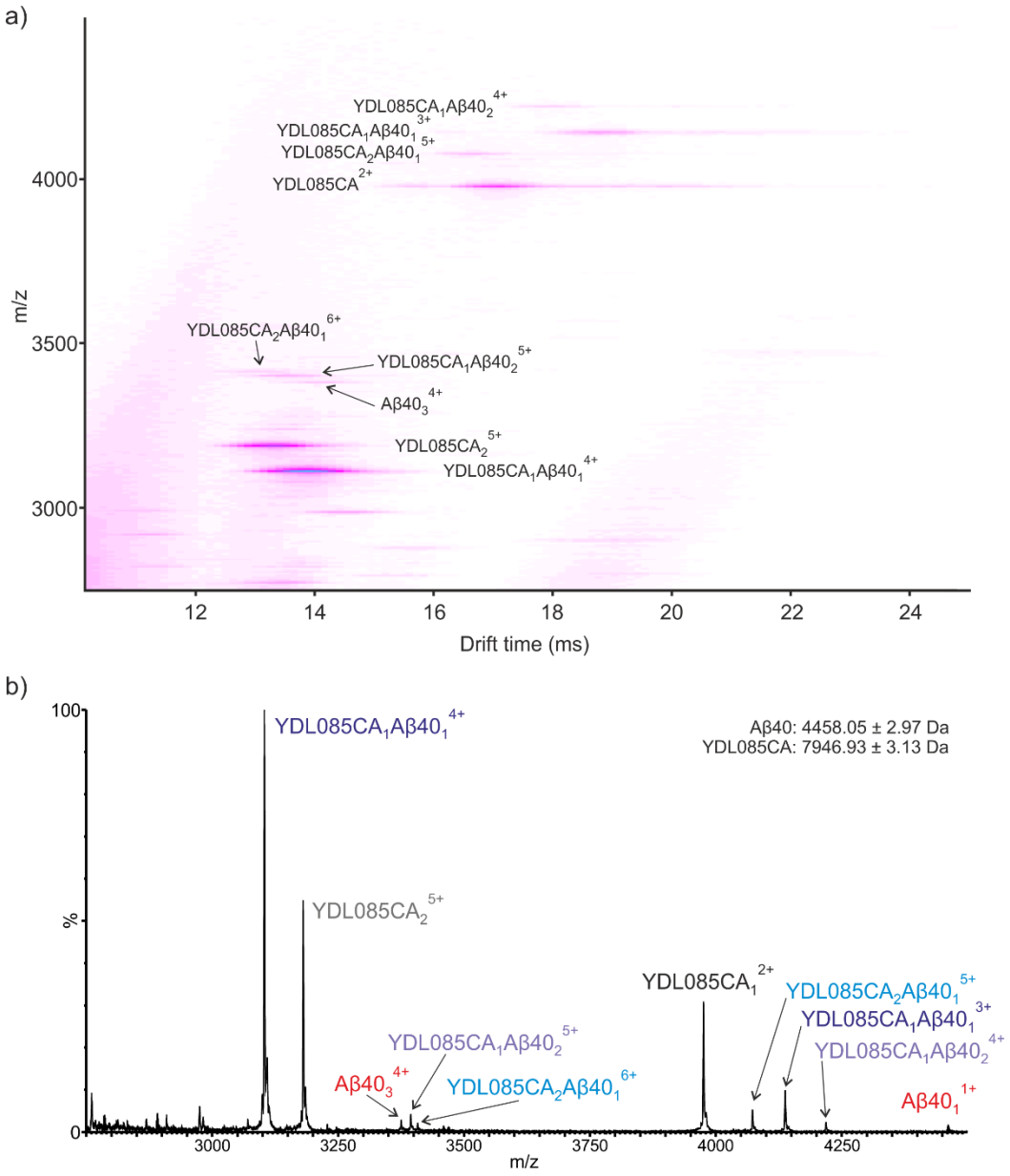


Figure 3.15: Native nESI-IMS-MS of a mixture Aβ40 and YDL085CA reveals a population of complexes. IMS drift time plot (a) and MS spectrum (b) of 20 µM Aβ40 and 20 µM YDL085CA in 50 mM ammonium bicarbonate, pH 7.4. In both cases the m/z range from 2750 to 4500 is shown as this region contains the complex ions. The peaks are labelled according to the ion or complex to which they correspond with the charge state indicated. In addition to monomeric Aβ40 and YDL085CA, a dimer of YDL085CA, a trimer of Aβ40 and a range of complexes between the two proteins (1:1, 1:2 and 2:1) are observed. The drift time plot (a) uses a square root intensity scale. The theoretical masses (average mass/monoisotopic mass) for Aβ40 and YDL085CA are 4461.06/4458.19 Da and 7949.20/7944.32 Da respectively.

Native nESI-MS of α-synuclein revealed a possible interaction between YDL085CA and some charge states of α-synuclein (Figure 3.16b). Ions corresponding to the 1:1 complex are observed for the M+8H^{8+} to M+19H^{19+} ions of α-synuclein. The intensity of these ions is relatively low compared to the ions of the non-interacting proteins. IMS of the complex ions shows that for the most intense complex ion that could easily be isolated from other peaks, M+10H^{10+}, there are three confirmations while for the M+10H^{10+} ion of free α-synuclein (Figure 3.17) only two conformations are observed. This is consistent with previously published evidence that α-synuclein becomes less compact upon binding to MOAG-4 (Yoshimura *et al.* 2017).

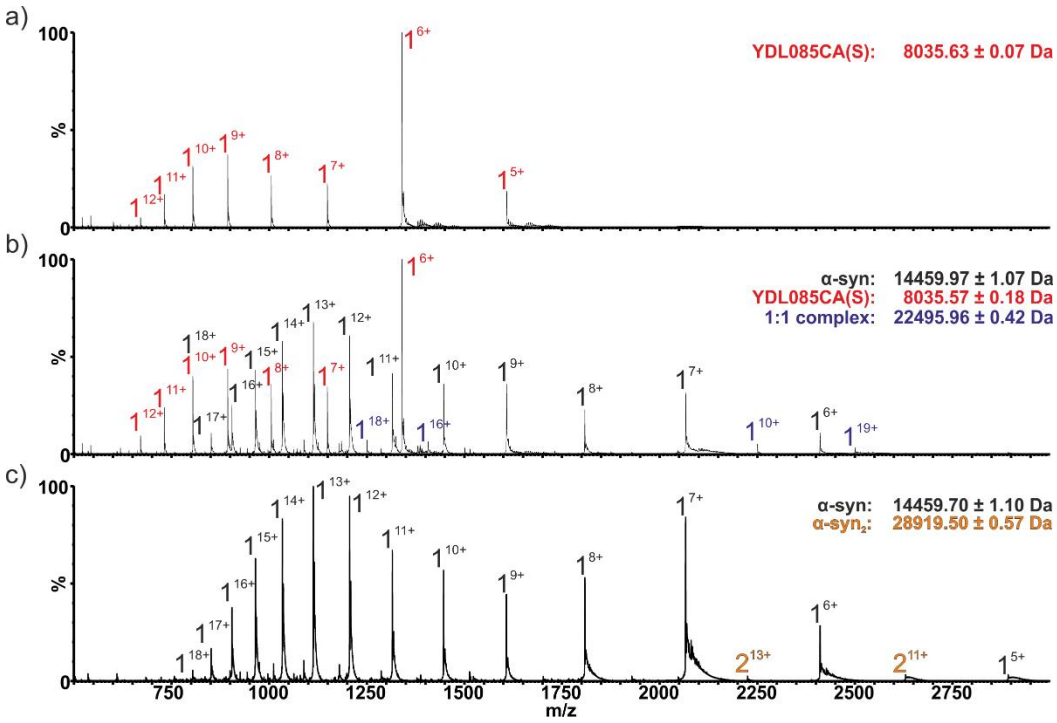


Figure 3.16: Native MS reveals a 1:1 complex formed by α-synuclein and YDL085CA(S). Native MS spectra for (a) YDL085CA(S) alone, (b) a 1:1 mixture of YDL085CA(S) and α-synuclein and (c) α-synuclein alone. The masses of the observed proteins and complexes are indicated. The theoretical average masses of each construct are: YDL085CA(S) 8036.28 Da and α-synuclein 14460.18 Da. YL085CA(S) and α-synuclein were 50 µM and 10 µM respectively in 50 mM ammonium acetate, pH 7.6.

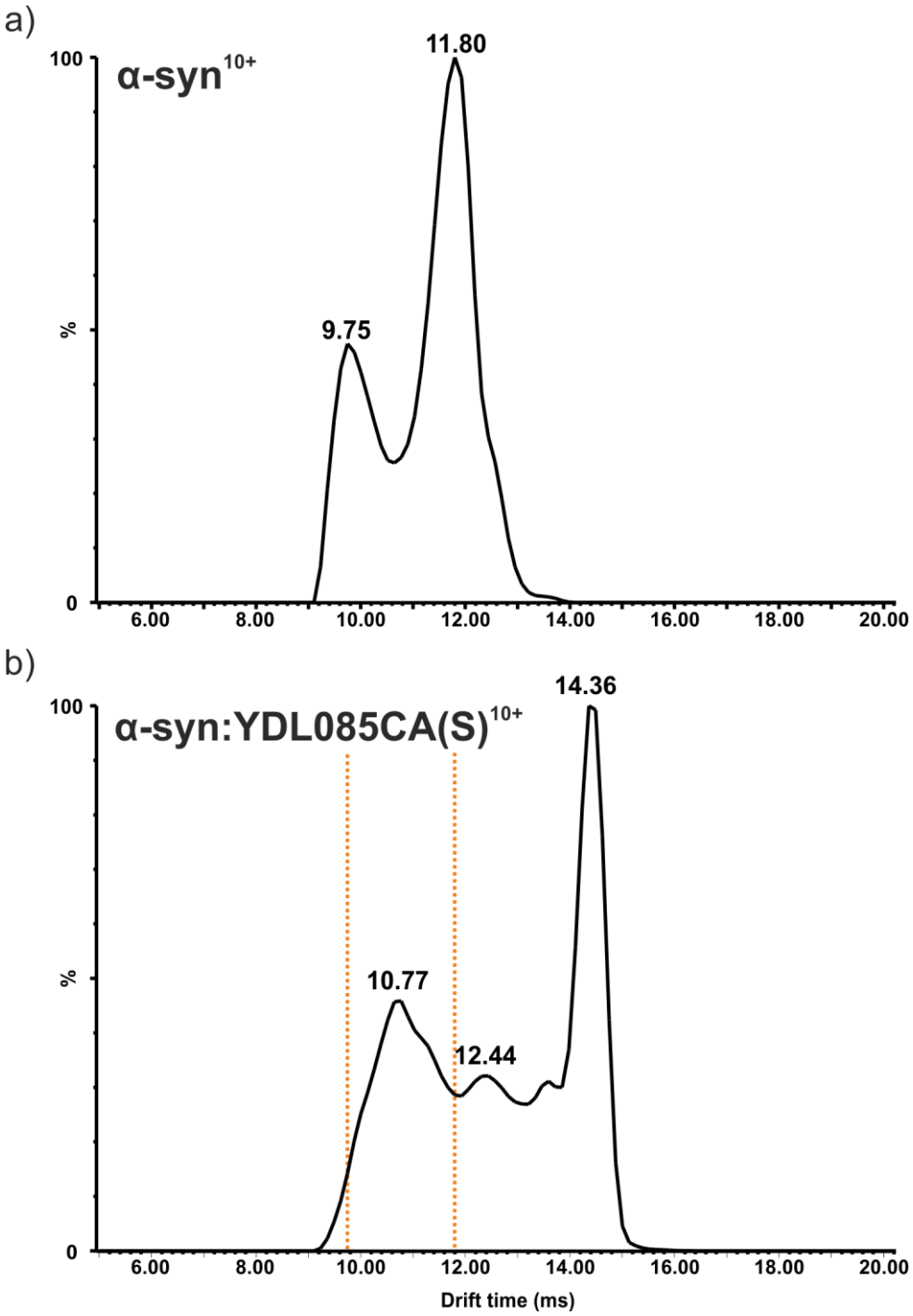


Figure 3.17: nESI-IMS-MS arrival time distributions for the M+10H^{10+} ion of (a) α-synuclein alone and (b) the α-synuclein:YDL085CA(S) complex. The dotted lines in (b) indicate the position of maxima in the absence of YDL085CA. The drift time (in ms) is indicated above major peaks.

3.6 Chemical crosslinking studies of the YDL085CA:Aβ40 complex

Native nESI-MS can provide a great deal of information regarding the interaction of proteins and peptides (stoichiometries, relative abundances and conformational information to name three examples). However, native MS does not inherently provide information regarding the site of an interaction nor is it always practical to investigate an intact complex via native methods. The YDL085CA:Aβ40 complex was observed by native MS only at very low levels (Figure 3.15) and as such further investigation of the site of interaction by techniques such as ETD would be difficult. Therefore the investigation of this complex was continued by means of crosslinking MS.

Chemical crosslinking of an equimolar ratio of YDL085CA and Aβ40 with BS^3 reveals several interesting effects (Figure 3.18). Under the conditions employed here Aβ40 forms no crosslinked species that can be detected by SDS-PAGE at any of the concentrations of BS^3 used. The crosslinking reactions was performed at room temperature for one hour, a time point that, under these conditions, is very early in the lag phase of fibril formation. It has been suggested that at early points in Aβ aggregation the population is dominated at low concentrations by monomer, with a reported apparent equilibrium constant of 54 µM (Garai and Frieden 2013). As such, the levels of oligomeric species may be extremely low and thus below the limit of detection of this experiment. However, at the concentrations used in the crosslinking experiments, oligomers have been observed (see Figure 3.15 and the literature (Bernstein *et al.* 2009b)). A second possibility is that the reactive residues in Aβ40 (K16, K28 and M0) may be distal from each another in the oligomeric conformation and thus crosslinking would be unlikely to occur. However, the broad spread of reactive residues throughout the N-terminus and central region of Aβ40 sequence makes this seem unlikely. Libeu at al. reported observing dimeric and trimeric Aβ40 under similar conditions (10 µM Aβ40 in PBS incubated at 37 °C for 30 min); however, they enriched oligomeric species by the use of size exclusion chromatography and Western blot detection, the latter of which can be more sensitive than the Coomassie stain used here (Libeu *et al.* 2011).

At the lowest concentration of BS^3, YDL085CA appears only as monomer although the intensity of the band observed is lower than for other samples suggesting that the recovery of this sample may have been poor. At higher excess of BS^3 a ladder of oligomeric species of YDL085CA is observed with approximate masses corresponding to as many as six monomer subunits. Conversely in the native MS experiments YDL085CA was only observed as monomer and possibly dimer in the presence of Aβ40. Whether this pattern of crosslinked species represents a real distribution of oligomers or is indicative of transient interactions or collisions which are falsely enriched due to the presence of large numbers of reactive residues is unknown.

YDL085CA and Aβ40 together show bands characteristic of both proteins alone (i.e. monomers and the YDL085CA ladder are still observed). A band is observed between 10 kDa and 15 kDa which is unique to the mixed reaction and approximately the correct mass for a 1:1 complex (indicated by the arrowhead in Figure 3.18). The intensity of this band increases as the amount of BS^3 is raised. A possible second band is observable between 20 kDa and 25 kDa and slightly below the YDL085CA trimer. This band could correspond to a number of theoretical species (for example an approximate molecular weight of 24 kDa could correspond to a YDL085CA trimer, a $YDL085CA_1Aβ40_3$ complex or a $YDL085CA_2Aβ40_2$ complex or a mixture of these species) and due to the low intensity it was not possible to extract this band for further analysis.

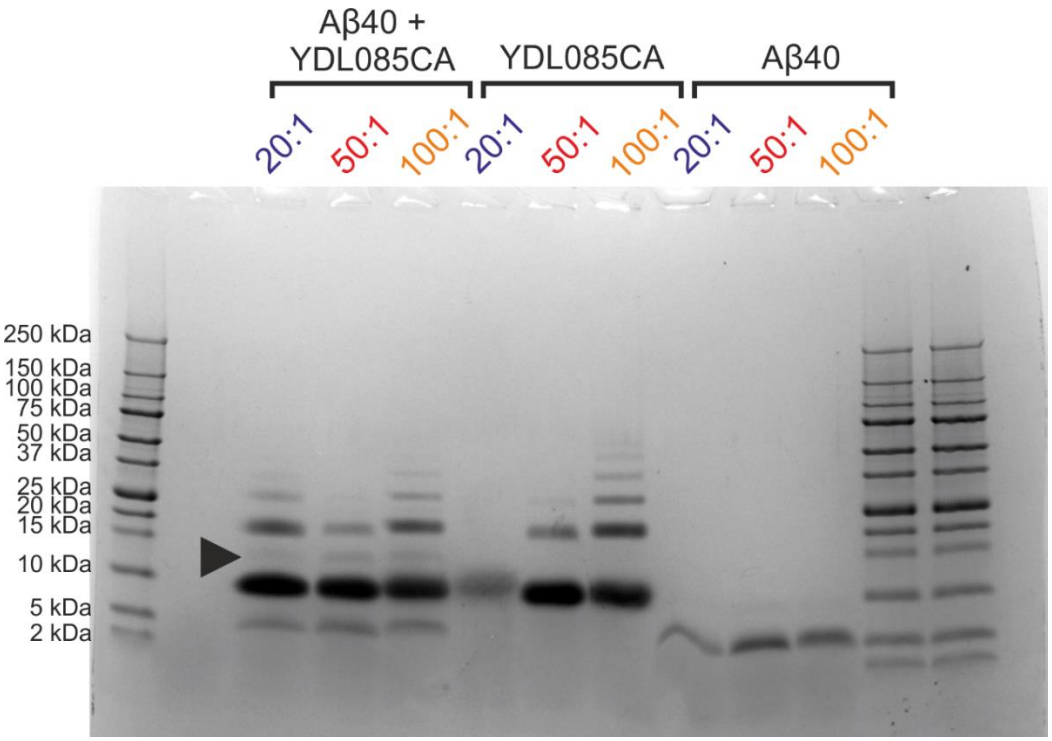


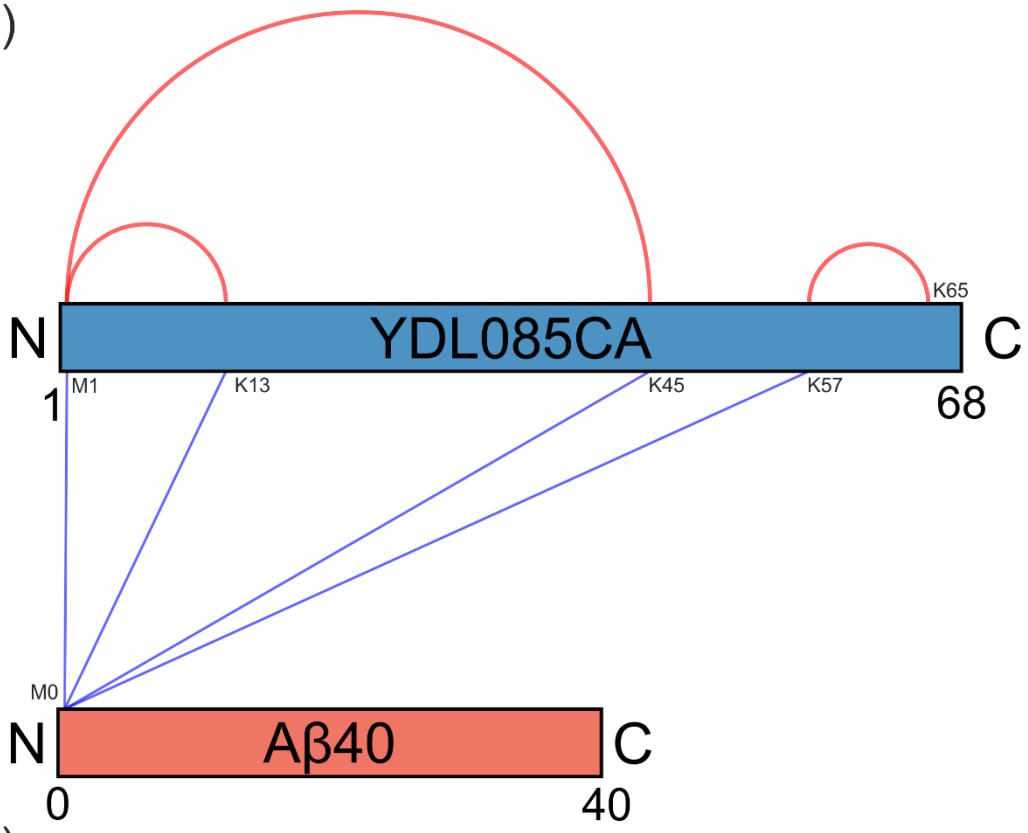
Figure 3.18: SDS-PAGE of Aβ40 and YDL085CA crosslinking at different concentrations of BS^3 after one hour of incubation. The contents of each lane are indicated above the figure. Aβ40 alone does not result in any observable crosslinked complexes at any of the BS^3 concentrations used. YDL085CA alone results in the formation of a ladder of crosslinked species. Co-incubation of Aβ40 and YDL085CA results in the formation of a 1:1 complex (indicated by the black arrowhead). In all cases it was observed that the top of the wells did not stain indicating that large SDS-stable or crosslinked fibrillar species were not generated during the experiment.

The bands for the monomeric species, the YDL085CA dimer and the putative 1:1 complex were excised from the gel and treated as described in the Chapter 2. Briefly, the excised gel bands were destained in 30% ethanol followed by sequential washes in acetonitrile to dehydrate the gel pieces. The dehydrated samples were digested overnight (18 hours) with trypsin. The resulting peptides were recovered by a series of washes in acetonitrile-formic acid solutions and the solution was evaporated to concentrate the peptides. While a tryptic digest of YDL085CA results in many short peptides, the presence of covalent BS^3 modification blocks cleavage resulting in longer peptides with missed cleavages. These longer peptides can be analysed more easily than the short peptides produced from unmodified protein.

The recovered peptides were analysed by LC-MS/MS. Samples were introduced into a M-class ACQUITY UPLC (equipped with a Symmetry C18 trap column and a Waters HSS T3 C18 75µM internal diameter x 150 mm column) attached to a Xevo G2XS or Synapt G2Si MS instrument. The peptides were eluted using a gradient of acetonitrile and infused into the MS by nESI.

All of the inter-protein crosslinks observed are generated by the reaction of the N-terminus of Aβ40 (i.e. the N-terminal primary amine of the methionine residue) and a variety of sites on YDL085CA (Figure 3.19). While there are two other primary amines in Aβ40 (lysine residues 16 and 28) neither is seen to crosslink to YDL085CA. Lysine 16 is observed to form dead end crosslinks by reacting with water from the buffer. This suggests that the Aβ40:YDL085CA reaction is mediated by the N-terminal segment of Aβ40 with little or no involvement from the central region. The crosslinking distance of BS^3 is 11.4 Å, therefore it can be suggested that no crosslinkable residues of YDL085CA are within 11.4 Å of either lysine residues of Aβ40 when the complex is formed.

a)



b)

Aβ40	YDL085CA	m/z	Charge	Observed Mass	Calculated Mass	Deviation (ppm)	FDR	Score
0-6	0-3	649.80	2	1298.59	1298.59	0.89	<0.005	29
0-6	0-3	649.80	2	1298.59	1298.59	0.89	<0.005	29
0-6	12-16	512.60	3	1535.79	1535.79	2.02	<0.005	32
0-6	44-52	621.97	3	1863.90	1863.90	-0.82	<0.005	57
0-6	44-52	627.30	3	1879.90	1879.90	-0.55	0.01	18
0-6	54-60	584.62	3	1751.85	1751.85	-1.46	<0.005	66

Figure 3.19: BS^3 crosslinking map for the 1:1 complex of Aβ40 and YDL085CA. (a) The inter- (blue) and intra- (red) protein crosslinks mapped onto the sequence of Aβ40 and YDL085CA. (b) A table showing the inter-protein crosslinks with false discovery rate (FDR) below 5% as calculated in the StavroX software (Gotze *et al.* 2012). Note that all inter-protein crosslinks involved the N-terminus of Aβ40. Figure (a) prepared using xVis (Grimm *et al.* 2015).

The N-terminus of YDL08CA crosslinks to Aβ40, as do peptides corresponding to residues 12-16, 44-52 and 54-60 of YDL085CA. Aside from the N-terminus all of these residues fall in conserved regions of YDL085CA that are predicted to form helical structure (Figure 3.1). In MOAG-4 both of these sections form helices and the first helix, containing YDL085CA 12-16 has been shown to undergo a large change in structure upon binding to a substrate (Yoshimura *et al.* 2017).

A number of intra-protein crosslinks are also observed for YDL085CA. Intra-protein crosslinks, in particular those between defined domains, can be used to infer information about a protein's tertiary structure. However, in this case these crosslinks are scarce and as such little reliable information may be gained from them. The scarcity of data for this experiment is in part due to the inherent difficulties of correctly assigning which peaks in an MS/MS spectrum are from crosslinked samples. This issue can be resolved unequivocally and thus the number of identified crosslinked peptides increased, by use of isotopically labelled crosslinking reagents (Figure 3.20a and b) (Collins *et al.* 2003; Muller *et al.* 2001; Pearson, Pannell and Fales 2002).

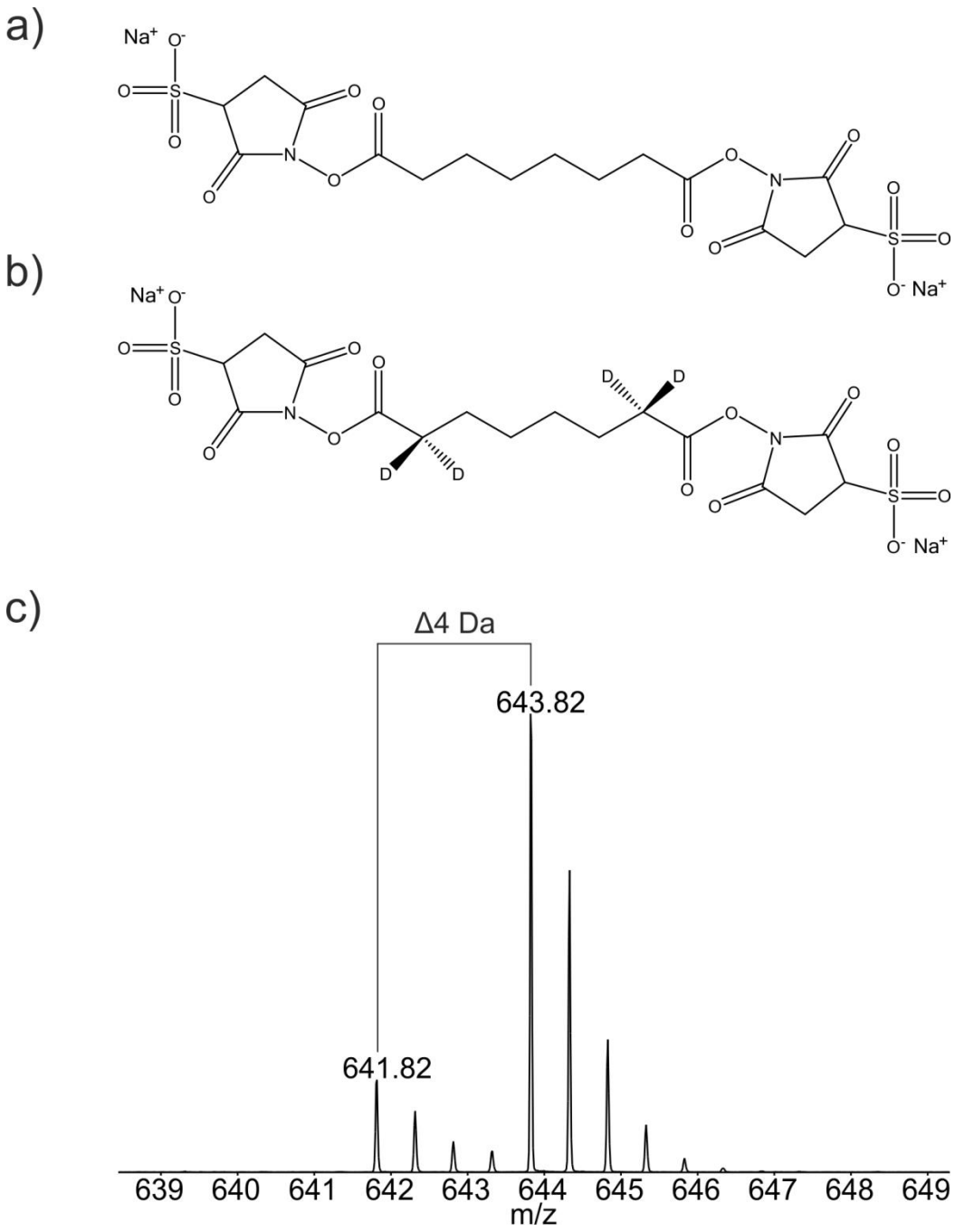


Figure 3.20: Crosslinking with isotopically labelled BS^3. (a) The structure of BS^3 and (b) the structure of BS^3 H/D_4 indicating the position of the stably incorporated deuterium atoms. (c) A doublet of peaks is observed for a crosslink formed in the presence of mixed BS^3 D_4 and BS^3 D_0 with a signature 4 Da mass difference (2 m/z for a $M+2H^{2+}$ ion as presented above).

In a system where deuterated and undeuterated crosslinkers are used in tandem the relative inclusion of the heavy and light isotopic variant should be reflective of the ratio of the two crosslinkers in solution (Figure 3.20c). For example, if the heavy and light crosslinker are mixed at a stoichiometric ratio the resulting

population crosslinked peptides should have a 50:50 ratio of peptides generated by heavy and light crosslinker interactions.

The benefit of using a mixed isotopically labelled crosslinker becomes clear when the MS of a putative crosslinked peptide is viewed. For BS^3 D_0/D_4, a doublet peak distribution separated by four Daltons is observed for the crosslinked species (Figure 3.20c). These characteristic doublet peaks allow for rapid removal of peptides that are generated from the digestion of non-crosslinked regions of a protein (Collins *et al.* 2003; Muller *et al.* 2001; Pearson, Pannell and Fales 2002). The presence of doublet peaks is of use both in computational analysis of data sets (which can be very large for the MS/MS data in a crosslinking MS experiment) and for manual validation of the results.

The crosslinking of Aβ40 and YDL085CA with BS^3 D_0/D_4 was performed at a 50:1 excess of crosslinker (Figure 3.21). The initial experiments with the undeuterated BS^3 revealed that while a 20:1 excess of BS^3 generated very low levels of crosslinked peptide, a 100:1 excess led to very poor recovery of peptides post digest.

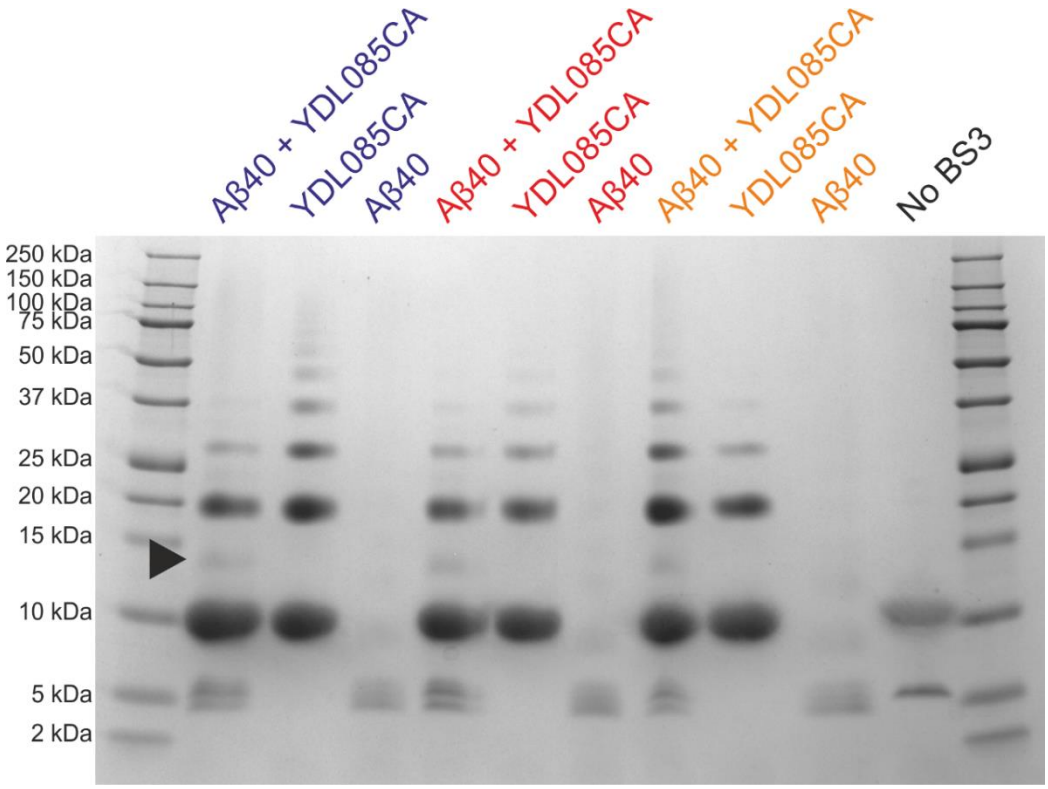


Figure 3.21: SDS-PAGE of three repeats of the Aβ40 and YDL085CA crosslinking experiment using a BS^3 D_0/D_4 mixture. The black arrowhead indicates the position of the 1:1 complex. A control well containing Aβ40 and YDL085CA without BS^3 is shown at the right hand side. The cross linking reaction was performed in 20 mM sodium phosphate, 0.2 mM EDTA, 0.02% sodium azide, pH 7.4 with 25 µM Aβ40 and 25 µM YDL085CA in the presence of a 50:1 excess of crosslinker (w/w) for 1 hour at room temperature.

As observed for the undeuterated BS^3 crosslinking experiments, Aβ40 appears to remain monomeric under the solution and crosslinking conditions employed here. YDL085CA once again forms a ladder of higher order structures when exposed to BS^3 (species up to pentamer are resolved with some indication of higher order species). When crosslinked, the mixed Aβ40:YDL085CA samples form a complex with the approximate molecular weight of a 1:1 complex (12 kDa, indicated in Figure 3.21 by the black arrowhead). It is also interesting to note that when Aβ40 without crosslinker is compared to the crosslinked Aβ40 samples the bands appear different. The crosslinked Aβ40 samples show evidence of two bands with different apparent molecular weight (slightly above and below the 5 kDa molecular weight marker). This may be due either to changes in the diffusion characteristics of the protein due to incorporation of crosslinker directly, or to

changes in the conformation of the crosslinked species due to the formation of intra molecular crosslinks which restrict the unfolding of the protein in SDS buffer. Due to the low intensity of these bands and the poor resolution between them they were extracted as a single sample and treated as such through the downstream processing.

The Aβ40:YDL085CA complex crosslinked with BS^3 D_0/D_4 results in a larger number of identified crosslinked peptides than for BS^3 D_0 alone. An example of this is shown in Figure 3.22 (the full list of identified crosslinks is shown in Appendix II). Once again, only the N-terminal primary amine group of Aβ40 forms any crosslinks with YDL085CA (Figure 3.22). Interestingly, no crosslinks are identified from either of the lysine residues in Aβ40 to YDL085CA, Aβ40 or dead end crosslinks. This may suggest that these sites are occluded in the complex and thus not accessible to the crosslinker.

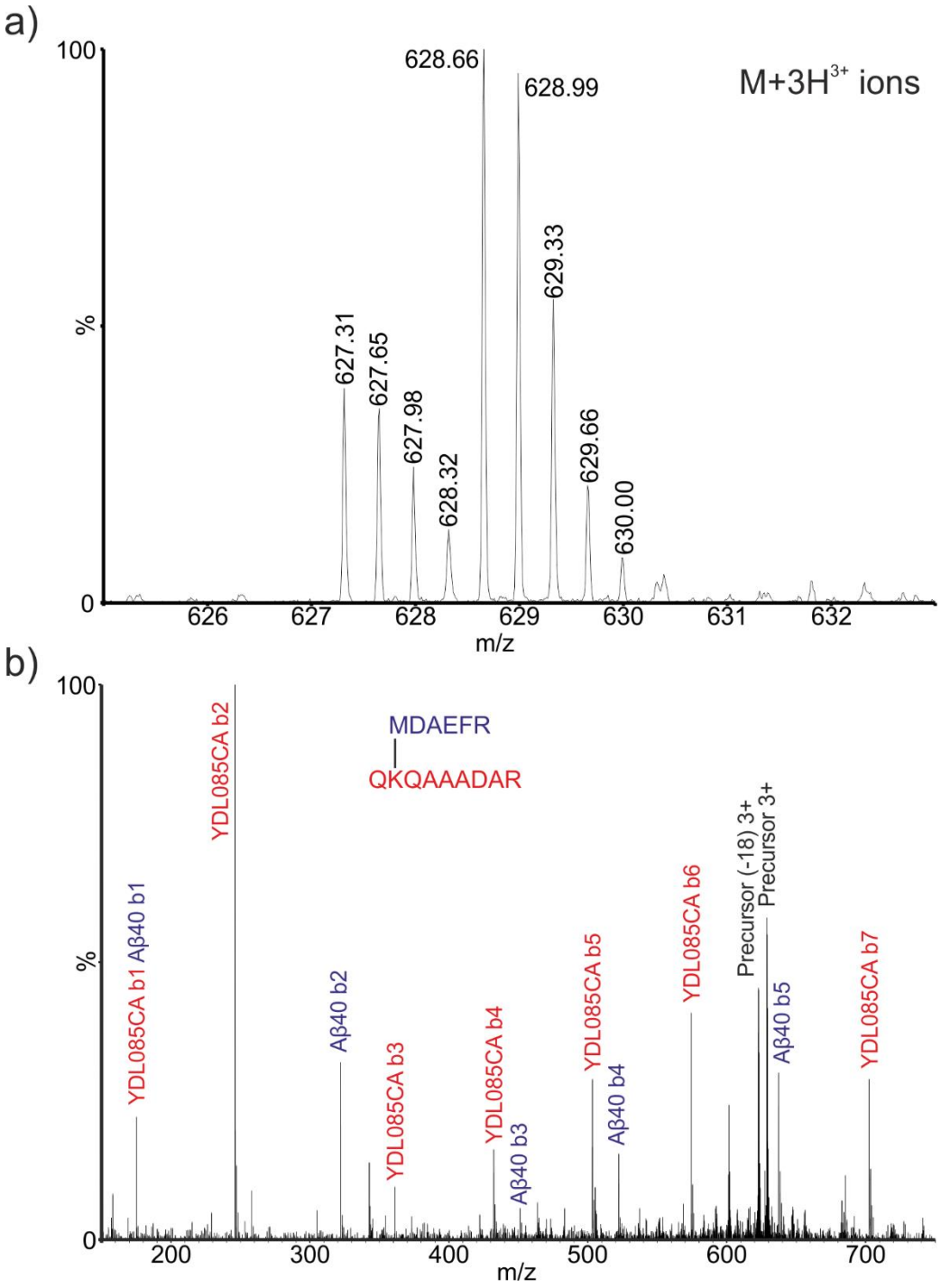


Figure 3.22: Example of data analysis for a crosslinked peptide from a YDL085CA:Aβ40 complex. (a) A doublet peak was identified at 627.3 and 628.6 m/z (corresponding to 1878.9 Da and 1883.0 Da, respectively). The ion corresponding to the heavy isotopically labelled BS^3 ion was isolated in the MS/MS spectrum (b). Fragments corresponding to peptide from YDL085CA and Aβ40 are shown in red and blue respectively (the full mass range is not shown). This peptide corresponds to residues 44-52 of YDL085CA and residues 0-5 of Aβ40.

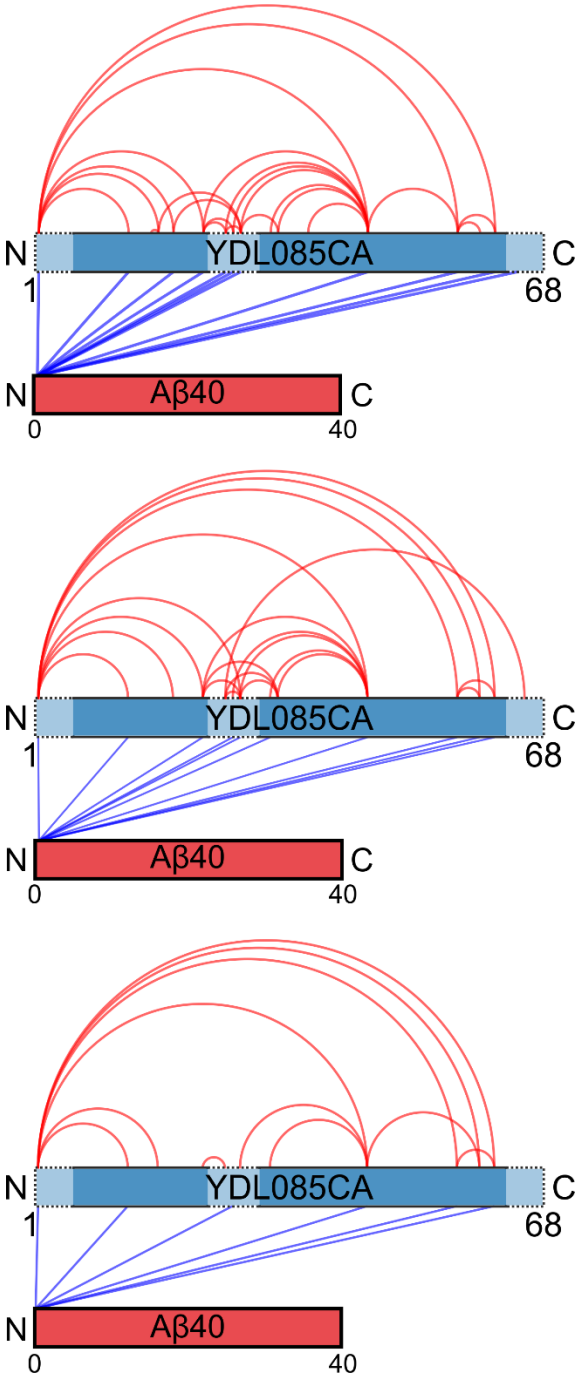


Figure 3.23: A crosslinking map of the 1:1 complex of Aβ40 and YDL085CA. Three repeats are shown. Inter-protein crosslinks are shown in blue and intra-protein crosslinks are indicated in red. A table of the crosslinks identified is shown in Appendix II. Crosslinks were identified using the StavroX software (Gotze *et al.* 2012). Figure prepared using xVis (Grimm *et al.* 2015). For YDL085CA, solid boxes represent predicted α-helical secondary structure. The crosslinking reaction was performed in 20 mM sodium phosphate, 0.2 mM EDTA, 0.02% sodium azide, pH 7.4 with 25 µM Aβ40 and 25 µM YDL085CA in the presence of a 50:1 excess of crosslinker (w/w) for 1 hour at room temperature.

The sites of the crosslinks between YDL085CA and Aβ40 can be localised to several regions of the proteins (Figure 3.24 and Appendix II). Once again, the N-terminus of YDL085CA interacts with Aβ40. The central region of the proposed transient α-helix (residues 8-22) also crosslinks to Aβ40 via the lysine in position 13 though not those at 16 and 17. Lysine 19 crosslinking was observed in only one data set where it forms an intermolecular crosslink to Aβ40. Lysines 23 and 27 crosslink to Aβ40 as does Serine 28 and lie respectively at the end of a proposed α-helix and in a region proposed to lack secondary structure. Lysines 32 and 33 are not observed to crosslink to Aβ40 suggesting that this region of the unstructured inter helix linker or N-terminal end of the second helix does not form a stable interaction with Aβ40. Residues 45, 57 and 62 of the second proposed α-helix also crosslink to Aβ40 while the nearby lysine 60 does not. A helical wheel projection of this helix (Figure 3.24) places residues which crosslink to the substrate separated by large distances (140 ° for K57 and K62) and does not present and obvious interface for the interaction to occur on. The NMR structure of MOAG-4 suggests the second helix may be exposed on all sides (Yoshimura *et al.* 2017) thus allowing a bound protein to wrap around the interface.

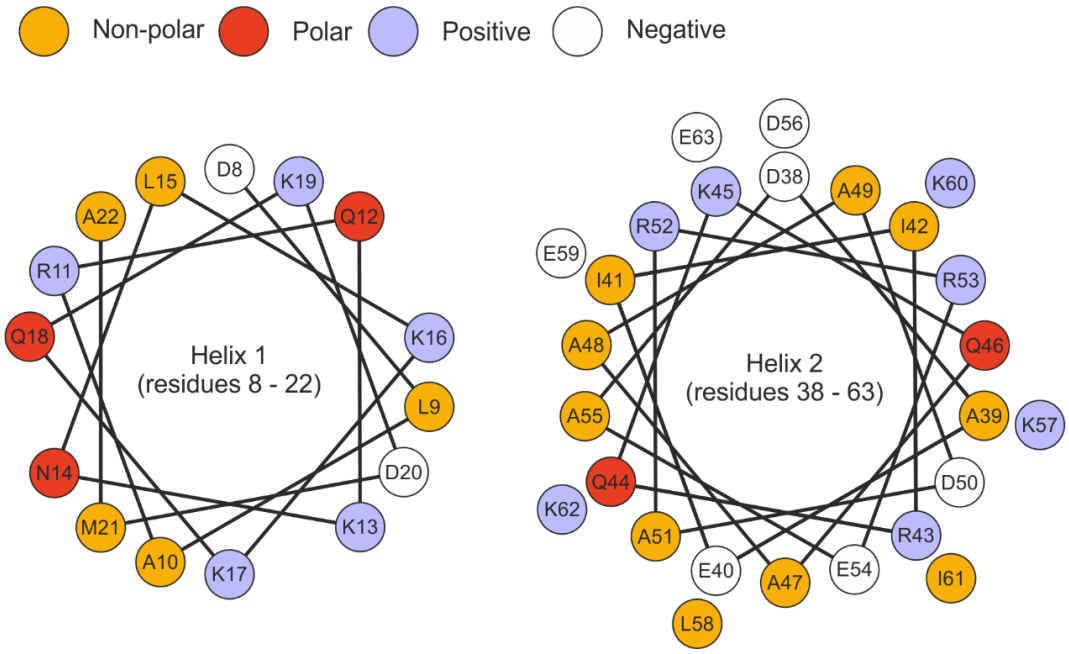


Figure 3.24: Helical wheel projections of helix 1 (residues 8-22) and helix 2 (residues 38-63) of YDL085CA. Sequences processed using the Wheel tool from Rapheal Zidovetzki (Zidovetzki 2009).

These interactions point to both the importance of helix one and to the previously unreported importance of helix two and part of the linking region of YDL085CA in the binding interface with Aβ40 (Figure 3.25). It is interesting to note that the helix two crosslinks, in particular K57 and 62, lie in regions that are not conserved between MOAG-4 and YDL085CA. This may indicate that YDL085CA has evolved a different interaction domain to MOAG-4.

Figure 3.25: The sequence of YDL085CA indicating the position of reactive residues. Reactive lysine residues are indicated in black boxes. Solid arrowheads above the residues indicate that an inter-protein crosslink was observed while an open arrowhead indicates that no inter-protein crosslinks were observed. Shown below are the secondary structure of MOAG-4 (brown) (Yoshimura *et al.* 2017) and the predicted secondary structure of YDL085CA (blue) repeated from Figure 3.1. The boxes indicate helical structure.

It is interesting to note that, in the absence of YDL085CA, Aβ40 forms several intramolecular crosslinks while none are observed in the presence of YDL085CA (Figure 3.23 and 3.26). Crosslinks are observed between several residues in the

N-terminus as well as a group of amino acids between residues 16 and 29. In addition, the N-terminus appears to interact with the region surrounding residue 16, a region which has been suggested to have helical secondary structure (Vivekanandan *et al.* 2011). The apparent abrogation of these interactions upon binding to YDL085CA suggests that Aβ40 may enter a more expanded conformation. Interaction with MOAG-4 has been shown previously to cause the sampling of an expanded conformation of α-synuclein (Yoshimura *et al.* 2017). This expanded conformation may expose otherwise protected amyloidogenic residues or alter the dynamics of the Aβ40 structural ensemble.

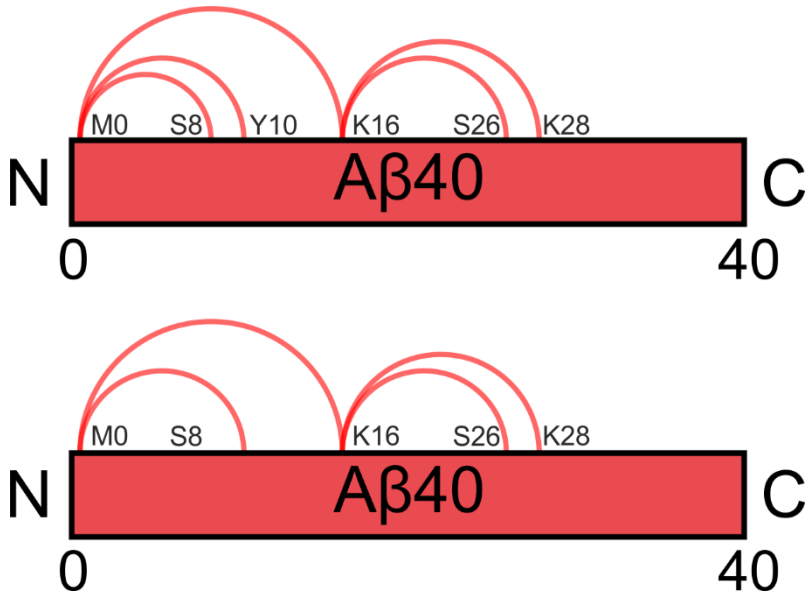


Figure 3.26: A crosslinking map of Aβ40 monomer. Three repeats are shown. Intra-protein crosslinks are indicated in red. Note that in the third repeat no crosslinks could be identified. The crosslinking reaction was performed in 20 mM sodium phosphate, 0.2 mM EDTA, 0.02% sodium azide, pH 7.4 with 25 µM Aβ40 in the presence of a 50:1 excess of crosslinker (w/w) for 1 hour at room temperature.

Several of the residues crosslinked in the Aβ40 samples did not contain primary amine groups (S8, Y11, and S26) and as such are not traditionally held to be reactive residues in BS^3 crosslinking. However, as discussed in Chapter 3.1.2, there is significant evidence for BS^3 reacting with hydroxyl groups in the residues serine and tyrosine as observed here as well as threonine (Madler *et al.* 2009).

Crosslinking of the YDL085CA monomer reveals a complex network of crosslinks between many residues of the protein (Figure 3.27). Given the evidence from structural predictions, NMR data and CD analysis (Figure 3.28) that YDL085CA is extremely flexible and may lack structure in large areas (approximately 40% of the protein is predicted to lack structure in the NMR experiments on MOAG-4 while the secondary structure predictions and CD analysis of YDL085CA suggests approximately 20% of the protein is unstructured) the wide range of long distance crosslinks is perhaps unsurprising (Yoshimura *et al.* 2017).

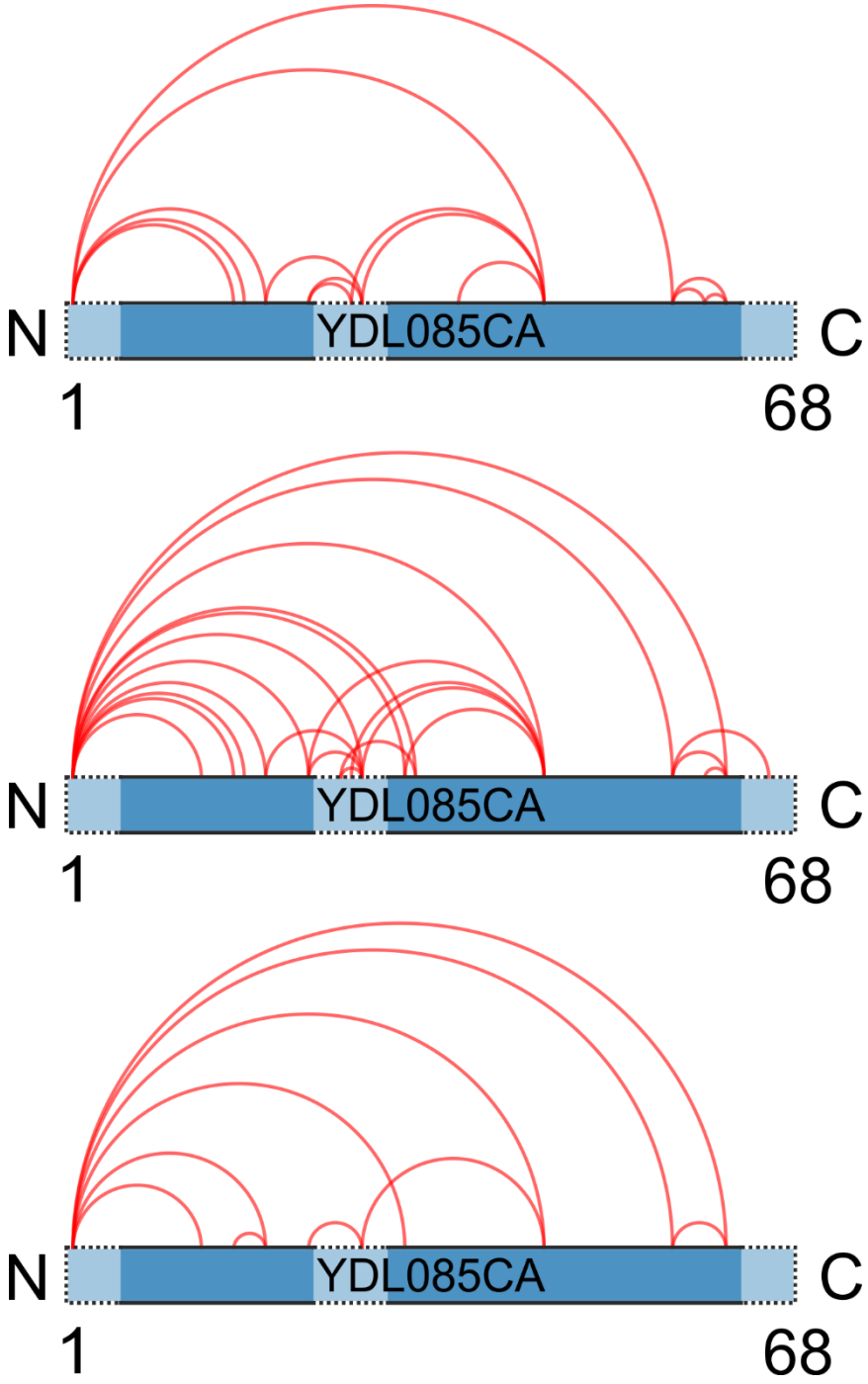
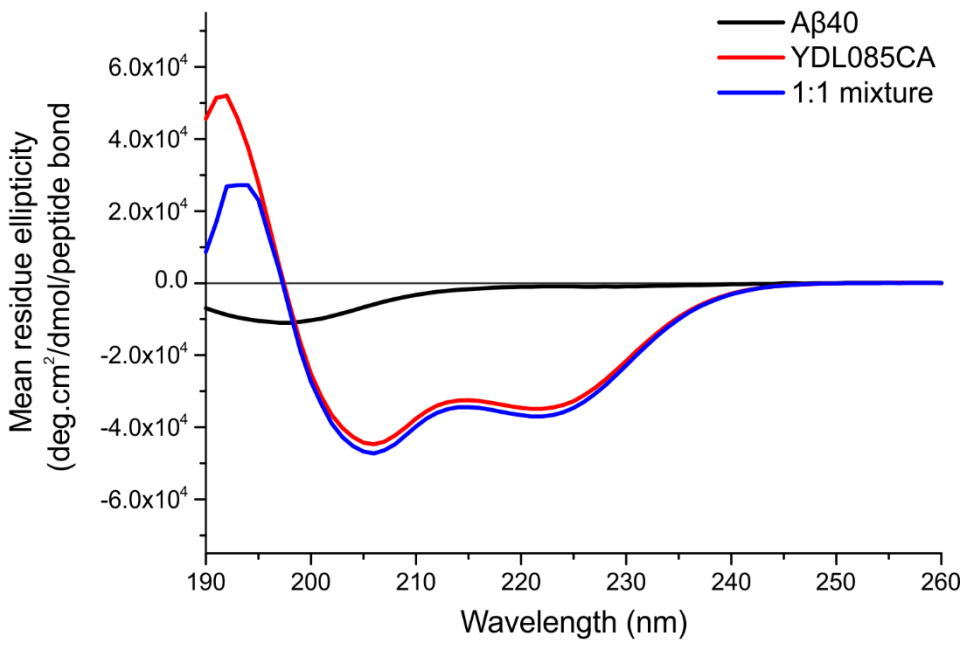


Figure 3.27: A crosslinking map of YDL085CA monomer. Three repeats are shown. Intra-protein crosslinks are indicated in red. The position of the crosslinks suggests that the N-terminus can interact with sites throughout the protein while the C-terminus does not interact with the central region. Solid boxes represent predicted α-helical secondary structure. The crosslinking reaction was performed in 20 mM sodium phosphate, 0.2 mM EDTA, 0.02% sodium azide, pH 7.4 with 25 µM YDL085CA in the presence of a 50:1 excess of crosslinker (w/w) for 1 hour at room temperature.

a)



b)

	α-helix	β-strand	Other
Aβ40	0.03	0.24	0.67
YDL085CA	0.73	0.08	0.18
YDL085CA from 1:1 mixture	0.77	0.06	0.16

Figure 3.28: Far UV CD analysis of YDL085CA and Aβ40. (a) The CD spectra over the range 190 nm to 260 nm of Aβ40 alone (black), YDL085CA alone (red) and YDL085CA in the 1:1 mixture (blue). (b) CDSSTR analysis of the secondary structure content of the proteins in (a). The proteins in (a) were 22 µM in 20 mM sodium phosphate, pH 7.4 at 20 °C. The signal from YDL085CA in the 1:1 mixture was calculated by subtracting the signal for Aβ40 alone from the total signal.

Essentially all crosslinkable residues (M1, K13, K16, K17, K19, K23, K26, K27, S28, K32, K33, S37, K45, K57, K60, and K62) in YDL085CA from intra molecular crosslinks. The C-terminal reactive residue, K65, was observed to form crosslinks in only one sample while T66 was not observed to form any crosslinks. Interestingly no dead end crosslinks were observed for any residue suggesting that YDL085CA samples a conformation(s) in which all reactive residues are within 11.4 Å of a second reactive residues.

The N-terminus of YDL085CA was observed to interact with many residues situated throughout the sequence of YDL085CA. The region consisting of residues

15-35 contains many reactive residues (9 lysine residues and 1 serine residue). These residues were observed to form crosslinks with each other and with the N-terminus but not with the reactive residues in the C-terminal region, such as K57, K60, K62, K65 and T66. This suggests that, while the N-terminus is free to interact with many residues, something prevents residues 15-35 and the C-terminal region from coming into close proximity. The predicted secondary structure of YDL085CA (and observed NMR data for MOAG-4 (Yoshimura *et al.* 2017)) shows that residues 15-35 lie at the end of the transient helix one and either in the helix one-helix two linker or the N-terminal end of helix two (Figure 3.1b-d). The C-terminal reactive residues meanwhile lie in regions thought to form the C-terminal end of the second helix and in the short sequence immediately after the helix. Therefore, the constraints of the helices' flexibility likely prevent interactions between these two groups of residues. Somewhat unusually, K45 crosslinks to residues in the 15-35 cluster suggesting that the region consisting of residues 15-35 can come into close contact with the middle of helix two. This may be explained by the first helix forming transiently and thus sampling a conformation(s) in which residues 15-35 have considerable freedom to interact with helix two.

There is little change in the crosslinking pattern of YDL085CA upon the addition of Aβ40 (Figure 3.23 and 3.27). This may be suggestive of there being a low degree of conformational rearrangement of YDL085CA upon binding to its substrate. Similarly, only small differences in the structure of MOAG-4 were observed upon its interaction with α-synuclein (Yoshimura *et al.* 2017). CD analysis of YDL085CA (Figure 3.28) suggests that the addition of Aβ40 causes little change in the secondary structure of YDL085CA.

3.7 Discussion

In this Chapter it has been shown that the MOAG-4/SERF1 ortholog, YDL085CA, has a very different effect on Aβ40 aggregation than does SERF1. In addition, a crosslinking-MS approach was employed and used successfully to map the sites of interaction between Aβ40 and YDL085CA to the N-terminus of Aβ40 and helices of YDL085CA, respectively.

Investigation of the effect of YDL085CA on the aggregation kinetics of Aβ40 resulted in several interesting observations. The first observation is that the effect of YDL085CA is highly dependent upon the buffer. As discussed above, it is thought that the conformations of highly disordered proteins, such as Aβ40 and YDL085CA, are extremely sensitive to buffer conditions, namely the ions present in solution (Muller-Spath *et al.* 2010; Liao *et al.* 2013; Wicky, Shammas and Clarke 2017). While the ionic strengths of the two buffers employed in this study are similar, the ions present are different consisting primarily of Na^{2+} and PO_4^{3-} ions in the case of the sodium phosphate buffer and NH_4^+ and HCO_3^- in the case of the ammonium bicarbonate buffer. The conformation of a protein, even an IDP, can play an important role in its behaviour and as such the effect of the buffer ions on the conformation of the proteins may account for this variation (Hofmann, Nettels and Schuler 2013; Muller-Spath *et al.* 2010; Wicky, Shammas and Clarke 2017).

An additional consideration in the interaction of YDL085CA and Aβ40 is the abundance of charges in solution. The highly charged sequence of YDL085CA (which has a net charge of 12+ at pH 7.4 and pI 11.1) and the evidence, from the crosslinking experiments, that the interaction between the two proteins is mediated by the charged N-terminal region of Aβ40 suggest that the interaction may be electrostatically driven. Previously published work on MOAG-4 has also demonstrated that an electrostatic interaction is likely (Falsone *et al.* 2012; Yoshimura *et al.* 2017). As such the presence of ions which can shield charged sites in solution could alter the effect of YDL085CA on the aggregation process.

A further factor, not explored in this investigation, is the potential effect of pH upon the YDL085CA:Aβ40 interaction.

The second surprising observation in the Aβ40 aggregation assays was that YDL085CA behaves very differently to SERF1 under similar conditions (Falsone *et al.* 2012). While SERF1 has been shown to reduce the lag time of Aβ40 aggregation by approximately 50% at an equimolar ratio (Falsone *et al.* 2012), under the conditions employed here (20 mM sodium phosphate, 0.2 mM EDTA, 0.02% sodium azide, pH 7.4), YDL085CA at equimolar concentration was shown to have a mild inhibitory effect. There are a number of possible reasons for this effect. As mentioned above, YDL085CA appears to be extremely sensitive to buffer conditions. This buffer sensitivity might be expected to be common to members of the MOAG-4 family and therefore the addition of sodium in the buffer used in the SERF1 experiments could contribute to the variation in the effect observed. A second possible reason for the variation observed in the effect of YDL085CA and SERF1 is the difference in the sequence of the two proteins. While MOAG-4, SERF1 and YDL085CA share several conserved regions their overall sequence identity is less than 50% with the C-terminus in particular showing variation. These additional regions and sequence variations present in YDL085CA may play a role in mediating or regulating the effect of YDL085CA. Several crosslinks to Aβ40 were observed in this C-terminal region indicating that this region comes into close proximity to Aβ40 during interactions and as such provide evidence of a potential interaction.

The concentration dependence of YDL085CA activity is evidently more complicated than might be expected for a simple activity (Arosio *et al.* 2014). This analysis was not used here as it does not sufficiently describe the observed results in that it does not describe non-linear concentration dependence (in addition it requires complex data fitting which functions only on purely sigmoidal curves). Analysis of the total data set suggest that the addition of YDL085CA has little effect on the aggregation of Aβ40 until a fourfold excess is added at which point substantial inhibition is observed as shown by the increase in lag time and T_{50}. There are several possible explanations of this observation. A high K_d for the

Aβ40:YDL085CA interaction could result in poor inhibition at low concentrations. A second possible explanation would be that the effect of YDL085CA is mediated by a higher order species and therefore the K_d of that theoretical species would be a factor in any observed effect along with the reduction in effective concentration of active YDL085CA sites. The nESI-MS experiments presented in Figure 3.15 and the presence of crosslinked species of YDL085CA in the SDS-PAGE gels shown in Figures 3.18 and 3.21 provide some evidence for this possibility.

Native nESI-IMS-MS investigation of the formation of complexes during the co-incubation of YDL085CA and Aβ40 was found to be extremely challenging. YDL085CA monomer and some evidence of YDL085CA dimer could be observed along with a distribution of Aβ40 oligomeric species up to tetramer. However, observation of a complex between YDL085CA and Aβ40 required very long acquisition times (up to 45 min) and resulted in low signal intensity indicating that the complexes were either low in abundance, ionised poorly or were unstable in the gas phase (and possibly in solution as well). The complexes observed were 1:1, 1:2 and 2:1, a distribution which may be indicative of a 2:2 complex undergoing gas phase dissociation during the experiment. Native MS of YDL085CA and α-synuclein revealed a 1:1 complex. Once again this interaction was observed at a low intensity relative to the monomeric ions of YDL085CA and α-synuclein. The difficulty of acquiring such data demonstrated that a systematic study of truncations, as employed in the investigation of the ataxin-3:QBP1 interaction, or of mutants would be impractical and as such crosslinking MS was employed for further study of the interaction.

Crosslinking MS of YDL085CA and Aβ40 provided a great deal of information about both the complex and the monomers of the two proteins. The most striking observation in the crosslinking experiments was the localisation of the Aβ40:YDL085CA crosslinks to the N-terminus of Aβ40. Aβ40 has three primary amines, contributed by the N-terminus and two lysine residues, as well as two serine residues and a tyrosine residue which may be reactive under certain conditions. All of the observed inter-protein crosslinks in the 1:1 complex were formed by the N-terminus. This is strongly suggestive that the N-terminal region

plays an important role in the formation of the complex. Given the length of the crosslinker, including the spacer, is 11.4 Å this suggests that the two lysine residues of Aβ40 are either not within 11.4 Å of YDL085CA (or at least of a reactive residue) or that they are non-exposed to solvent and therefore the crosslinker cannot access the reactive sites (an eventuality that is unlikely in an intrinsically disordered protein). As previously described, the N-terminus of Aβ40 contains several charged residues and may therefore form an electrostatic interaction with the highly charged YDL085CA (Figure 3.19). The crosslinks do not necessarily indicate the direct involvement of the crosslinked residues in the formation of the complex however they do provide distance constraints for the location(s) of the interface.

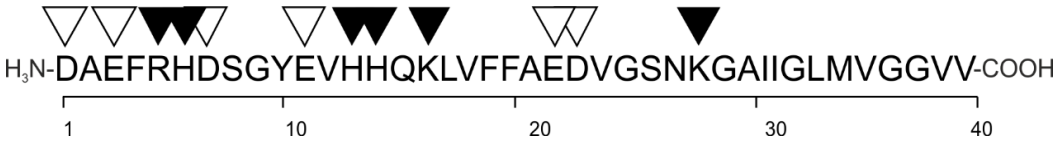


Figure 3.29: The position of charged residues in Aβ40. The sequence of Aβ40 with the position of charged residues shown by the presence above the sequence of a closed arrowhead for positive charges and an open arrowhead for negative charges. Figure adapted from (Stewart and Radford 2017).

The inter-protein crosslinks mapped onto the sequence of YDL085CA cluster to two regions. The first region, residues 8-28 of YDL085CA, forms the central region of a region proposed to from a transient α-helix and immediately C-terminal to this in a region predicted to be unstructured. NMR experiments with MOAG-4 and α-synuclein indicate that this well conserved region undergoes a conformational change upon binding (Yoshimura *et al.* 2017), however CD analysis (Figure 3.28) suggests that little change in secondary structure is observed in YDL585CA upon interaction with Aβ40. The second region, residues 45-62 also shows a cluster of crosslinks to Aβ40. This region consists of a predicted α-helix also observed in the NMR structure of MOAG-4 (Yoshimura *et al.* 2017). Interestingly K60, also in this region is not observed to form an interaction with Aβ40. The reactive residues K32 and K33, situated in a region predicted to form the N-terminal end of the second α-helical region although

observed to be disordered in the MOAG-4 structure, do not form crosslinks with Aβ40.

These findings suggest a model in which the charged N-terminal region of Aβ40 interacts with the two helical regions of YDL085CA which themselves contain an abundance of charged residues being predominantly positively charged in the first helix and containing a mixture of charges in the second helix. Interestingly the C-terminal region of the second helix is less well conserved than other regions of the protein and may play a role in the differences in activity observed between YDL085CA and SERF1. Electrostatic interactions, such as those suggested here and elsewhere (Yoshimura *et al.* 2017), can be abrogated by increasing the concentration of charged species in solution, either via the addition of salts or by changes in pH. Future studies should investigate the effect of salt and pH on the interaction between MOAG-4/SERF1/YDL085CA and amyloid proteins.

Observation of crosslinked monomers of YL085CA and Aβ40 and comparison of the crosslinking observed in the monomer with that seen in the complex suggests that YDL085CA undergoes little change in conformation upon binding Aβ40, while Aβ40 either enters an extended conformation or one in which the reactive sites are not exposed. The NMR data for the interaction of MOAG-4 and α-synuclein suggest that upon binding α-synuclein enters an extended conformation in which the highly charged C-terminus no longer shields the aggregation prone central region (Yoshimura *et al.* 2017). An extended conformation of Aβ40 could explain the loss of intra molecular crosslinking observed for Aβ40 in complex with YDL085CA. It is worth noting that the crosslinked monomer band may not have undergone crosslinking as a monomer. The crosslinking reaction is assumed to be a relatively rare event and as such a subunit in a complex may from an intra molecular crosslink and then dissociate during the SDS-PAGE step prior to being extract and analysed as monomer derived peptides. However the striking differences observed in the crosslinking of the monomer band and the Aβ40 in the 1:1 complex suggest that any contribution from dissociated complex is minimal.

In conclusion, here it has have demonstrated that YDL085CA can inhibit the formation of ThT-positive amyloid species by Aβ40. In addition it has been shown that an interaction occurs between the N-terminus of Aβ40 and the charged helices of YDL085CA. While not conclusive it seems likely that a change in the conformation of Aβ40 upon binding mediates the change in the kinetics of Aβ40. The differences between the effects observed for YDL085CA and the related MOAG-4 and SERF1a may be either the result of the apparent sensitivity of the YDL085CA:Aβ40 kinetics to buffer conditions or to the effect of the poorly conserved C-terminal region of YDL085CA.

4 Inhibition of the aggregation Ataxin-3 by polyglutamine binding peptide 1 (QBP1)

4.1 Introduction

As discussed in the Chapter 1.2, polyQ diseases such as MJD involve the aggregation of proteins containing abnormally expanded polyQ domains (Fan *et al.* 2014). Currently polyQ diseases lack effective therapeutics (Long, Tang and Jiang 2014). One molecule shown to have therapeutic potential against polyQ diseases is the undeca-peptide SNWKWWPGIFD, known as QBP1 (Nagai *et al.* 2000).

Burke and co-workers screened an M13 expressed combinatorial peptide library for a peptide which preferentially bound to glutathione S-transferase (GST)-polyQ fusion proteins containing a pathogenic length polyQ expansion (Q_{62}) (Nagai *et al.* 2000). Several peptides were isolated, all of which were tryptophan-rich and of which QBP1 showed the greatest specificity for the pathogenic length polyQ domains in a phage display screen. QBP1 was shown to inhibit thioredoxin-polyQ Q_{62} and thioredoxin-polyQ Q_{81} aggregation *in vitro*, as measured by turbidity assay. In the above assay complete inhibition of thioredoxin-polyQ Q_{62} aggregation required substoichiometric amounts of QBP1 (3:1 thioredoxin-polyQ to QBP1) while thioredoxin-polyQ Q_{81} required a much higher ratio (1:10 thioredoxin-polyQ to QBP1). A computationally scrambled peptide consisting of the same amino acid residues as QBP1 in a different order, SP1, was shown not to affect the aggregation under the same conditions. In COS-7 cells the QBP1 peptide inhibited the formation of polyQ aggregates formed by polyQ-YFP fusion proteins and, co-localised with the aggregates observed (Nagai *et al.* 2000). In addition QBP1 treated cells expressing the polyQ-YFP construct were observed to have a lower occurrence of cell death than those treated with SP1 or those untreated.

Further investigation of the sequence requirements of QBP1-mediated inhibition of polyQ aggregation revealed that the sequence of QBP1 can be shortened to WKWWPGIF with no loss of effectiveness both in *in vitro* thioredoxin-polyQ Q_{62}

aggregation assays and in *in vivo* cell survival assays with the COS-7 cells described above (Ren *et al.* 2001). In addition the GIF sequence can be replaced with a second WKWW and the spacer amino acid P can be altered to a V residue with little effect on the activity of the peptide, indicating that a proline kink is not required for QBP1 function (Ren *et al.* 2001). The WKWW domain and a second GIF/WKWW were observed to be essential to QBP1's inhibitory activity suggesting they may play an important role in mediated aggregation inhibition (Ren *et al.* 2001). A further study has identified Trp3, Trp5, Trp6, Ile9 and Phe10 as key pharmocophores for the inhibitory activity of QBP1 (Tomita *et al.* 2009).

Direct expression of a tandem repeat of QBP1 in the eyes of a *Drosophila* model of FLAG-polyQ Q_{92} both suppressed polyQ aggregation and increased the median life span of the treated animals from 5.5 days to 52 days indicating that QBP1 is an effective therapeutic *in vivo* (Nagai *et al.* 2003). The modification of the QBP1 peptide with HIV-derived trans-activator of transcription (TAT) protein domains to aid in intracellular delivery did not abrogate the function of QBP1 *in vivo* thus allowing effective delivery of the peptide into cells in more complex systems such as animal models (Popiel *et al.* 2007). To validate further the potential of QBP1 as a therapeutic agent, several QBP1-protein transduction domain fusion proteins (including QBP1-TAT) were delivered orally to polyQ expressing *Drosophila* and once again rescued the disease phenotype (Popiel *et al.* 2007).

The mechanism of QBP1 function is still poorly understood. It appears that QBP1 interacts with monomeric polyQ proteins containing extended polyQ stretches, though its effect may be mediated either by interaction with monomers or via a previously unobserved interaction with higher order species. As QBP1 does not reverse the formation of pre-existing polyQ fibrils its effect must be mediated presumably by an interaction prior to the formation of mature fibrils (Nagai *et al.* 2000).

A study of the binding of QBP1/SP1 to a thioredoxin-polyQ construct demonstrated, by means of surface plasmon resonance (SPR), that QBP1 interacts with thioredoxin-$polyQ_{62}$ with a K_d in the low µM range (5.6 µM) while

no K_d was reported for SP1 (Okamoto *et al.* 2009). This suggests a model where the lack of activity from SP1 is a result of its failure to bind to the polyQ protein.

QBP1 has been shown to prevent the transition of thioredoxin-polyQ Q_{62} proteins from their native state to the β-sheet confirmation observed in mature fibrils (Nagai *et al.* 2007). It has been suggested that the WKWW sequence can interact with an uncommon aggregation-competent conformation of the polyQ domain and that the presence of the adjacent proline interferes with the subsequent formation of β-sheet (Ren *et al.* 2001). However, the observation that the proline can be replaced with other residues with little or no loss of function suggests that at least part of this hypothesis is incorrect (Ren *et al.* 2001; Tomita *et al.* 2009). Daggett and co-workers suggested, from simulations, that QBP1 may interact with a rarely sampled α-sheet conformation (similar to a β-sheet with all of the carbonyl groups projected in one direction and all of the amino groups in the opposite direction creating two charged faces (Pauling and Corey 1951)) of polyQ proteins (Q_{40}) and thus prevent the formation or elongation of an α-sheet nucleus (Armen *et al.* 2005). However CD data suggest that the QBP1-bound thioredoxin-polyQ Q_{62} has a helical secondary structure (Nagai *et al.* 2007). Hervás *et al.* have demonstrated that QBP1 reduces the occurrence of hypermechanostable conformers of expanded polyQ ($(I27)_4$-Q_{62}-$(I27)_3$) which are observed at higher levels in longer (and therefore more aggregation prone) polyQ constructs suggesting that QBP1 may alter the conformation or stability of polyQ monomers (Hervas *et al.* 2012).

A combination of CD, fluorescence spectroscopy and NMR analyses (Figure 4.1) has revealed several interesting differences between the structures of QBP1 and SP1 (Ramos-Martin *et al.* 2014). CD data suggest that QBP1 has secondary structure consistent with a partially ordered peptide while SP1 could not be fitted to any recognised secondary structure. It was also observed that while the W4-P5 peptide bond of QBP1 is in the trans conformation the W1-P2 peptide bond in SP1 is a mixture of cis and trans. However, the location of the proline so close to the N-terminus likely reduces its effect on the conformation of the peptide. P5-G6 are shown to form a type II turn resulting in a “u-shaped” peptide

conformation. The QBP1 side-chains were observed to form a hydrophobic cluster involving Trp 4, Ile7, Lys2, Trp3 and Phe8 several of which have been identified as possible pharmocophores. In the same experiments SP1 was observed not to adopt a preferred conformation. Unfortunately little evidence for the mechanism of QBP1 function was gained from this study.

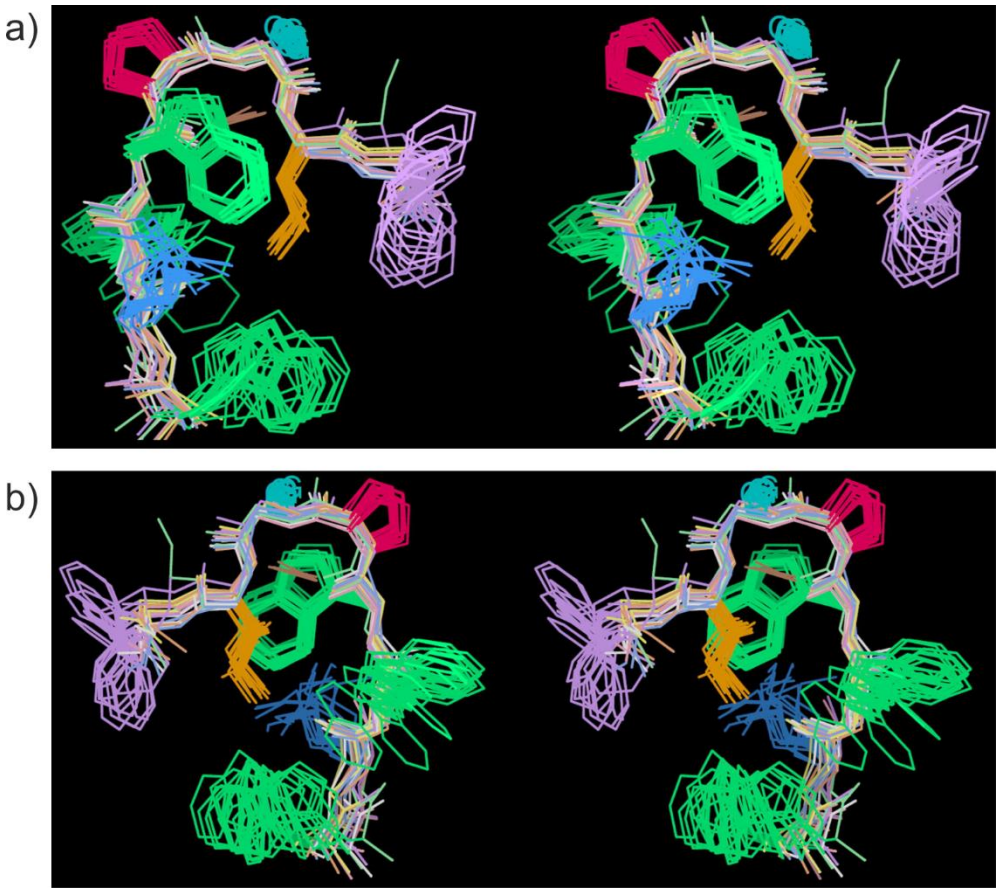


Figure 4.1: NMR derived 3D solution structures of QBP1. (a) Twenty solution structures (shown in two groups of ten for clarity) of QBP1. (b) The same twenty structures rotated 180 °. Side chains are show in colours: Trp = green, Lys = blue, Pro = red, Gly = sky blue ''G'', Ile = orange and Phe = lilac. Figure from (Ramos-Martin *et al.* 2014).

All of the above makes QBP1 appear an excellent candidate for polyQ disease therapy. However, despite promising activity in disease models it has yet to translate into an effective therapeutic. While protein transduction domain-QBP1 fusions were effective in cultured cells and drosophila polyQ disease models (Popiel *et al.* 2007), intraperitoneal administration of protein transduction domain-QBP1 fusions resulted in no effect on the motor deficiency phenotypes

of polyQ disease mice and neuronal aggregate formation was still observed (Popiel *et al.* 2009). It has been suggested that this is due to QBP1 crossing the blood-brain barrier with low efficiency (Popiel *et al.* 2009; Popiel *et al.* 2013). Viral delivery has shown promise in several instances, however, safety issues surrounding the use of viral vectors have yet to be overcome (Bauer *et al.* 2010; Popiel *et al.* 2013; Popiel *et al.* 2012). Two current areas of development are improvements to the ability of protein transduction domain modified QBP1 to cross the blood-brain barrier and the development of low molecular weight analogues of QBP1 with higher blood-brain barrier permeability (Popiel *et al.* 2013). A further problem surrounding the translation of QBP1 from the laboratory to the clinic is the current dearth of information regarding the mechanism by which it exerts its activity. As such QBP1 remains a research tool, at least for the immediate future.

The polyQ protein ataxin-3 was chosen here as the subject of the investigation into QBP1 function. A more detailed description of ataxin-3 and its aggregation can be found in Chapter 1.2.6.

The interaction of QBP1 and ataxin-3 is somewhat unusual (Figure 1.31). While ataxin-3 is a polyQ protein it also undergoes a polyQ independent aggregation step mediated by the JD to from protofibrils, as described in the Introduction (Ellisdon, Thomas and Bottomley 2006). The addition of QBP1 does not prevent this polyQ independent step of aggregation but does inhibit the subsequent polyQ dependent step (Ellisdon, Thomas and Bottomley 2006).

As with other polyQ proteins the mechanism by which QBP1 prevents the aggregation of ataxin-3 is unknown. Here we attempt to elucidate the mechanism of inhibition by means of MS and supporting biophysical techniques.

4.2 Ataxin-3 constructs

In order to elucidate the site of interaction between ataxin-3 and QBP1 a range of ataxin-3 constructs were expressed and purified as described in the Chapter 2.9.2.1 (Figure 4.2). The ataxin-3 used in this study (shown in Figure 4.2) consist of a globular N-terminal Josephin domain followed by two ubiquitin interacting motifs, a polyQ domain and a third UIM, all separated by disordered linker regions of varying length. Residues -17 to 0 consist of a hexa-histidine tag and a linker-region containing a TEV cleavage site that was not utilised in this investigation.

Two full length ataxin-3 constructs, ataxin-3 78Q and ataxin-3 14Q, were used to investigate the effect of the expansion of the polyQ domain on QBP1 binding. Ataxin-3 78Q (Figure 4.2a) contains a pathogenic length polyQ domain consisting of 77 glutamine residues and a single, naturally occurring, lysine at position 4. Ataxin-3 14Q (Figure 4.2b) contains a non-pathogenic length polyQ domain consisting of 13 glutamine residues and a lysine at position 4.

Three truncations of ataxin-3 were used in this investigation. These proteins consist of the N-terminus of ataxin-3 terminating at residues 241 (Figure 4.2c), 221 (Figure 4.2d) and 182 (Figure 4.2e) in the cases of JDU1, JD+ and JD, respectively. These proteins were used primarily to investigate the location of the QBP1 binding site.

MBP+ (Figure 4.2f) consists of the ataxin-3 residues 183-221 appended C-terminally to Maltose Binding Protein (MBP) separated by a TEV protease cleavable linker. This construct was used to confirm the binding site of QBP1 and to demonstrate that it retains its affinity in isolation.

Further information on all of the constructs can be found in Chapter 2.

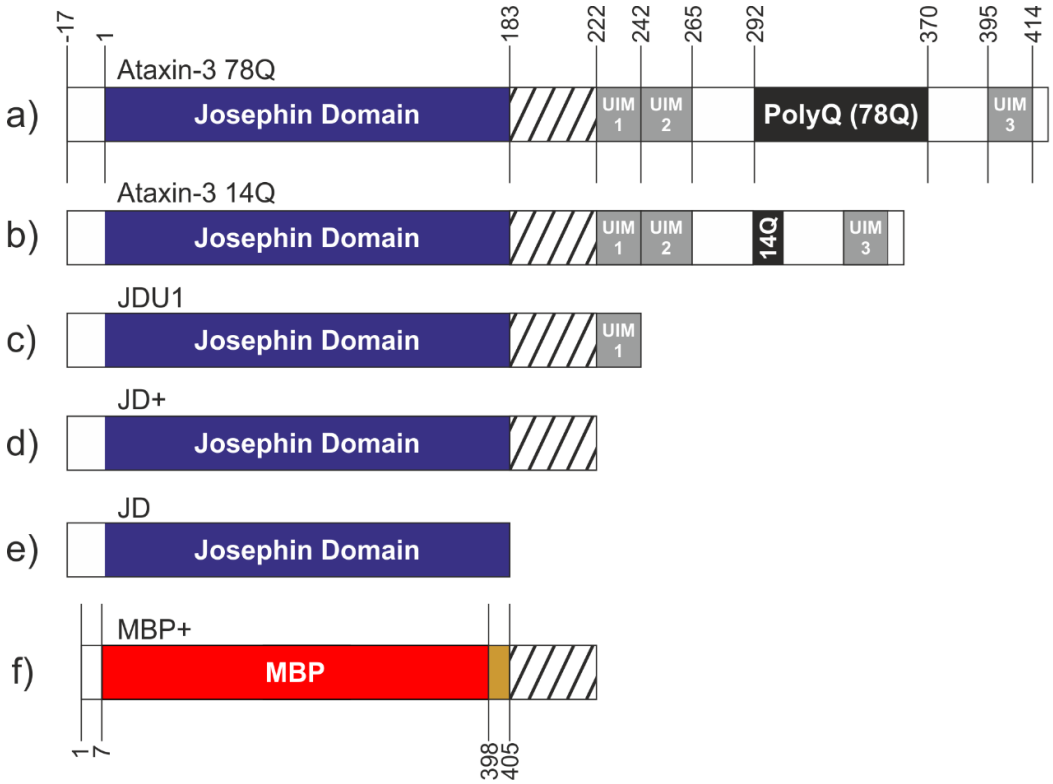


Figure 4.2: Ataxin-3 constructs used in this investigation. Ataxin-3 consists of a globular N-terminal domain (the Josephin domain (JD), residues 1-182) and a largely disordered C-terminal tail containing three ubiquitin interacting motifs (UIMs) plus a polyQ domain. The constructs used are (a) ataxin-3 78Q; (b) ataxin-3 14Q; (c) JDU1, the Josephin domain plus the first UIM (residues 1-241); (d) JD+, the Josephin domain plus residues 183-221, (e) the Josephin domain alone (residues 1-182), and (f) MBP+, maltose binding protein linked to residues 183-221 of ataxin-3, separated by a TEV cleavage site. All ataxin-3 constructs (a-e) contain an N-terminal hexahistidine tag followed by a TEV cleavage site (LENLYFQG).

4.3 QBP1 prevents polyQ dependent fibril maturation of ataxin-3 under MS-compatible conditions

Previous work on the effect of QBP1 upon the aggregation of ataxin-3 has been undertaken in conditions that are not amenable to study by MS. For example, the seminal work by Ellisdon *et al.* uses 30 µM protein in 100 mM Tris, 80 mM NaCl, 10% glycerol (v/v), 2mM PMSF, 5 mM EDTA and 15 mM β-mercaptoethanol (Ellisdon, Thomas and Bottomley 2006). Both the presence of high concentrations of salt and the increase in viscosity due to the inclusion of glycerol would make direct infusion MS problematic. As such it was necessary to make use of an alternative buffer and to validate the action of QBP1 was maintained under these conditions.

The buffer conditions decided upon were 250 mM ammonium bicarbonate, 1 mM DTT, pH 7.8. These conditions mimic the cellular pH and have been demonstrated previously to be compatible with ataxin-3 aggregation (Scarff *et al.* 2015). Doubling the DTT concentration resulted in an approximately 50% increase in the lag time of the ThT assay for JD. This is an interesting observation as it suggests that the oxidation state of cysteine (C31, C35, C131, C189 and C397) may play a role in the aggregation of ataxin-3.

The aggregation of the ataxin-3 constructs used in this investigation was examined by means of the ThT fluorescence assay. As has already been noted, in the case of ataxin-3 aggregation, it is the formation of protofibrils that results in an increase in ThT fluorescence and, therefore, the ThT assay reports on the polyQ independent aggregation step only (Ellisdon, Thomas and Bottomley 2006).

Under the conditions employed in the ThT assays, which are identical to those employed in the later MS studies with the exception of the presence of 20 µM ThT, all ataxin-3 constructs undergo aggregation with a lag time in the range of 2.5 to 7 hours (Figure 4.3 and Table 4.1). They display the typical sigmoidal profile of amyloid formation and have similar gradients indicating a similar rate of aggregation.

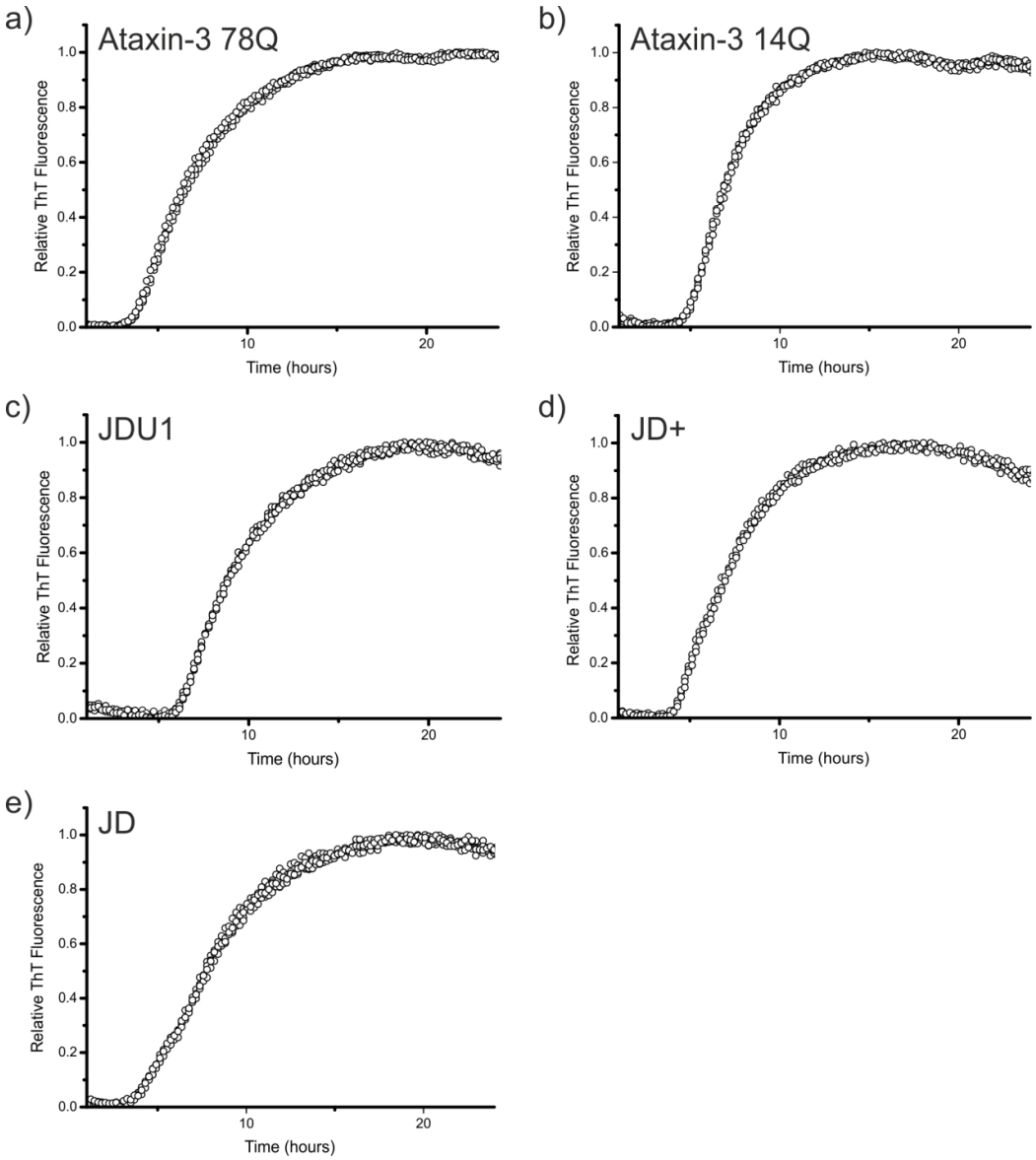


Figure 4.3: The aggregation characteristics of ataxin-3 are preserved in MS compatible buffer conditions. Ataxin-3 aggregation kinetics, measured by means of the ThT fluorescence assay, are shown for (a) ataxin-3 78Q (b) ataxin-3 14Q (c) JDU1 (d) JD+ and (e) JD. In all cases the protein was 10 µM in 250 mM ammonium bicarbonate, 1 mM DTT, 20 µM ThT, pH 7.8. The assay was performed at 37 °C without agitation. For data shown n=4.

Table 4.1: Kinetic analyses of the data presented in Figure 4.3. The lag time, T_{50} and elongation rate are shown.

	Lag time (hours)	T_{50} (hours)	Elongation rate (arbitrary units hour^{-1})
JD	**3.908 ± 0.075**	**7.775 ± 0.100**	**0.129 ± 0.004**
JD+	**3.345 ± 0.471**	**6.442 ± 0.849**	**0.164 ± 0.024**
JDU1	**5.611 ± 0.113**	**8.896 ± 0.062**	**0.153 ± 0.008**
Ataxin-3 14Q	**4.598 ± 0.032**	**6.903 ± 0.080**	**0.217 ± 0.005**
Ataxin-3 78Q	**3.249 ± 0.031**	**6.497 ± 0.132**	**0.154 ± 0.006**

Ataxin-3 78Q has been reported previously to undergo faster aggregation that ataxin-3 14Q, which has a shorter (non-pathogenic) polyQ stretch, as measured by the ThT assay (Scarff *et al.* 2015). In these data, while ataxin-3 78Q appears to aggregate faster than ataxin-3 14Q, JD aggregates with a similar rate to ataxin-3 78Q (Figure 4.3a, b, e and Table 4.1). While this may be an effect of the truncation of the protein it should be noted that there is both inter-batch and inter-experiment variation in the lag times of aggregation for the ataxin-3 constructs. This variation in the lag time may be caused by the extreme sensitivity of the kinetics of aggregation to small changes in protein concentration, buffer contaminants and temperature as well as to the stochastic nature of the nucleation events that determine the length of the lag time (Hortschansky *et al.* 2005; Morel *et al.* 2010). For the constructs JDU1 and JD+, these are the first reported aggregation kinetics.

The effect of co-incubation of ataxin-3 and QBP1 during the aggregation assays was investigated in two ways. The effect on the polyQ independent aggregation step was observed by means of the ThT assay, while the polyQ dependent fibril maturation was investigated by means of transmission electron microscopy (TEM).

In all cases the lag time and aggregation rate of polyQ independent aggregation, as measured by the ThT assay, were unchanged or, in the case of JDU1, slightly accelerated by the addition of QBP1 ($p = 0.002$) (Figure 4.4 and Table 4.2). This suggests that the polyQ independent step of ataxin-3 aggregation is unchanged by the addition of polyQ. This finding is consistent with literature reports under

other conditions (Ellisdon, Thomas and Bottomley 2006). The variation in values between Table 4.1 and Table 4.2 are a result of inter experiment variation, however, within an experiment the values are consistent.

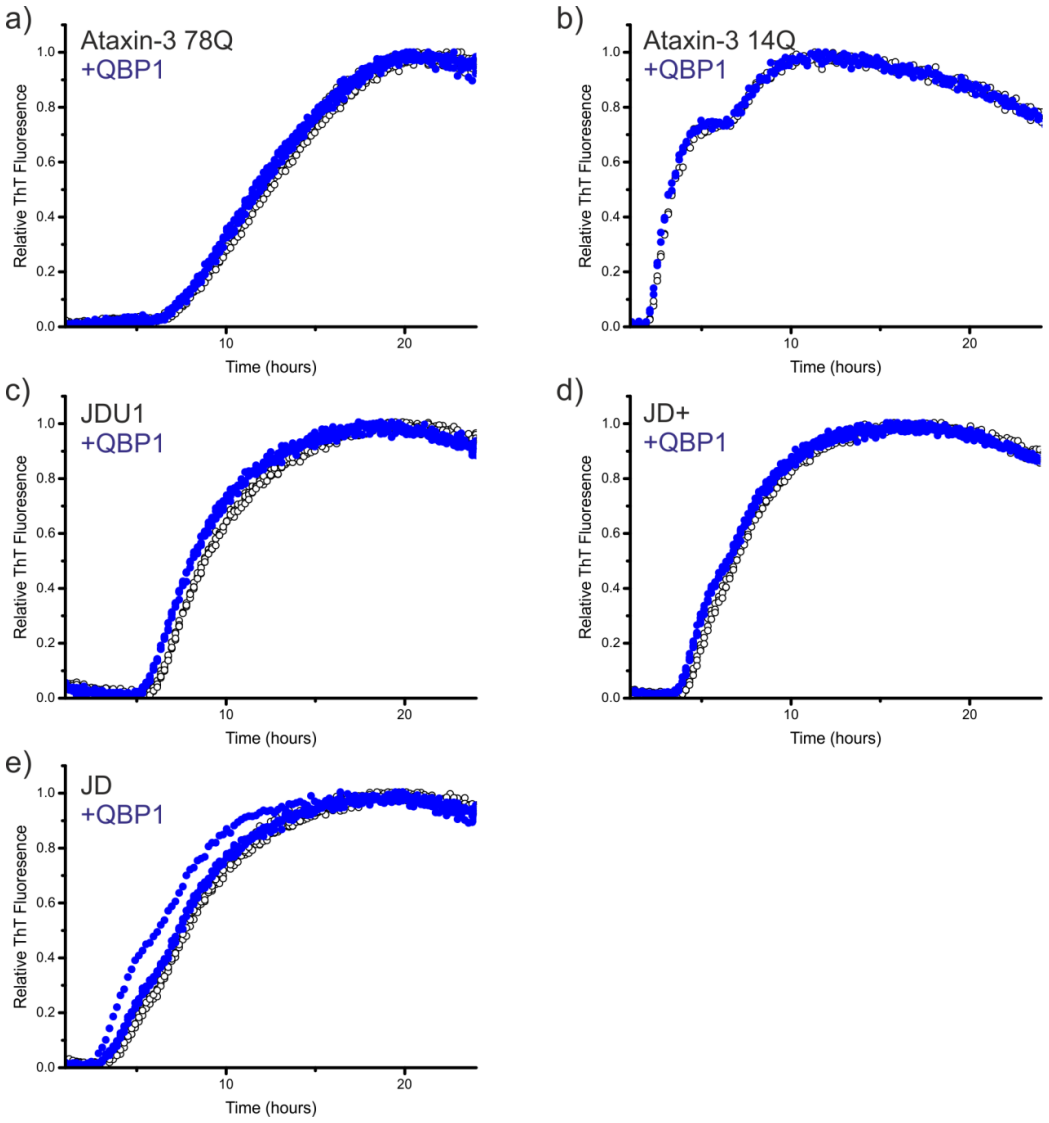


Figure 4.4: QBP1 does not alter the polyQ independent step of ataxin-3 aggregation. The aggregation kinetics of each of the ataxin-3 constructs ataxin-3 78Q (a), ataxin-3 14Q (b), JDU1 (c), JD+ (d) and JD (e) are shown. The aggregation of the protein was measured by means of the ThT assay. The protein alone is shown in black while the protein co-incubated with 50 µM QBP1 is shown in blue. In all cases the protein was 10 µM in 250 mM ammonium bicarbonate, 1 mM DTT, 20 µM ThT, pH 7.8 and the assay was performed at 37 °C without agitation (a, c-e n=4, b n=3).

Table 4.2: Analysis of the aggregation kinetics presented in Figure 4.4.

	Lag time (hours)	T_{50} (hours)	Elongation rate (arbitrary units $hour^{-1}$)
Ataxin-3 78Q	**6.8 ± 0.3**	**12.3 ± 0.1**	**0.091 ± 0.001**
Ataxin-3 78Q + QBP1	**6.7 ± 0.1**	**11.8 ± 0.2**	**0.099 ± 0.004**
Ataxin-3 14Q	**1.7 ± 0.1**	**3.4 ± 0.0**	**0.292 ± 0.021**
Ataxin-3 14Q + QBP1	**1.6 ± 0.0**	**3.3 ± 0.0**	**0.307 ± 0.001**
JDU1	**5.6 ± 0.1**	**8.9 ± 0.0**	**0.150 ± 0.005**
JDU1 + QBP1	**5.2 ± 0.0**	**8.1 ± 0.1**	**0.168 ± 0.004**
JD+	**3.3 ± 0.5**	**6.4 ± 0.8**	**0.164 ± 0.024**
JD+ + QBP1	**3.9 ± 0.0**	**6.5 ± 0.1**	**0.139 ± 0.002**
JD	**3.9 ± 0.1**	**7.8 ± 0.1**	**0.129 ± 0.005**
JD + QBP1	**3.1 ± 0.1**	**7.4 ± 0.0**	**0.116 ± 0.002**

It has been reported previously that QBP1 prevents the transition between short protofibrils and mature fibrils for ataxin-3 containing an expanded polyQ domain (Ellisdon, Thomas and Bottomley 2006). The ThT assay, shown in Figure 4.3 and Figure 4.4, measures only the formation of protofibrils. In order to observe the effect on fibril formation transmission electron micrographs were taken at the end point of the fibrillation assay (51h).

When incubated alone, ataxin-3 78Q forms long fibrilliar structures similar to those observed for other amyloid species (Figure 4.5a) (Bevivino and Loll 2001; Chow *et al.* 2004). The fibrils observed are typically 500 nm to 1 µm in length, although some form large bundles making length estimates difficult. In addition at this time point smaller protofibrilliar structures are also observed. When incubated with QBP1 at a five-fold excess, ataxin-3 78Q is observed to form only protofibrilliar species at this time point (Figure 4.5b).

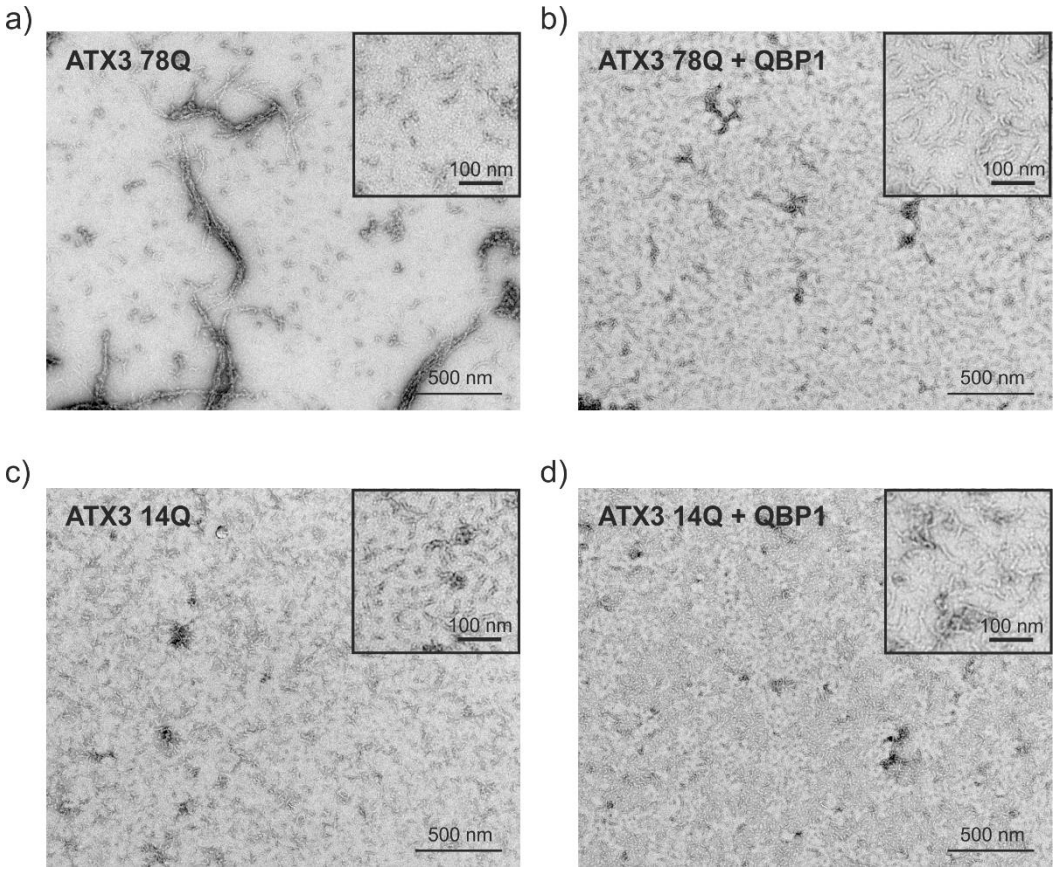


Figure 4.5: The presence of QBP1 inhibits the polyQ dependent transition from protofibrils to fibrils. TEM images of (a) ataxin-3 78Q alone and (b) co-incubated in the presence of QBP1 show that fibrils are formed at this time point in the absence of QBP1. Ataxin-3 14Q, which does not form fibrils, is observed to form protofibrils both alone (c) and in the presence of QBP1 (d). Insets show protofibrils present in all samples. In all cases the protein was 10 µM in 250 mM ammonium bicarbonate, 1 mM DTT, 20 µM ThT, pH 7.8. The samples were prepared after 51 hours of incubation at 37 °C without agitation.

Ataxin-3 14Q, which has been shown not to undergo the polyQ dependent stage of fibril maturation (Ellisdon, Thomas and Bottomley 2006), is observed to form protofibrillar structures which appear similar to those observed in the ataxin-3 78Q samples (Figure 4.5c). Co-incubation of ataxin-3 14Q with a five-fold excess of QBP1 causes no morphological changes in the protofibril that can be observed by TEM (Figure 4.5d).

These data suggest that ataxin-3 aggregation proceeds in MS-compatible conditions in a manner comparable to that observed in other buffers (Ellisdon, Thomas and Bottomley 2006). In addition, the effect of QBP1 on fibril formation

is maintained under the conditions employed here, at least at the time points observed. Therefore, the later experiments under these conditions may be considered reflective of the situation in other experimental conditions previously reported.

4.4 Native MS reveals a novel site of interaction between ataxin-3 and QBP1

As discussed previously, QBP1 was developed by screening a peptide library for peptides which bound to GST-polyQ fusion proteins with greater affinity when the polyQ region was expanded (Q_{62}) than when it was short (Q_{19}) (Nagai *et al.* 2000). Given the differences in the lengths of polyQ expansion required to cause disease and the obvious influence of the surrounding sequence and structure on polyQ aggregation, exactly where the threshold for QBP1 binding might lie for ataxin-3 is unknown. Native nESI-MS provides a powerful tool to investigate the binding of QBP1 to different ataxin-3 constructs.

The ataxin-3 aggregation kinetics and TEM data provide evidence that QBP1 interacts with ataxin-3 78Q, while Ellisdon *et al.* have demonstrated that QBP1 inhibits the second stage of the aggregation of ataxin-3 64Q suggesting that this construct also interacts with QBP1 (Ellisdon, Thomas and Bottomley 2006). As such the first construct characterised by native nESI-MS was ataxin-3 78Q (Figure 4.6).

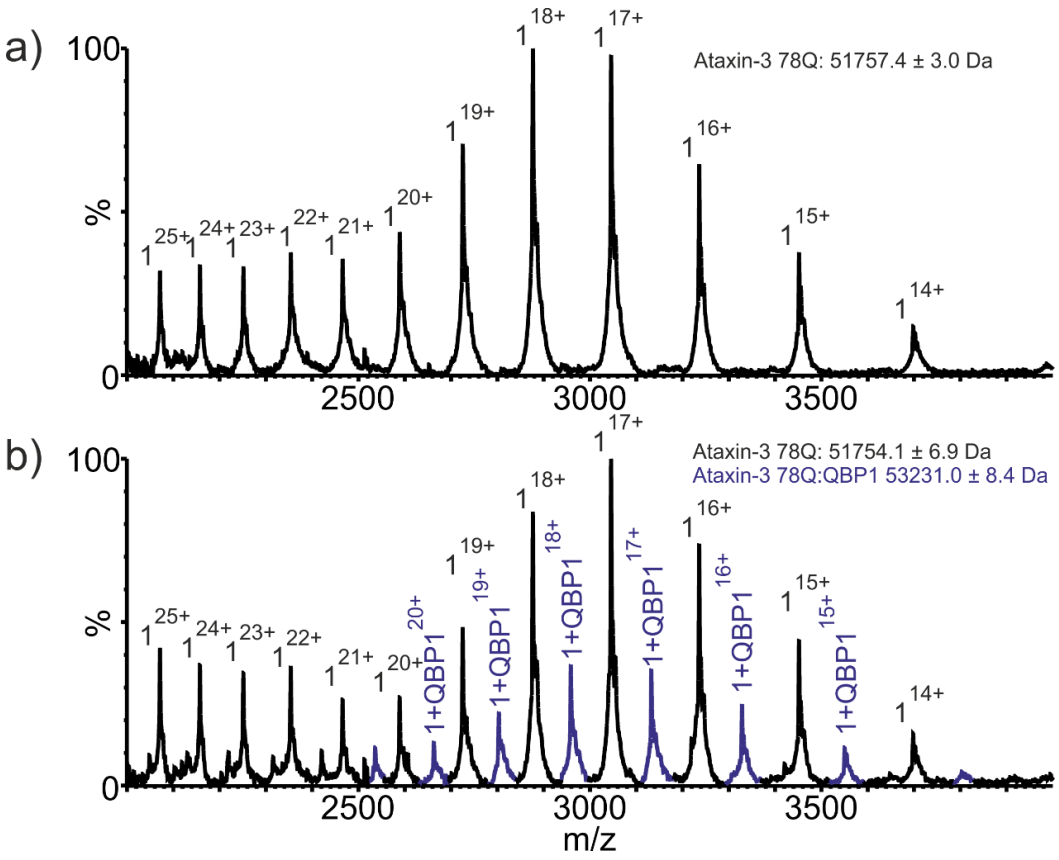


Figure 4.6: Native nESI-MS reveals a 1:1 complex between ataxin-3 78Q and QBP1. (a) The native nESI-MS spectrum of ataxin-3 78Q. (b) The native nESI-MS spectrum of ataxin-3 78Q in the presence of QBP1. A 1:1 complex is observed as indicated. Ataxin-3 78Q ions are labelled in black and ions corresponding to the 1:1 complex are labelled in blue. In both cases the protein was 10 µM in 250 mM ammonium bicarbonate, 1 mM DTT, pH 7.8. QBP1 was added at 50 µM.

Under the conditions employed here the nESI-MS spectrum of ataxin-3 78Q consists of two broad distributions of ions (Figure 4.6a). The ions M+14H^{14+} to M+20H^{20+} form a Gaussian distribution typical of a folded protein. A second charge state distribution, from M+21H^{21+} and above, has been observed previously for ataxin-3 by use of MS (Scarff *et al.* 2015; Scarff *et al.* 2013) and is attributed to the presence of the largely disordered tail region (residue 182 to the C-terminus). Under these conditions dimer and higher order structures are not observed.

Upon the addition of QBP1 (Figure 4.6b), peaks corresponding to a 1:1 complex of molecular mass 53231.0 Da are observed (the theoretical mass of a 1:1 complex is 53228.6 Da). These ions are observed primarily in the compact charge

state range $M+14H^{14+}$ to $M+20H^{20+}$. No evidence is observed of complexes containing higher numbers of QBP1 molecules. A small shift in charge state distribution is observed in some samples with the primary ion moving from $M+18H^{18+}$ for ataxin-3 78Q alone to $M+17H^{17+}$ for the unbound ataxin-3 78Q in the presence of QBP1. These data suggest that the complex observed is specific. It is interesting to observe that a single QBP1 is observed to interact with ataxin-3 78Q. If the polyQ region formed the site of interaction it might be expected that multiple copies of QBP1 would bind along its length.

Like ataxin-3 78Q, nESI-MS shows that ataxin-3 14Q also gives rise to two charge state distributions when analysed using nESI-MS (Figure 4.7a). The compact CSD is centred on $M+16H^{16+}$ and a more extended CSD from $M+20H^{20+}$ and upwards. In addition to the monomeric ions, dimeric ions are observed in the range above 3000 m/z. The addition of QBP1 once again results in the formation of a 1:1 complex between the peptide and ataxin-3 14Q (Figure 4.7b). The complex is formed primarily by the ions in the compact CSD. The intensity of dimeric ions is reduced slightly however this may be due to reduced ionisation efficiency in the presence of the high concentration of peptide and the resulting decrease in the signal-to-noise ratio.

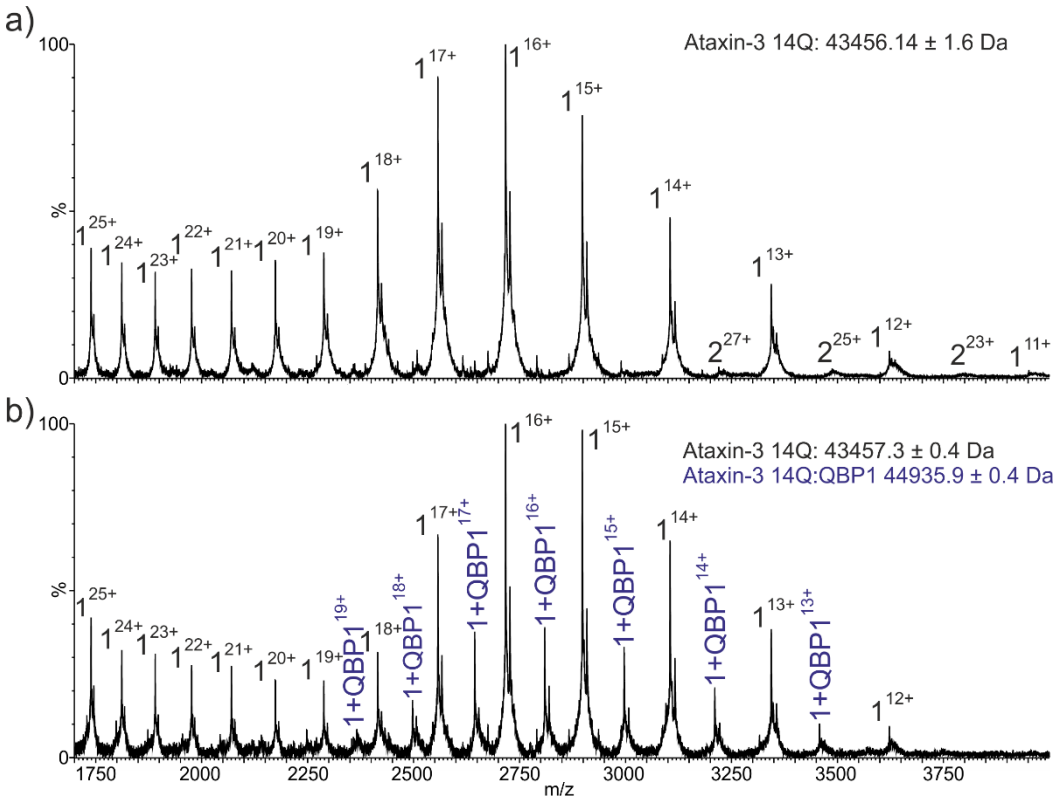


Figure 4.7: Native nESI-MS reveals the 1:1 complex between ataxin-3 and QBP1 is maintained for ataxin-3 constructs with sub-pathogenic length polyQ regions. (a) The native nESI-MS spectrum of ataxin-3 14Q. (b) The native nESI-MS spectrum of ataxin-3 14Q in the presence of QBP1. A 1:1 complex is observed as indicated. Ataxin-3 14Q ions are labelled in black and ions corresponding to the 1:1 complex are labelled in blue. In both cases the protein was 10 µM in 250 mM ammonium bicarbonate, 1 mM DTT, pH 7.8. QBP1 was added at 50 µM.

The finding that ataxin-3 14Q is able to bind to QBP1 suggests that an expanded polyQ domain is not required for ataxin-3 to bind QBP1. It is worth noting QBP1 was counter-selected against a GST-fusion protein containing 19 Qs and that no measurable binding was reported to this glutathione S-transferase Q_{19} construct when measured using surface plasmon resonance (Nagai *et al.* 2000; Okamoto *et al.* 2009). The observation of such an interaction here may indicate that the interaction is mediated by a region or structure that is not present in the glutathione S-transferase Q_{19} construct.

Given that an expanded polyQ domain does not appear to be necessary for ataxin-3 to bind to QBP1, a truncated ataxin-3 construct consisting of residues 1-221 was created by inserting a stop codon C-terminal of the first UIM. This

construct, known as JDU1, was then investigated in the same manner as the full length ataxin-3 variants already described.

Native nESI-MS of JDU1 (Figure 4.8a) reveals that under the conditions employed here the protein is primarily monomeric, with two CSDs, one centred on M+13H^{13+} and the other forming an extended distribution at and above M+16H^{16+}. The length of this extended distribution is more limited than that observed for the full length ataxin-3 constructs, terminating at M+20H^{20+}. This is perhaps indicative of the more limited range of conformations formed by the shorter disordered region compared to the full length protein.

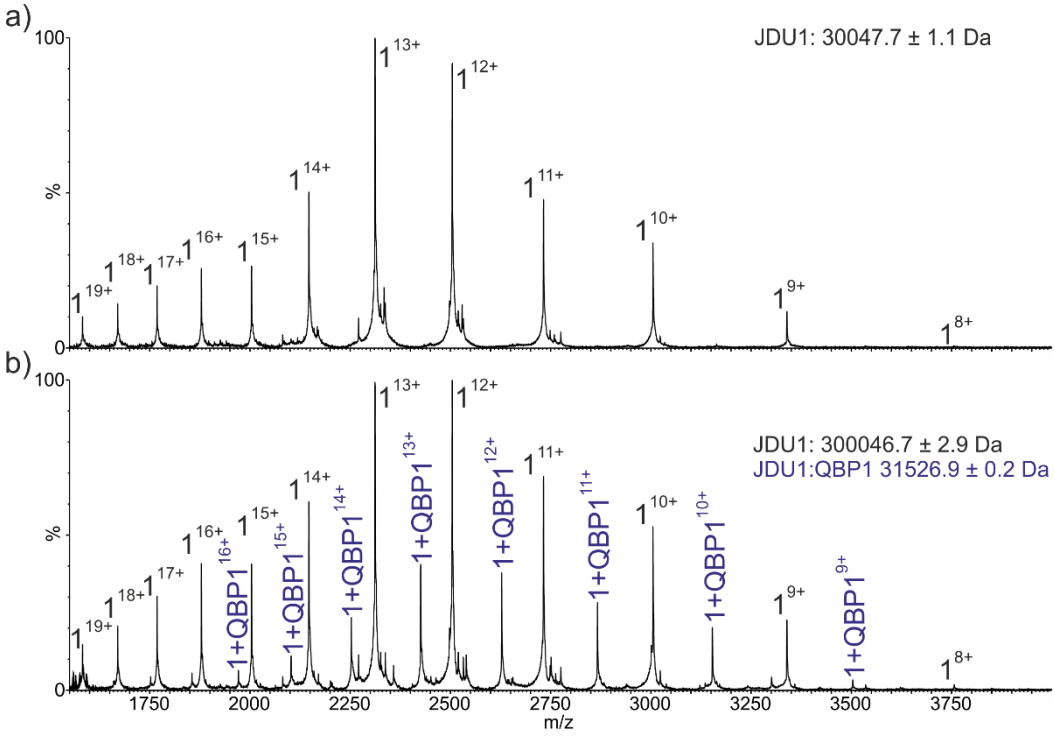


Figure 4.8: The interaction between ataxin-3 and QBP1 is mediated by a region outside of the polyQ domain (a) The native nESI-MS spectrum of ataxin-3 truncation protein JDU1. (b) The native nESI-MS spectrum of JDU1 in the presence of QBP1. A 1:1 complex is observed as indicated. JDU1 ions are labelled in black and ions corresponding to the 1:1 complex are labelled in blue. In both cases the protein was 10 µM in 250 mM ammonium bicarbonate, 1 mM DTT, pH 7.8. QBP1 was added at 50 µM.

Unexpectedly, after the addition of QBP1 a 1:1 complex is observed. As with the full length ataxin-3 constructs ions corresponding to the complex are primarily observed from the compact CSD. This finding suggests that the complex being

observed in these experiments must form at a novel site, outside of the polyQ domain, in the region 1-221. Given that both complexes formed by the full length protein and the truncated variant, JDU1, observed in the nESI-MS experiments are 1:1 complexes, it seems likely that the complex formation is mediated by the same site. Further support for this hypothesis is given by the fact that there is little evidence for the formation of significant amounts of higher order complexes between ataxin-3 and QBP1.

In the native nESI-MS spectrum of the JD (Figure 4.9a) a single CSD is observed for monomeric ions, centred on the $M+9H^{9+}$ ion. While the other ataxin-3 constructs examined here have a second expanded CSD the JD does not. The JD lacks the large disordered regions of the other constructs and is globular (Nicastro *et al.* 2005). A minor population of dimer can also be observed in the range above 3000 m/z.

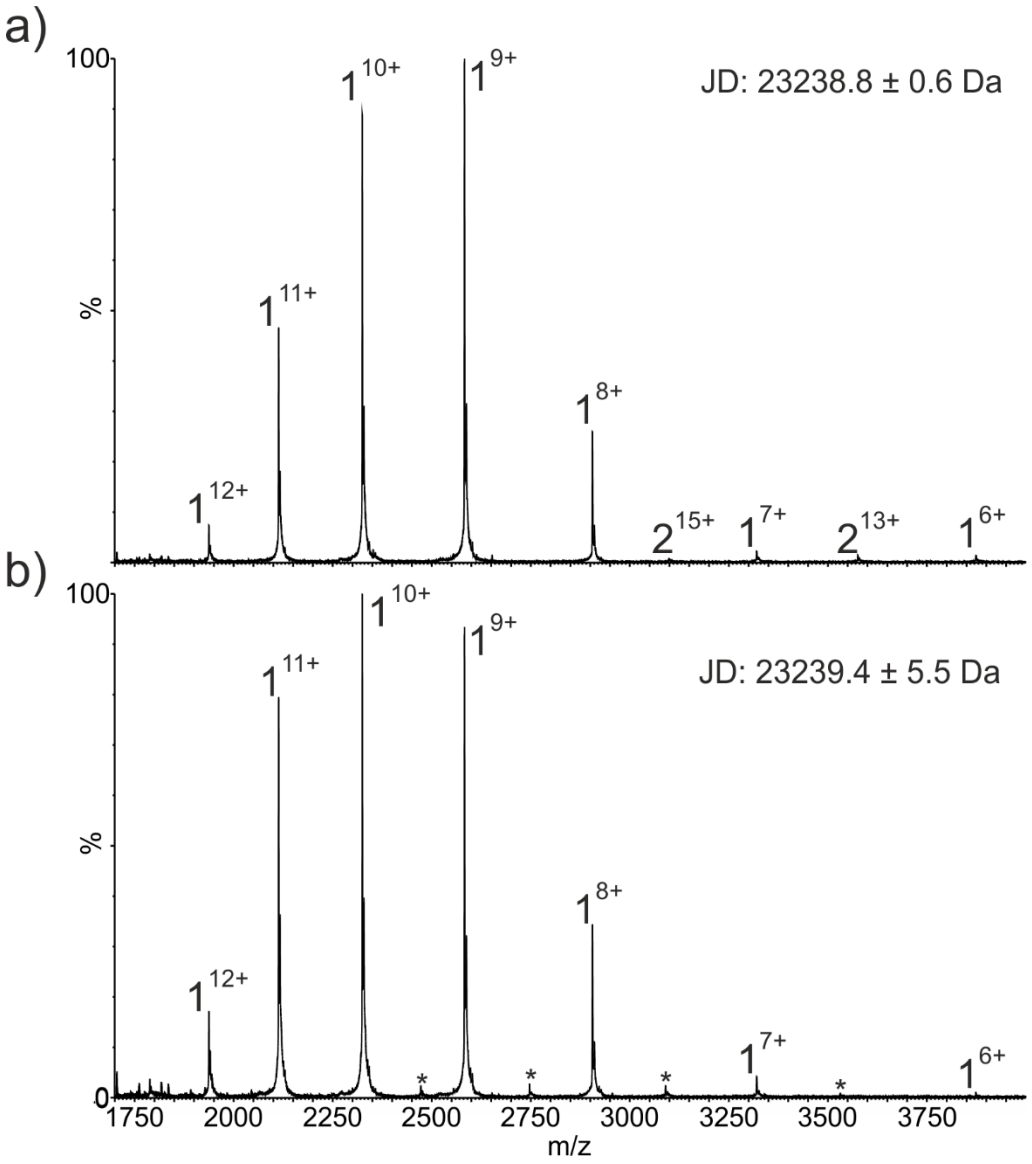


Figure 4.9: The Josephin domain alone does not interact with QBP1. (a) The native nESI-MS spectrum of the JD. (b) The native nESI-MS spectrum of JD in the presence of QBP1. Only low intensity ions corresponding to a 1:1 complex are observed (marked with *). JD ions are labelled in black. In both cases the protein was 10 µM in 250 mM ammonium bicarbonate, 1 mM DTT, pH 7.8. QBP1 was added at 50 µM.

In contrast to the observed interaction between atatxin-3 variants and QBP1 for the larger constructs, the JD alone does not interact with QBP1 (Figure 4.9b). Upon addition of QBP1 very low intensity ions corresponding to a 1:1 complex are observed in some spectra, these may be non-specific interactions. The loss of QBP1 binding upon deletion of residues outside of the JD suggests that the site of the interaction wholly or partially lies outside of the JD in the region 183-242.

It is also possible that the site of interaction is exposed only when a "tail" is present to cause the formation of the binding competent conformation.

4.5 The ubiquitin interacting motifs of QBP1 do not mediate the interaction with QBP1

There is evidence that the first two UIMs of ataxin-3 (UIM1 and UIM2) may play a role in the aggregation of the full length protein (Invernizzi *et al.* 2013; Santambrogio *et al.* 2012). Residues 182-291 of ataxin-3 were shown to increase the aggregation of ataxin-3 constructs when present in those constructs (Santambrogio *et al.* 2012) and the addition of the same residues as an isolated sequence was shown to inhibit the aggregation of full length ataxin-3 Q_{55} (Invernizzi *et al.* 2013).

The ataxin-3 constructs which bind QBP1 in the MS screen (ATX3 78Q, ATX3 14Q and JDU1) all contain UIM1. This suggests that while UIM2 is not essential for QBP1 binding (given that JDU1 binds QBP1 but contains UIM1 but not UIM2), UIM1 may be a requisite feature for the interaction to occur. With this in mind peptides were generated equivalent to UIM1 (residues 222-241) and the region 222 to 264, which comprises UIM1 and UIM2 with the connecting linker region (Figure 4.10).

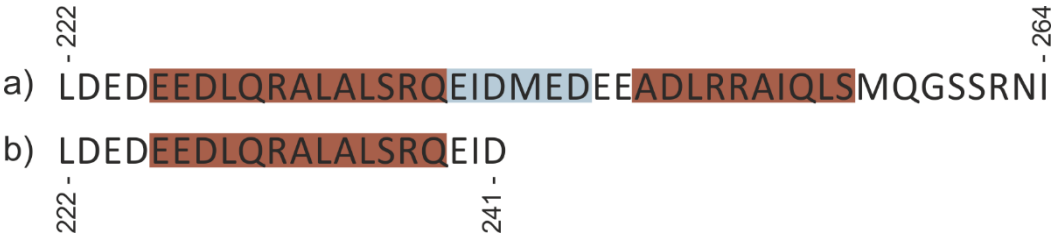


Figure 4.10: The sequences of ataxin-3 derived UIM1 and UIM12 peptides. (a) The sequence of UIM12 consisting of residues 222-264. (b) The sequence of UIM1 consisting of residues 222-241. In both cases helices are marked in red. The flexible linker region of UIM12 is shown in blue. The secondary structure details are taken from (Song *et al.* 2010).

The structure of UIM1 and UIM2 of ataxin-3 has been solved by NMR (Song *et al.* 2010). The UIMs each consist of a single helix (residues 226-238 and 247-256) and

are linked by an unstructured linker region (residues 239-244). Upon binding to ubiquitin hydrophobic contacts are formed between the two helices.

Far UV circular dichroism (CD) of the two peptides (Figure 4.11 and Table 4.3) reveals that they maintain similar secondary structure to that observed in the native context of the peptides (Song *et al.* 2010). In both cases negative maxima are at 205 nm and 221 nm indicate a mixture of helical secondary structure and random coil. Analysis of these data by means of the CDSSTR method and Dichroweb server (Table 4.3) shows that the UIM1 and UIM12 peptides are 64% and 59% helical, respectively (Compton and Johnson 1986; Whitmore and Wallace 2004). The NMR structure of the UIM12 region (Figure 1.30c) contains 53% helical secondary structure while the UIM1 region alone contains 65%. The far UV CD spectra of UIM1 and UIM12 are consistent with those observed for the related UIM from the hepatocyte growth factor-regulated tyrosine kinase substrate, Hrs (Shekhtman and Cowburn 2002).

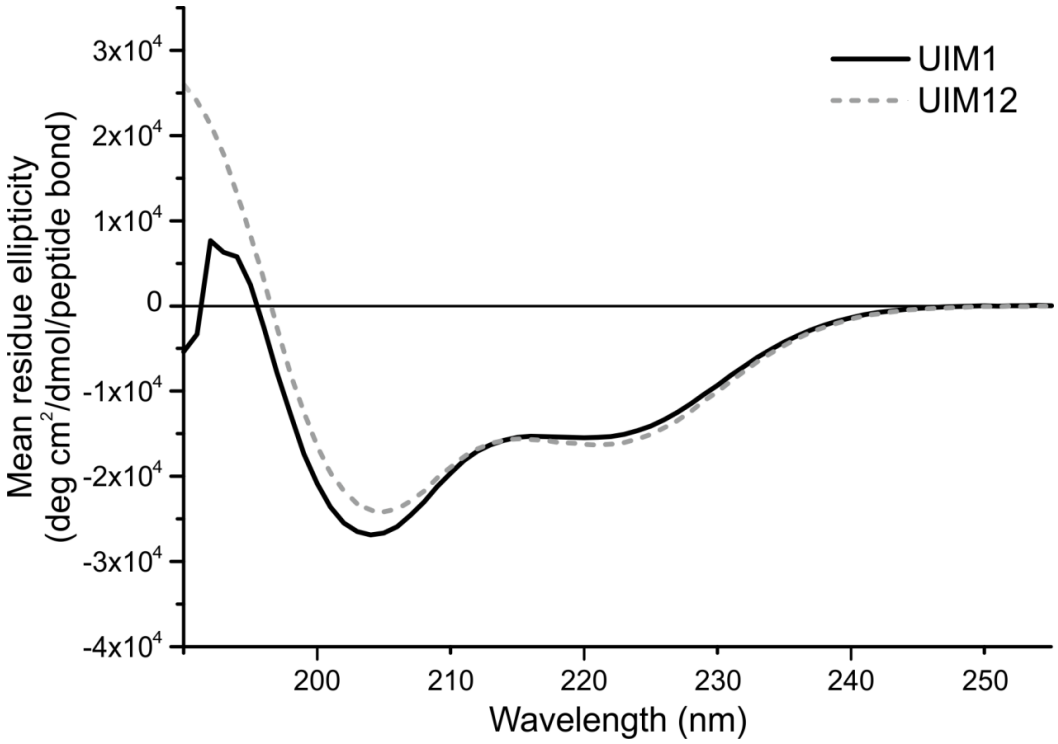


Figure 4.11: Far UV CD data reveals that secondary structure is maintained in the isolated UIM peptides. The far UV CD spectra of UIM1 (black solid line) and UIM12 (grey dashed line) over the range 190 nm to 255 nm. The negative maxima occur at 205 nm and 221 nm for both peptides. UIM1 and UIM12 were suspended at 0.25 mg/mL and 0.11 mg/mL respectively in 50 mM sodium phosphate buffer, pH 8.0 at 20 °C.

Table 4.3: Helical secondary structure content of the peptides UIM1 and UIM12. The secondary structure content of each peptide was determined using CDSSTR (Compton and Johnson 1986) and Dichroweb (Whitmore and Wallace 2004) from the data presented in Figure 4.11. These data are consistent with the secondary structure observed in the NMR structure (Song *et al.* 2010).

	Experimental α-helical content	Measured α-helical content (*Song et al.*)
UIM1	64%	65%
UIM12	59%	53%

MS of UIM1 under denaturing conditions reveals that it has the correct mass (theoretical monoisotopic mass 2357.1 Da, observed mass 2357.1 Da, Figure 4.12). It is interesting to observe that under the denaturing conditions used UIM1 formed a dimer and trimer that are not observed under native conditions (see Figure 4.13). Native MS of UIM1 was first attempted in the ammonium

bicarbonate buffer used for the investigation of other ataxin-3 constructs. However, under these conditions UIM1 proved difficult to observe. The same experiment undertaken in ammonium acetate proved more successful with peaks corresponding to the $M+3H^{3+}$ and $M+2H^{2+}$ charge states of UIM1 observed (Figure 4.13a). Native MS of a mixture of UIM1 and QBP1 at equimolar ratios (Figure 4.13b) or at fivefold molar excess of QBP1 revealed no evidence of an interaction between the peptides (Figure 4.13c).

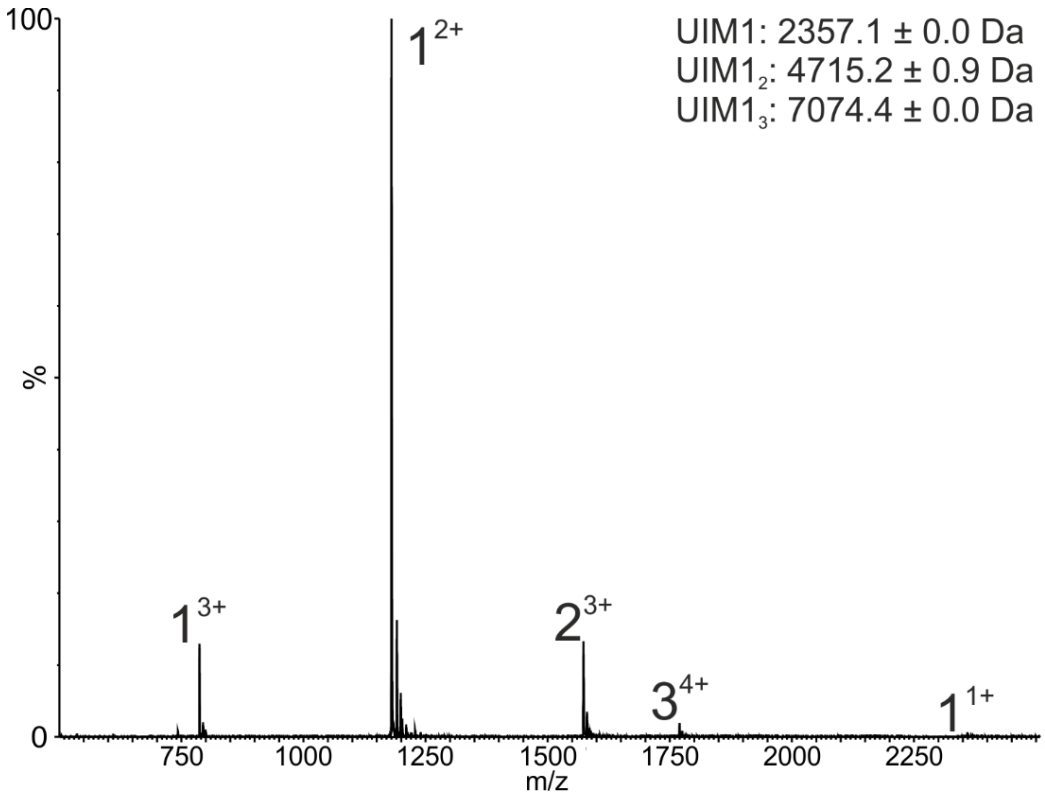


Figure 4.12: Denaturing MS of the peptide UIM1. Denaturing nESI-MS of the peptide UIM1 reveals that the peptide has the expected molecular mass. Unusually for a denatured MS spectrum, higher order species (dimers and trimers) are also observed. UIM1 peptide was suspended at 20 µM in 50% acetonitrile (v/v), 1% formic acid (v/v).

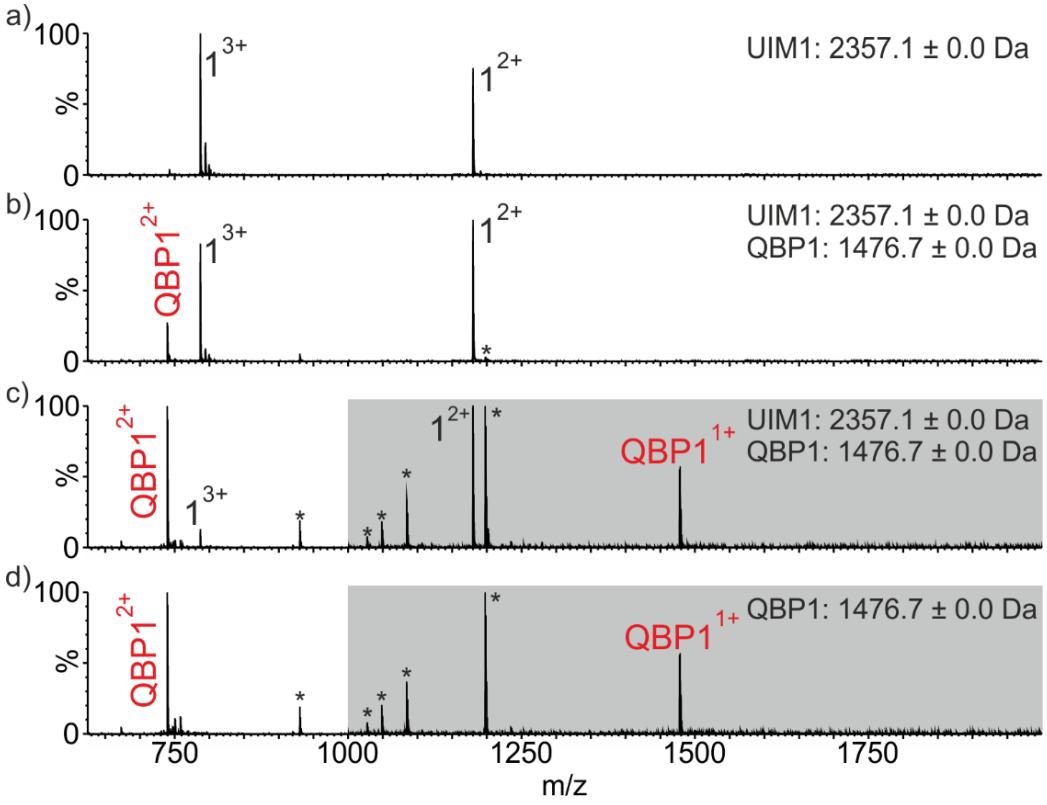


Figure 4.13: Native nESI-MS shows no evidence of a UIM1:QBP1 complex. Native nESI-MS spectra of (a) UIM1 alone, (b) UIM1 in the presence of equimolar QBP1 and (c) UIM1 in the presence of a fivefold molar excess of QBP1 show no evidence of a complex between UIM1 and QBP1. QBP1 alone (d) is shown for reference. The area shaded grey was magnified ten-fold post acquisition. A series of contaminant ions found in the QBP1 sample are indicated by the *. UIM1 peptide was suspended at 20 µM in 20 µM in 50 mM ammonium acetate, pH 7.4 to which QBP1 was added. The QBP1 in (d) was 100 µM as in (c).

The results observed with UIM12 were similar to those described for UIM1 (Figure 4.14). Native MS of UIM12 (Figure 4.14a) results in the observation of peaks corresponding to the $M+3H^{3+}$, $M+4H^{4+}$ and $M+5H^{5+}$ ions with a very low intensity peak corresponding to the $2M+7H^{7+}$ ion. No evidence of an interaction is observed in the native MS of a mixture of UIM12 and QBP1, again at equimolar ratios (Figure 4.14b) or at five-fold molar excess of QBP1 (Figure 4.14c).

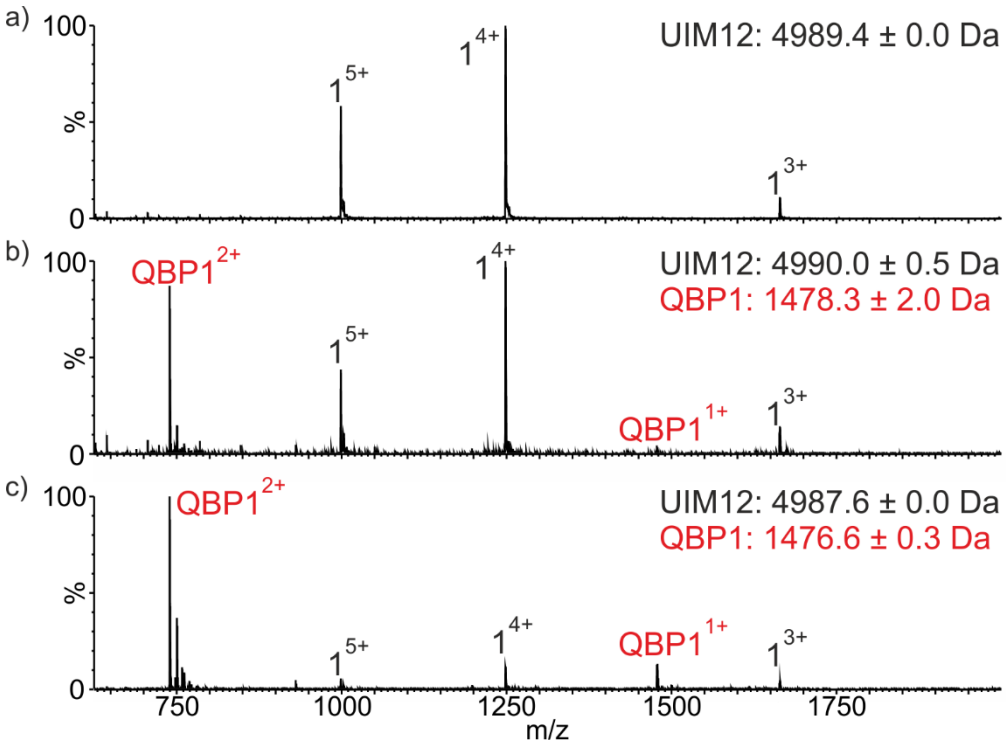


Figure 4.14: Native nESI-MS of UIM12 show no evidence of a complex with QBP1. Native nESI-MS of UIM12 alone (a) confirms that is has the correct molecular mass. After the addition of equimolar (b) or a five-fold excess of (c) QBP1 no evidence of a UIM1:QBP1 complex is observed. UIM12 peptide was suspended at 20 µM in 20 µM in 50 mM ammonium acetate, pH 7.4 to which QBP1 was added.

These data lead to the conclusion that UIM1 and UIM2 do not form a binding site for QBP1. These data do not, however, rule out the involvement of the UIMs in binding QBP1 in the context of the full length protein. In order to investigate whether the UIMs are necessary for QBP1 binding in the context of the full length protein generation of further ataxin-3 derived constructs was required.

4.6 Ataxin-3 residues 183-221 are necessary and sufficient for binding QBP1

In light of the finding that that UIM1 and UIM2 do not appear to bind to QBP1 in isolation, the sequence 183-221, consisting of the sequence between the JD and the first UIM (Figure 4.2) may be the site of interaction.

A C-terminally truncated ataxin-3 construct consisting of residues 1-221 and known as JD+ was generated as described in Chapter 2. The mass of this construct

was confirmed by MS (theoretical mass 27704.4 Da, observed mass 27704.0 ± 2.8 Da).

In the native nESI-MS spectrum of JD+ both monomer and dimer are observed (Figure 4.15a). The charge state distribution of the monomer is primarily distributed around the $M+12H^{12+}$ ion. Interestingly with the addition of the 39 residues to the C-terminal end of the JD an unfolded distribution can once again be observed i.e. the $M+15H^{15+}$ ions and above. This elongated CSD was absent in the JD mass spectra but is observed with all other ataxin-3 constructs (Figures 4.6 - 4.9). Dimer ions can be observed at higher m/z.

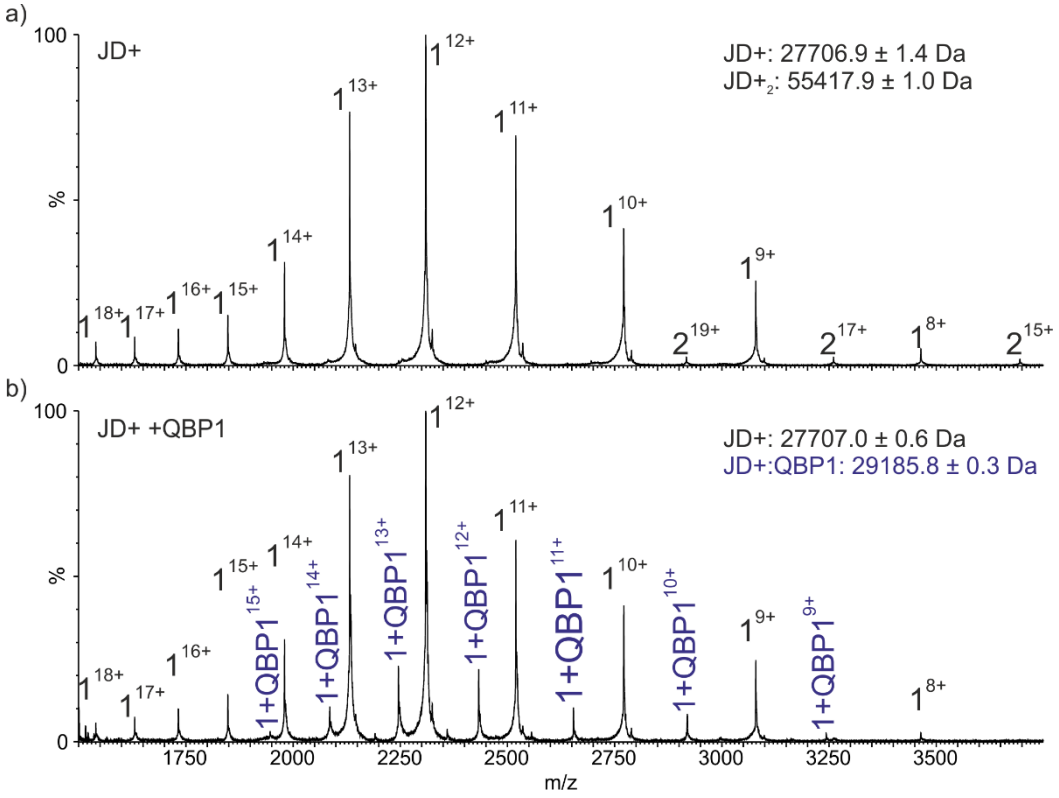


Figure 4.15: The addition of residues 183-221 restores QBP1 binding in ataxin-3 (a) The native nESI-MS spectrum of ataxin-3 truncation protein JD+. (b) The native nESI-MS spectrum of JD+ in the presence of QBP1. A 1:1 complex is observed as indicated. JD+ ions are labelled in black and ions corresponding to the 1:1 complex are labelled in blue. In both cases the protein was 10 µM in 250 mM ammonium bicarbonate, 1 mM DTT, pH 7.8. QBP1 was added at 50 µM.

The addition of QBP1 to JD+ results in the formation of a 1:1 complex which is readily observed by MS (Figure 4.15b). The CSD of the monomer is unchanged by the addition of the QBP1 peptide. Dimer is still observed in the presence of QBP1

although the signal is lower than in the absence of the peptide. This may be due to the peptide supressing the ionisation of the larger complex. Further work would be required to determine if this change is dimer distribution is a direct effect of QBP1 on the population in solution.

These data suggest that the addition of residues 183-221 renders the ataxin-3 construct able to bind QBP1. To further validate this interaction a construct consisting of residues 183-221 appended to the C-terminus of the maltose binding protein (MBP) was made. This construct is referred to as MBP+. MBP has previously been used as a solubilisation domain for fusion protein expression and to increase expression of problematic sequences (Lebendiker and Danieli 2011). MBP is unrelated to ataxin-3 or the original GST-polyQ fusion construct against which QBP1 was raised (Nagai *et al.* 2000) and has not been reported to aggregate to form amyloid fibrils. As such it is unlikely that MBP would be QBP1 binding competent therefore any binding observed to this construct can be confidently attributed to 183-221.

Native MS of MBP+ reveals a population consisting mostly of monomer with some dimer also observed (Figure 4.16a). Upon the addition of equimolar QBP1 a 1:1 complex is observed for almost all charge states of monomeric MBP+ (Figure 4.16b). The intensity of the dimer is too low for interactions with QBP1 to be detected. A small shift in the CSD of MBP+ monomer is observed upon the introduction of QBP1.

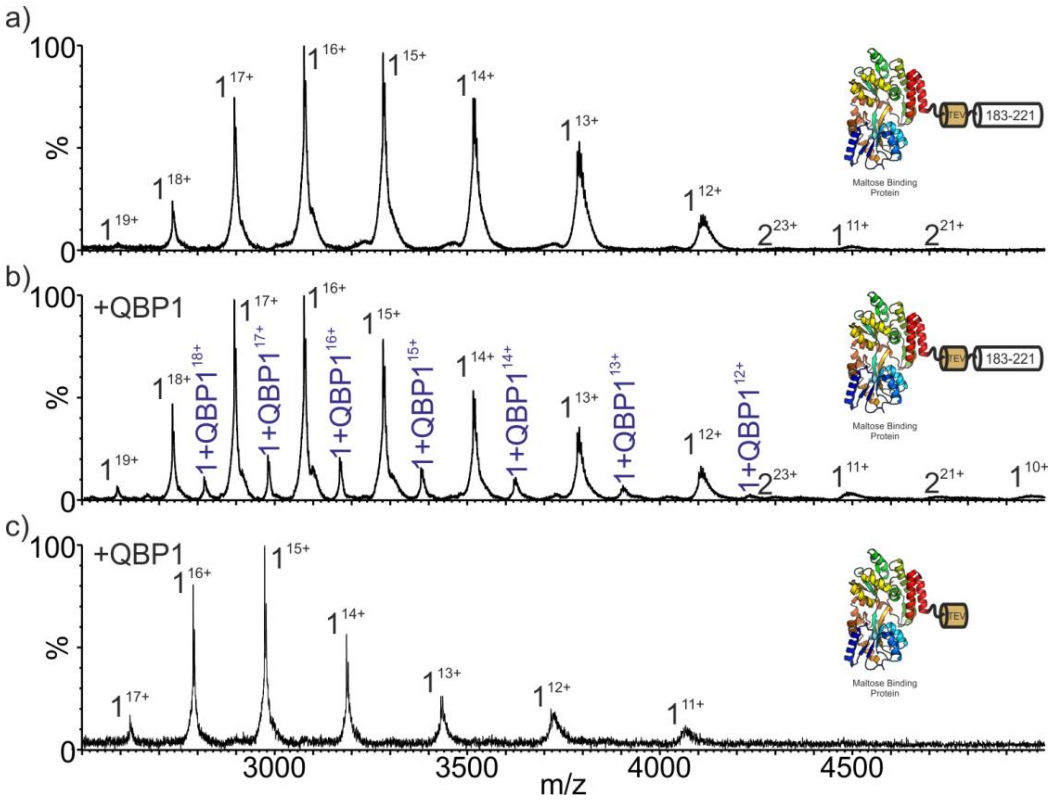


Figure 4.16: Ataxin-3 residues 183-221 enable QBP1 binding in the context of other proteins. Native nESI-MS spectra of (a) MBP+ alone, (b) MBP+ in the presence of equimolar QBP1 and (c) MBP+ TEV cleavage product in the presence of equimolar QBP1. Ions corresponding to QBP1:MBP+ complex are shown in blue. The masses of the observed ions were (a) MBP+: 49206.6 ± 1.1 Da (b) MBP+: 49206.5 ± 0.6 Da, MBP+:QBP1: 50686.8 ± 0.3 Da and (c) MBP: 44595.0 ± 13.9 Da. MBP structure adapted from (Rubin *et al.* 2002).

In order to confirm that QBP1 does not interact with MBP the MBP+ construct was cleaved with TEV. The TEV cleavage site in cleavage in MBP+ is immediately N-terminal of the ataxin-3 183-221 sequence and as such should produce two cleavage products: MBP and ataxin-3 183-221. Post cleavage, ataxin-3 183-221 was not observed, however, complete cleavage appear to have occurred as no intact MBP+ was observed (Figure 4.16c). Addition of QBP1 and MS analysis was also undertaken (Figure 4.16c). No peaks corresponding to ions formed by a MBP+ TEV cleavage product (MBP-) and QBP1 were observed which implies that the binding observed between MBP+ and QBP1 was mediated by the residues cleaved off by TEV (ataxin-3 183-221).

These findings indicate that residues 183-221 are capable of binding QBP1 both in the context of an ataxin-3 derived truncation product and in an unrelated MBP fusion protein. It seems likely, though not certain, that these residues are responsible for the interaction observed between the full length ataxin-3 and QBP1.

Ataxin-3 residues 183-221 contain a fairly equal amount of polar and non-polar residues (56% and 44%, respectively) which are disturbed throughout the peptide sequence, although a concentration of polar residues does occur in residues 200-208. The polar residues are evenly split between acidic and basic residues (18% and 21%, respectively) and are found in the centre of the region, residues 187 to 214. The net charge of the peptide at pH 7.8 is -1. Only two aromatic residues are found in 183-221. QBP1, on the other hand, contains a relatively large number of aromatics (27%) and non-polar residues (64%). The non-polar residue are cluster in the centre and C-terminal region of the sequence (residues 5-10) with polar residues situated at either terminus. The net charge of QBP1 at pH 7.8 is neutral with charges being contributed by K4 and D11. The sequences of the two peptides do not present an obvious binding motif and the lack of defined structure for 183-221 makes predicting a binding interface in three dimensions impractical.

4.7 NMR provides evidence of the QBP1:ataxin-3 interaction in solution

The data so far presented suggest that the interaction observed between ataxin-3 and QBP1 is mediated not by the polyQ domain but instead by the region 183-221 or some part thereof. In order to validate this interaction and also to confirm that the observed interaction occurs in solution as well as in the gas phase, an orthogonal technique was sought.

As noted in Chapter 1.3, the data presented here should not be treated as indicative of the relative strengths of interaction between QBP1 and the various complexes, but instead were used qualitatively to assess binding of QBP1 to the different ataxin-3 constructs in solution (Figure 4.17).

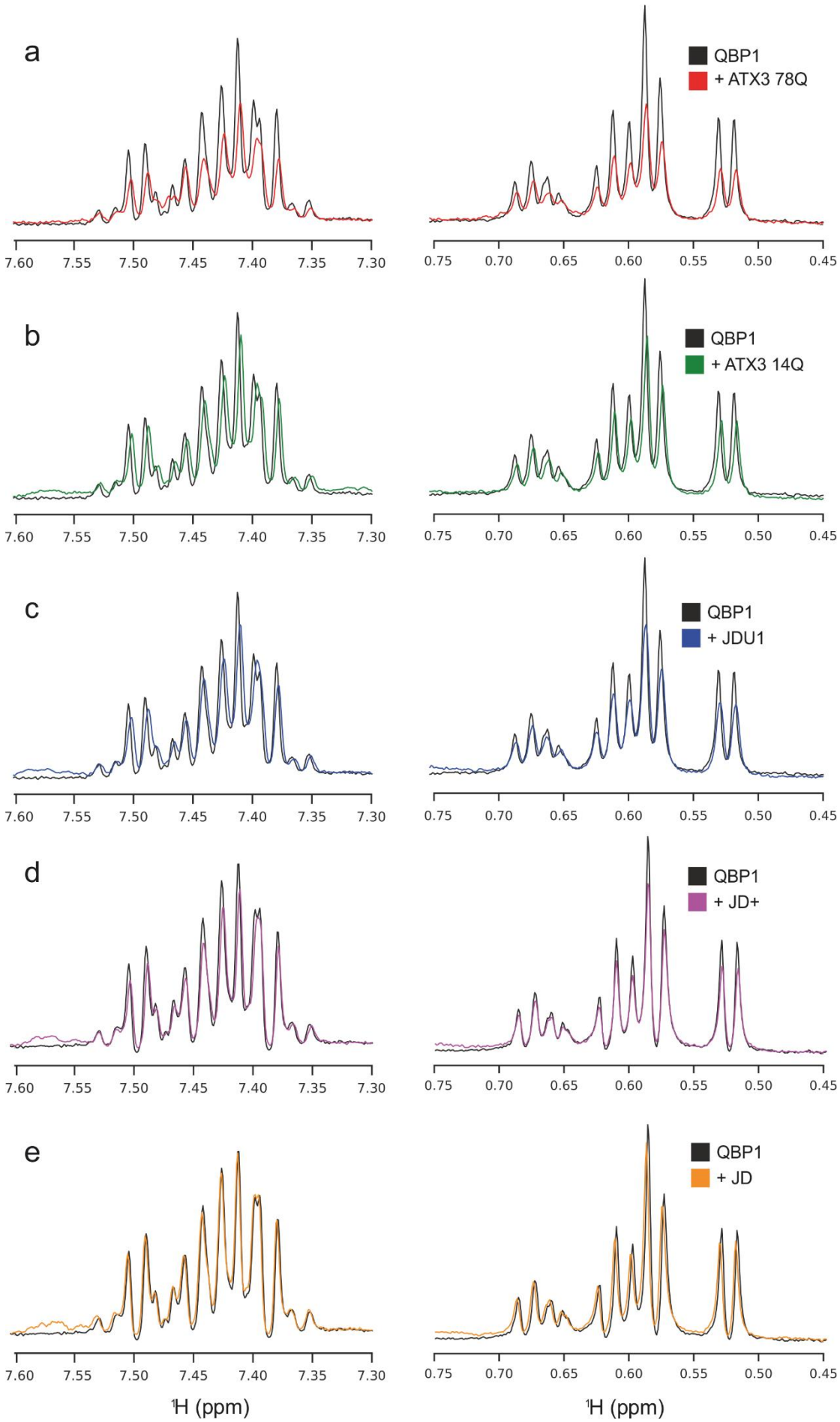

a
QBP1
+ ATX3 78Q
b
QBP1
+ ATX3 14Q
c
QBP1
+ JDU1
d
QBP1
+ JD+
e
QBP1
+ JD
7.60 7.55 7.50 7.45 7.40 7.35 7.30
0.75 0.70 0.65 0.60 0.55 0.50 0.45
¹H (ppm)
¹H (ppm)

Figure 4.17: ^{1}H-T1ρ NMR confirms that the interactions of QBP1 with ataxin-3 observed in the gas phase are reflective of interactions in solution. ^{1}H -T1ρ NMR for QBP1 with (a) ataxin-3 78Q (b) ataxin-3 14Q (c) JDU1 (d) JD+ and (e) JD. In all cases the spectrum for QBP1 alone is shown in black. Reductions in signal intensity compared to the QBP1 alone spectrum are indicative of binding. In all cases the samples were 200 µM QBP1 +/- 20 µM of the ataxin-3 variant. The data in (a) (b) and (c) were acquired by Dr T Karamanos (University of Leeds). The data in (d) and (e) were acquired by Dr A Kalverda (University of Leeds).

Table 4.4: Decrease in signal intensity in the methyl region upon the addition of ataxin-3 derived constructs relative to QBP1 alone.

Protein	Signal loss relative to QBP1 alone
Ataxin-3 78Q	24 ± 2.46%
Ataxin-3 14Q	22 ± 2.64%
JDU1	14 ± 2.75%
JD+	8 ± 1.99%
JD	-

All NMR data were acquired by Dr Theo Karamanos (University of Leeds) and Dr Arnout Kalverda (University of Leeds).

The T1ρ spectrum for QBP1 alone shows strong signals in the regions 0.45-0.75 ppm and 7.35-7.55 from methyl groups and backbone NH groups respectively (Figure 4.17). Upon the addition of ataxin-3 78Q (Figure 4.17a) a large decrease in signal is observed as detailed in Table 4.4. This is indicative of an interaction between the QBP1 and the ataxin-3 78Q protein. The addition of the other ataxin-3 constructs observed to interact with QBP1 in the MS experiments (ATX3 14Q, JDU1 and JD+) also result in a reduction in signal (Figure 4.17b-d). The magnitude of the signal reduction decreases as the C-terminus is progressively truncated this could be due to changes in the K_d of the interaction and/or the reduced size of the complex. The addition of JD causes no change in the signal intensity (Figure 4.17e) indicating that no interaction is formed between QBP1 and JD, consistent with the MS results presented above.

Under the conditions employed (temperature, protein concentration and time point) in this assay all ataxin-3 variants would be expected to be primarily

monomeric. As such the signal reduction in the T1ρ NMR data presented may be attributed to interactions of QBP1 with the ataxin-3 monomers as observed in the MS experiments.

The T1ρ NMR data presented here are consistent with the data acquired from the MS experiments. As such the two techniques serve to validate one another and provide clear evidence that the population of complexes observed in the gas phase are indicative of the state of the population in solution.

4.8 The effect of the scrambled QBP1 variant SP1 upon binding to ataxin-3

During the initial development and characterisation of QBP1 a computationally scrambled control peptide (SP1, sequence shown in Figure 4.18) was generated as a negative control (Nagai *et al.* 2000). SP1 was shown not to inhibit the fibrillation of thioredoxin-polyQ Q_{62} nor to recover the phenotype of COS-7 cells expressing cyan fluorescence protein-polyQ fusion proteins containing Q_{45}, Q_{57} or Q_{81} (Nagai *et al.* 2000). Published data did not report a K_d for any interaction of SP1 with a QBP1 substrate, thioredoxin-polyQ Q_{62} (Okamoto *et al.* 2009).

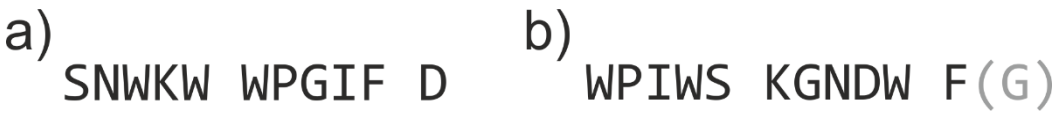


Figure 4.18: The sequences of QBP1 (a) and the scrambled peptide, SP1 (b). The additional C-terminal G in SP1G is shown in grey.

Like QBP1 (Figure 4.4), under the conditions employed here SP1 was shown to slightly inhibit the polyQ independent aggregation step as measured by the ThT assay (lag times for fibril formation: ataxin-3 78Q 2.85 ± 0.21 hours, ataxin-3 + SP1 2.2 ± 0.10 hours, p = 0.006, Figure AIII.6). However, while QBP1 has been shown to inhibit the polyQ dependent fibril maturation process (Figure 4.19a), TEMs of ataxin-3 78Q co-incubated with SP1 contain fibrils indicating that the polyQ dependent step is not inhibited (Figure 4.19b).

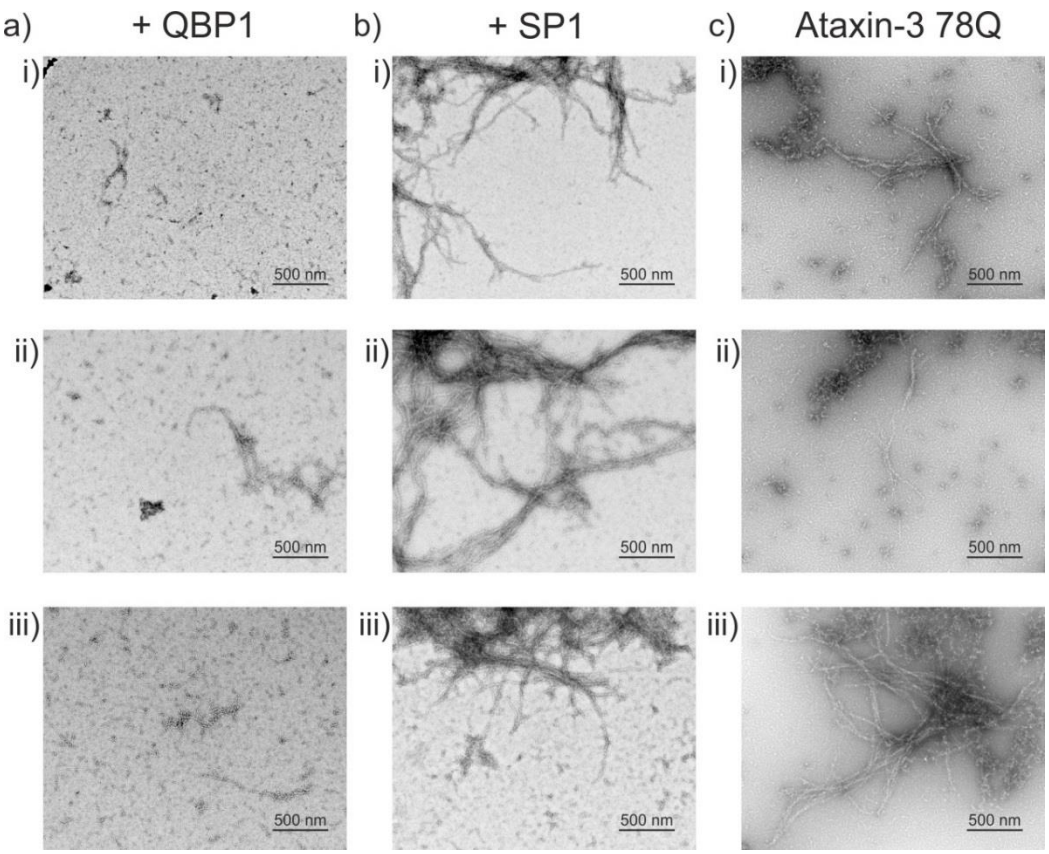


Figure 4.19: The scrambled peptide SP1 does not inhibit polyQ dependent fibril formation by ataxin-3 78Q. While the addition of QBP1 (a) inhibits the formation of fibrils by ataxin-3 78Q, the addition of SP1 (b) does not. Ataxin-3 78Q alone is shown for reference (c). i, ii and iii in each case are repeats. In all cases the protein was 10 µM in 250 mM ammonium bicarbonate, 1 mM DTT, pH 7.8. QBP1/SP1 was added at 50 µM. TEM grids were prepared after 51 hours incubation at 37 °C without shaking.

Unexpectedly, native nESI-MS of ataxin-3 78Q co-incubated with SP1 reveals the formation of a 1:1 complex (Figure 4.20a). The complex formed appears very similar to that formed by QBP1, sharing a similar CSD and occurring at similar intensities. MS spectra of other ataxin-3 constructs and SP1 reveal that the scrambled peptide interacts with ataxin-3 78Q, ataxin-3 14Q (Figure 4.20b) and JDU1 (figure 4.20c) but not with JD (Figure 4.20d).

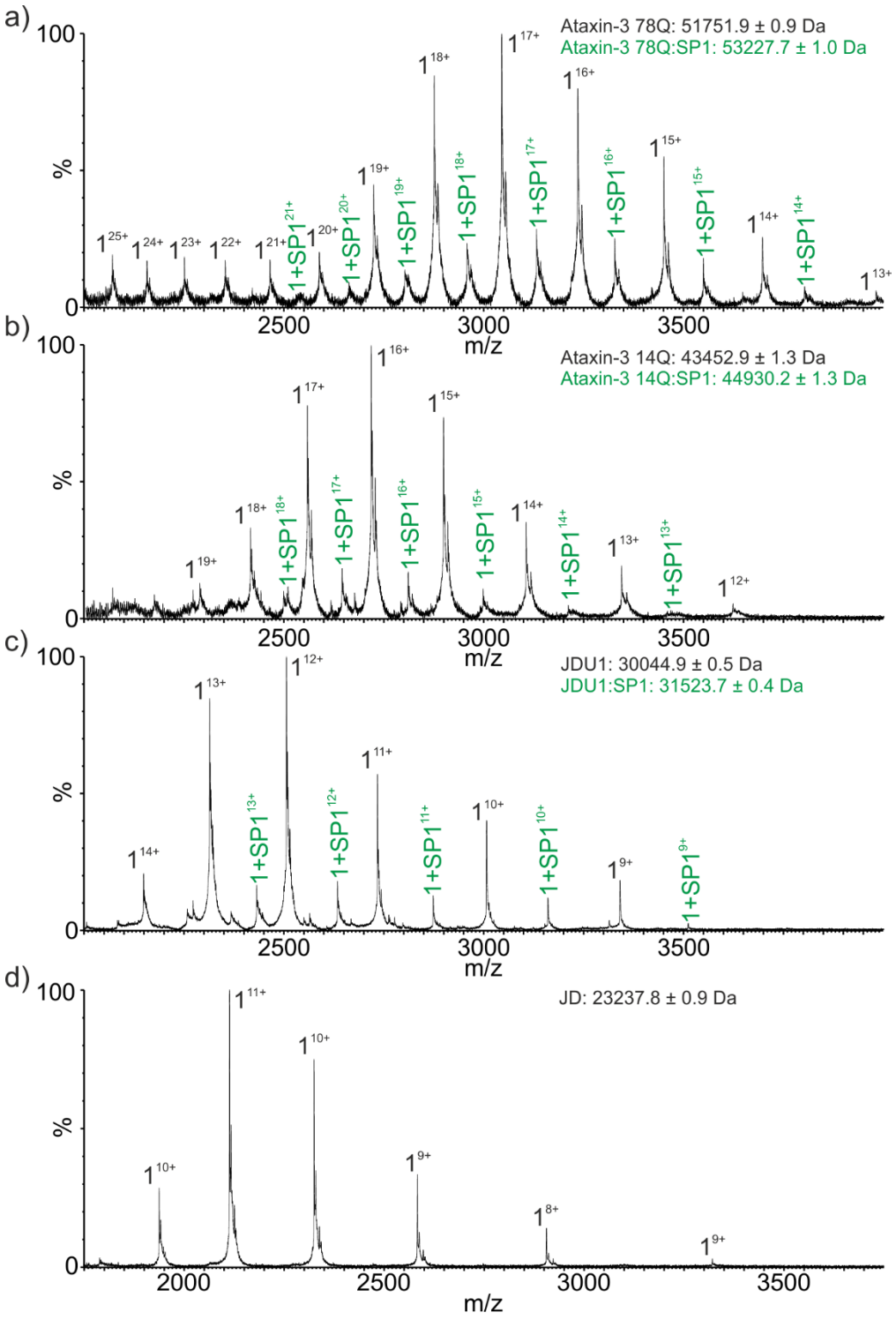


Figure 4.20: The scrambled peptide SP1 interacts with ataxin-3. Native nESI-MS spectra of (a) ataxin-3 78Q (b) ataxin-3 14Q (c) JDU1 and (d) JD in the presence of SP1. Ataxin-3 ions are labelled in black and ions corresponding to the 1:1 complex are labelled in green. In both cases the protein was 10 µM in 250 mM ammonium bicarbonate, 1 mM DTT, pH 7.8. QBP1 was added at 50 µM.

This finding raises the question of why one peptide prevents fibril maturation and the other does not given that both appear to bind in a similar fashion. In order to

investigate this further, the relative binding affinities of the two peptides with JDU1 were examined by native MS techniques.

In order to directly compare QBP1 and SP1 it was necessary to subject them to MS at the same time. QBP1 and the scrambled peptide SP1 consist of the same amino acid residues in different orders. As such, they have identical masses and so, if examined by MS together, they cannot be differentiated. An SP1 peptide with an additional glycine at the C-terminus (SP1G) was generated in order to allow simultaneous observation of QBP1 and SP1. Native nESI-MS of ataxin-3 derived protein JDU1 shows that SP1G and QBP1 can both be observed to form a 1:1 complex with ataxin-3 when co-incubated (Figure 4.21). The observed peaks have a similar intensity and occur for the same charge states. SP1G appears to form a similar distribution of complexes as the SP1 peptide indicating that the additional C-terminal residue does not affect binding.

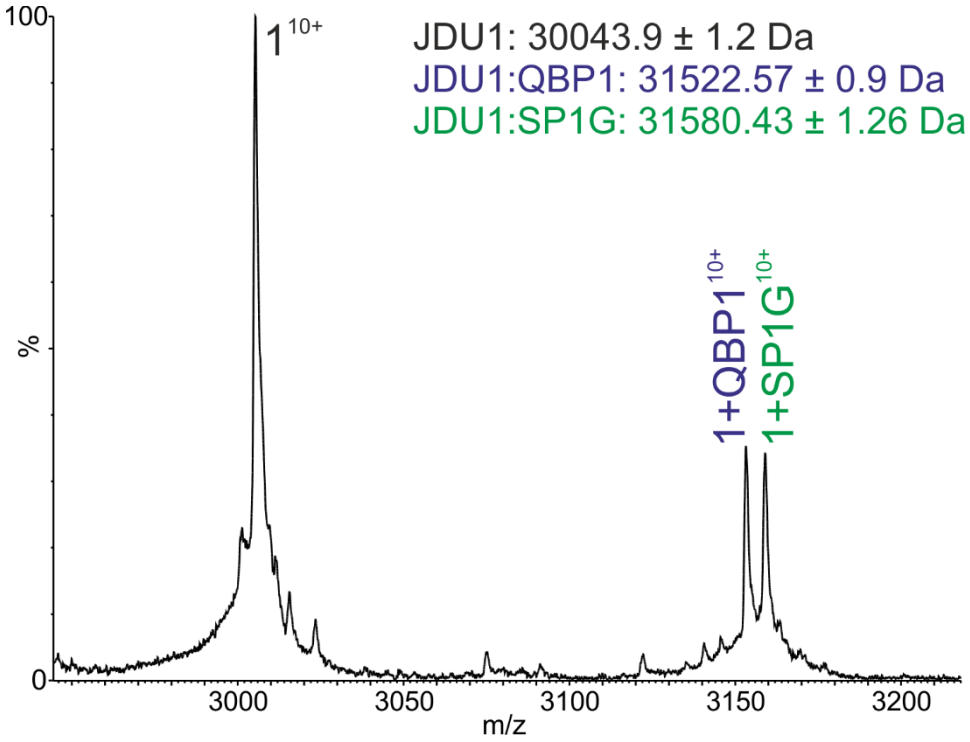


Figure 4.21: Native nESI-MS of JDU1 in the presence of both QBP1 and SP1G. Both QBP1 (blue) and SP1G (green) are observed to form a 1:1 complex with JDU1. The intensities of the two ions are comparable. The JDU1 protein was 5 µM in 250 mM ammonium bicarbonate, 1 mM DTT, pH 7.8. QBP1 and SP1G were added at 25 µM each.

Collision induced dissociation (CID) can be used to investigate the strength of non-covalent interactions in the gas phase (Chan *et al.* 2017; Hall *et al.* 2013; Han, Hyung and Ruotolo 2012; Zhang *et al.* 2014). The most abundant charge state complex of the JDU1:QBP1 complex, the $M+10H^{10+}$ ion under the conditions employed in this experiment (6 V trap collision energy, 10 V transfer collision energy and 10-fold excess of peptide), was isolated in the quadrupole. While maintaining all other instrumental conditions the trap voltage was increased in 4 V increments from 6 V to 62V at which point the complex was no longer observed. This process was repeated with the JDU1:SP1G complex. Comparison of the dissociation curves of QBP1 and SP1G (Figure 4.22) reveals that the JDU1:SP1G complex dissociates more readily than JDU1:QBP1 complex as indicated by the lower voltages required to result in a given reduction in signal. For example, the relative ion count for the JDU1:SP1G complex falls to 0.5 at a trap voltage of 22 V, while for JDU1:QBP1 the same drop in relative ion count takes place between 34 and 38 V).

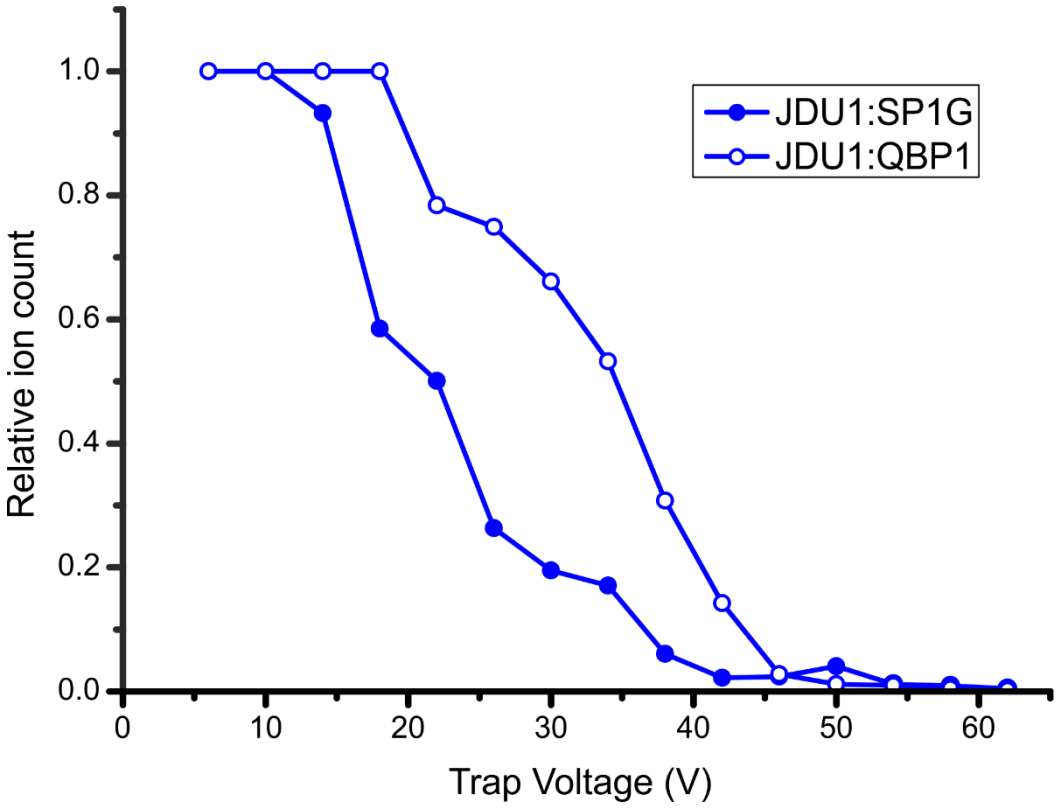


Figure 4.22: Collision induced dissociation shows that JDU1 forms a more stable complex with QBP1 than with SP1G. The M+10H^{10+} charge state of the JDU1:SP1G (open circles) and JDU1:QBP1 (closed circles) ion intensity reduces as the trap voltage is increased. The JDU1:SP1G complex dissociates at a lower voltage than the JDU1:QBP1 complex.

The lower affinity of the scrambled peptide for ataxin-3 suggests a possible reason for the lack of inhibition of fibril formation by SP1. The lower affinity may lead to lower occupancy in solution and thus reduce the effect on the equilibrium of ataxin-3 species.

Alongside the CID experiments, collision induced unfolding of the bound and free JDU1 M+10H^{10+} ion was also undertaken (Figure 4.23). Unfolding can be observed at collision energies of 30V for both complexes although JDU1:QBP1 may have undergone a greater degree of unfolding (Figure 4.23c). The unfolding profiles for the bound and free ataxin-3 are otherwise similar suggesting that QBP1 binding does not alter the inherent stability of ataxin-3 monomer in the gas phase.

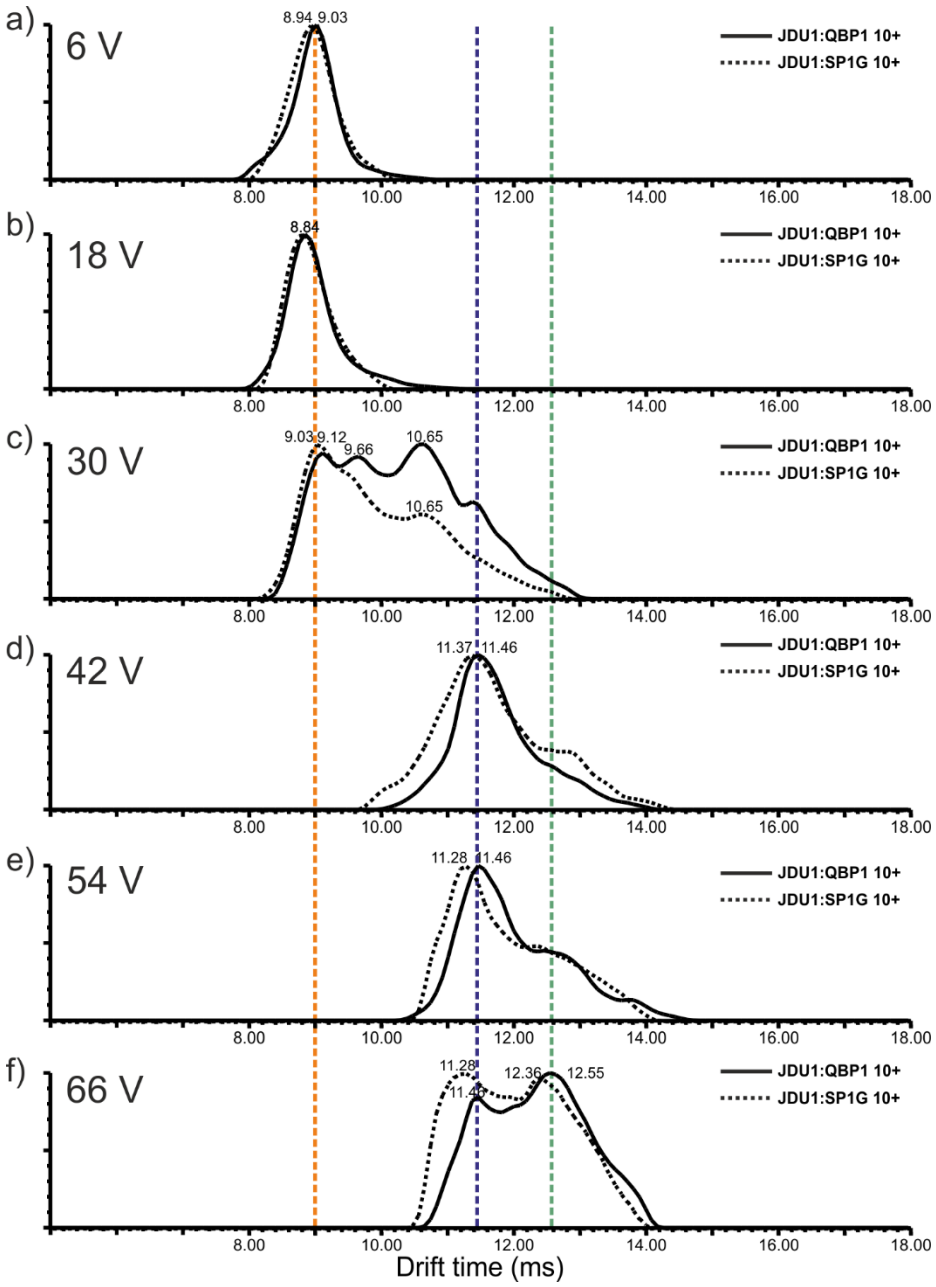


Figure 4.23: Collision induced unfolding of the JDU1:QBP1 and JDU1:SP1G complex $M+10H^{10+}$ reveals that the conformational stability of the two complexes are comparable. The arrival time distribution of the $M+10H^{10+}$ ion of the 1:1 complex of JDU1:QBP1 (solid line) and JDU1:SP1G (dashed line) are shown for increasing trap collision energies from 6V to 66V (a-f). The coloured lines indicate the arrival time of major pseudo-stable conformations.

The data presented here suggest that the binding of QBP1/SP1 is not sequence order specific. However the effect upon fibril maturation does appear to be sequence dependent. This reliance upon sequence may be due to the nature of the inhibitory interaction, possibly via a steric interference as suggested by

Daggett and co-workers (Armen *et al.* 2005) which may be missing in the scrambled peptide.

4.9 QBP1 binding does not alter the conformation of ataxin-3 monomers

One method by which an aggregation inhibitor can function is by altering the conformation of the aggregation-prone protein, or by altering a pre-existing conformational equilibrium. This has been demonstrated previously for amyloid aggregation inhibitors (Bazoti *et al.* 2008; Ehrnhoefer *et al.* 2008; Galanakis *et al.* 2011; Lu *et al.* 2013a; Shimmyo *et al.* 2008; Sivanesam *et al.* 2016; Young *et al.* 2014). While ataxin-3 has a disordered region it has been shown to occupy several distinct conformations (Scarff *et al.* 2013). To investigate this possibility that QBP1 alters the conformation of ataxin-3, nESI-IMS-MS was used to measure the collision cross-section of the ataxin-3 variants both bound to QBP1 and in their free state (Figure 4.24).

nESI-(TW)IMS-MS CCS estimation carries with it an inherent error, a value often quoted at 5% of the CCS (Bush *et al.* 2010; Smith *et al.* 2009). Measurements which vary within this range are generally not considered to be significant.

The estimated CCS for ataxin-3 14Q and the JD presented here (Figure 4.24b, e) are comparable with CCS values previously reported by Scarff *et al*. for ataxin-3 constructs analysed in 10 mM ammonium acetate (Scarff *et al.* 2013). The CCS of ataxin-3 14Q M+12H^{12+} reported by Scarff *et al.* has three conformations with CCS values of 3077 Å^2, 3395 Å^2 and 3673 Å^2 while the data here for the same charge state analysed here in 250 mM ammonium bicarbonate/1 mM DTT only two conformations are observed, with CCS values of 3054 Å^2 and 3234.74 Å^2 which are within error of the literature values. The absence of the third conformation in the data presented here may be an effect of the buffer. For the more globular JD the previously reported CCS values for the JD M+8H^{8+} ion are 1959 Å^2 and 2167 Å^2. At the same charge state the CCS determined from the data presented here are 1914 Å^2 and 2206 Å^2, within the error of the technique.

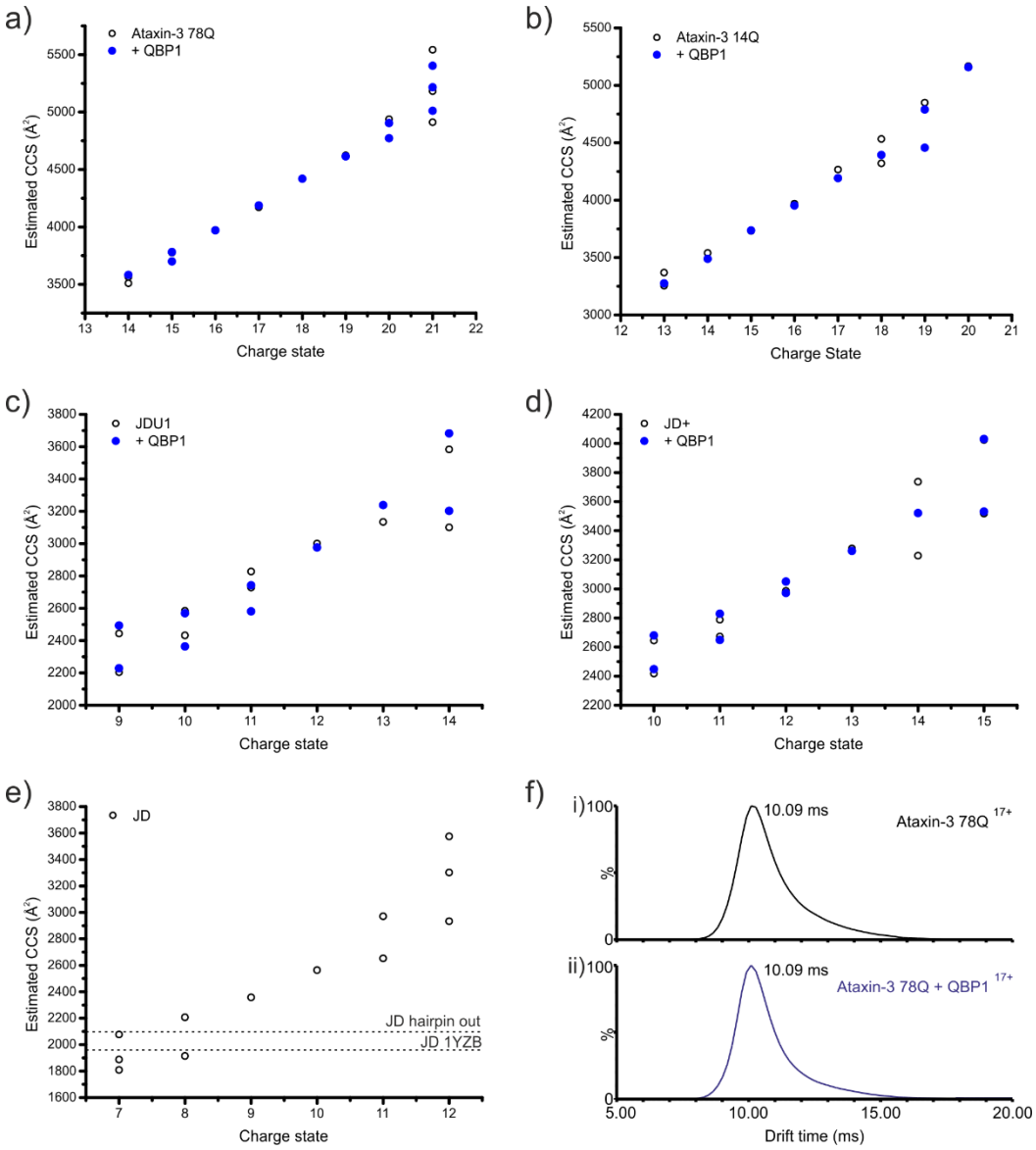


Figure 4.24: nESI-IMS-MS analyses reveals that the CCS of ataxin-3 constructs (ataxin-3 78Q, ataxin-3 14Q, JDU1 and JD+) are unchanged by QBP1 binding. Estimated CCS for each charge state of the dominant charge state distribution is shown for (a) ataxin-3 78Q, (b) ataxin-3 14Q, (c) JDU1, (d) JD+ and (e) JD (dashed lines show the predicted CCS from the PSA method, based on the structure of JD (Nicastro *et al.* 2005) and modelling of a JD with an outward pointing helical arm, first published by Scarff *et al.* (Scarff *et al.* 2013)). The ataxin-3 alone ions are shown in black and the QBP1 bound ions are shown in blue. The CCS values observed for each bound protein ion are all within 5% of the CSS of the unbound form with the exception of the JD+ $(M+14H)^{14+}$ ion. Figures quoted are the average of three replicates. (f) The arrival time distributions of the ataxin-3 78Q $M+17H^{17+}$ ion and the ataxin-3 78Q:QBP1 $(M+17H)^{17+}$ ion. The arrival time of the ion is indicated above. In all cases the protein was 10 µM in 250 mM ammonium bicarbonate, 1 mM DTT, pH 7.8. QBP1 was added at 50 µM.

Comparison of the CCS values obtained for the unbound ataxin-3 and the ataxin-3:QBP1 complexes (where they are observed) reveals that ataxin-3 does not undergo significant changes in CCS upon binding (Figure 4.23). The JD+ $M+14H^{14+}$ ion is an exception to this (Figure 4.24d), possibly due to low signal intensity leading to poor resolution of the two conformations in the samples containing QBP1 and leading to the two signals appearing as a single peak although a conformational change cannot be ruled out. This low signal for the complex is a product of the low relative signal intensity of the complex, ionisation suppression from the presence of high concentrations of the peptide and the reduced ion transmission inherent in IMS-MS experiments.

The CCS estimations from the nESI-IMS-MS data suggest that QBP1 binding does not alter the conformation of ataxin-3. Measurements by IMS lack the sensitivity to small scale changes in structure that can be observed by other techniques such as NMR and X-ray crystallography. It is possible, therefore, that small scale changes in structure do occur upon QBP1 binding however no evidence to support this is observed in these experiments.

4.10 QBP1 interacts with other amyloid proteins

Hervás *et al.* have demonstrated that QBP1 can alter the conformational properties of a range of amyloid protein monomers including Q_{62}, α-synuclein and Sup35 but not Aβ42 (Hervas *et al.* 2012). Given the interaction observed here appears to be mediated by regions outside of the polyQ domain it was decided to investigate the interaction with α-synuclein by nESI-MS as well.

Unlike the polyQ protein ataxin-3, α-synuclein undergoes a single stage aggregation process. As such the ThT aggregation assay can be used as directly indicative of the state of fibril formation. The aggregation kinetics of α-synuclein follow the characteristic sigmoidal growth of many other amyloid proteins (Buell, Dobson and Knowles 2014). Due to the high concentration of α-synuclein required to produce reproducible kinetics in a reasonable time scale (80 μM protein) the addition of QBP1 was limited to equimolar and sub-stoichiometric concentrations of peptide. Upon initial observation, the aggregation kinetics in

the presence of QBP1 appear to show no formation of ThT positive amyloid species (Figure 4.25a). However, closer inspection reveals that sigmoidal growth can be observed albeit with much reduced signal (Figure 4.25b). The lag time of fibril formation is reduced slightly by the presence of QBP1 (from 29.7 ± 7.0 hours to 13.8 ± 0.8 hours in the presence of equimolar QBP1). There is significant precedent for small molecules reducing the signal yield in the ThT assay via interference with the ThT binding or its fluorescence quantum yield rather than through a direct effect on aggregation (Coelho-Cerqueira, Pinheiro and Follmer 2014; Gade Malmos *et al.* 2017). In order to investigate this possibility TEM grids were prepared for the samples with and without QBP1. TEM (Figure 2.25c-e) reveals that after 180 hours fibrils are observed in samples of both α-synuclein alone and α-synuclein co-incubated with QBP1. The fibrils observed are straight and unbranched. Several morphologies are observed including flat ribbons. In addition to the fibrils, smaller objects are also observed which may correspond to oligomers.

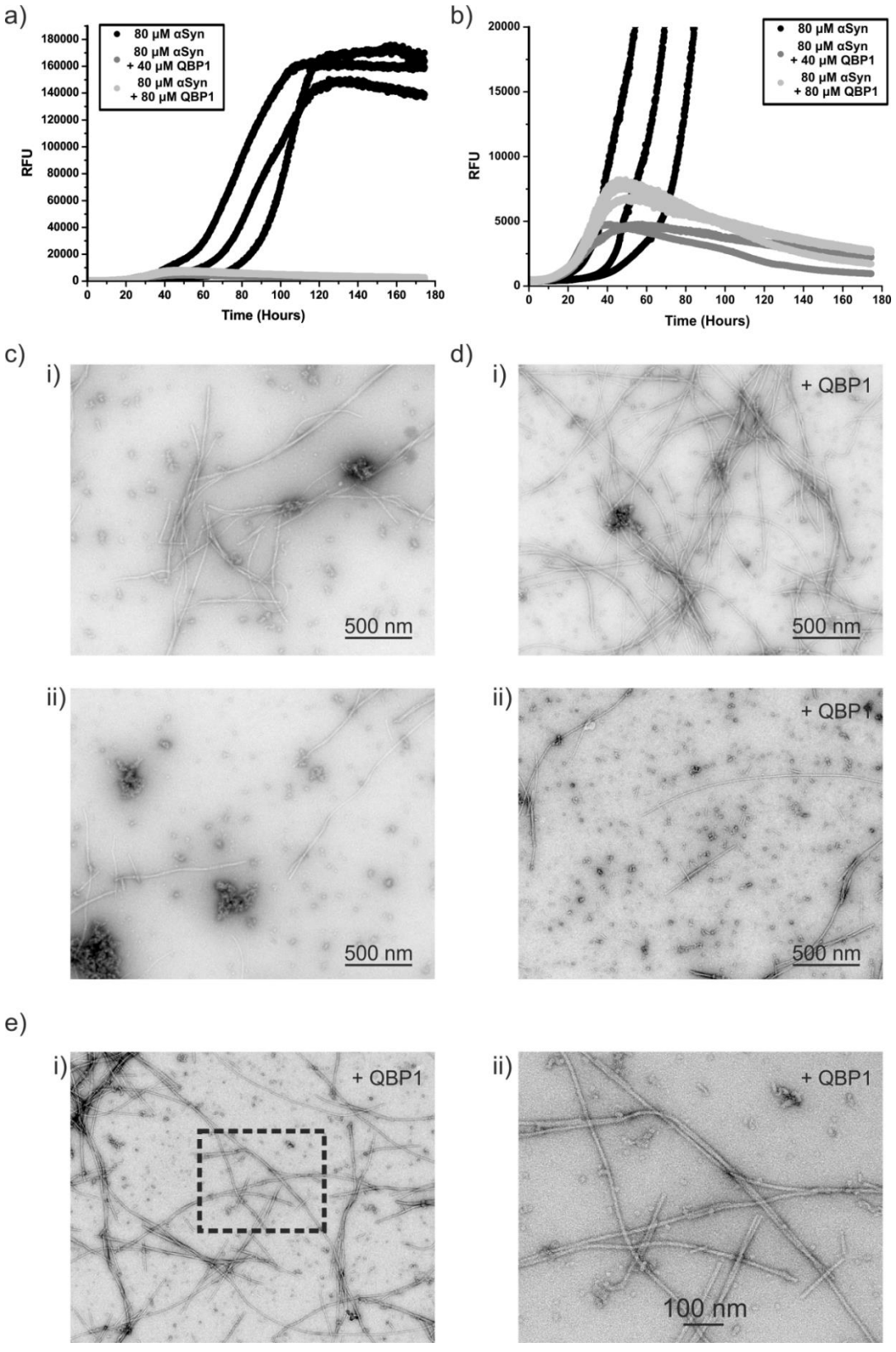

a)
80 μM αSyn
80 μM αSyn + 40 μM QBP1
80 μM αSyn + 80 μM QBP1
RFU
Time (Hours)
b)
80 μM αSyn
80 μM αSyn + 40 μM QBP1
80 μM αSyn + 80 μM QBP1
RFU
Time (Hours)
c)
i)
500 nm
ii)
500 nm
d)
i)
+ QBP1
500 nm
ii)
+ QBP1
500 nm
e)
i)
+ QBP1
ii)
+ QBP1
100 nm

Figure 4.25: QBP1 does not inhibit the formation of fibrils by α-synuclein. (a) The addition of equimolar (light grey) or sub-stoichiometric (dark grey) concentrations of QBP1 to α-synuclein reduces the observed ThT signal. However, rescaling of the axis (b) reveals that the lag time of aggregation is unchanged. TEM images of α-synuclein alone (c) and in the presence of sub-stoichiometric concentrations of QBP1 (d) both contain fibrils. (e) Observation of samples at increased magnification reveals the presence of several morphological species. The dashed box in e.i indicates the region magnified in e.ii. TEM samples were prepared at 180 hours. Samples prepared in 20mM Tris HCl, 200 mM NaCl pH 7.5 with 600 rpm shaking at 37 °C.

These data suggest that QBP1 does not inhibit the aggregation of α-synuclein at the concentrations used here. It is interesting to note that in these experiments QBP1 appeared to interfere with the ThT assay; however, the presence of QBP1 in the ataxin-3 aggregation assays did not cause a change in signal intensity (Figure 4.4). This may suggest that QBP1 interacts with α-Synuclein at a site that lies close to the ThT binding site or otherwise causes allosteric changes which alter ThT binding while in ataxin-3 the interaction appears at a site that does not alter ThT binding or fluorescence. In the case of ataxin-3 the ThT assay reports on the formation of polyQ independent protofibril formation so it is perhaps unsurprising that an inhibitor of the polyQ dependent stage should not interfere with this assay.

As Hervás *et al.* have reported evidence of an interaction between QBP1 and α-synuclein, native nESI-MS was used to investigate this possible interaction (Hervas *et al.* 2012).

In the native nESI-MS experiment α-synuclein forms at least two CSDs (Figure 4.26a). The first from $M+4H^{4+}$ to $M+9H^{9+}$ is taken to correspond to a compact conformation. A second CSD can be observed for $M+9H^{9+}$ to $M+17H^{17+}$. α-synuclein dimer is observed for many charge states in both CSDs indicating that α-synuclein readily forms dimers and that these species appear stable in MS conditions. Above 2000 m/z a number of low intensity peaks can be seen. These peaks correspond to dimer and trimer with the remaining peaks too broad for accurate mass determination. These unidentified peaks are thought to arise from other oligomeric species.

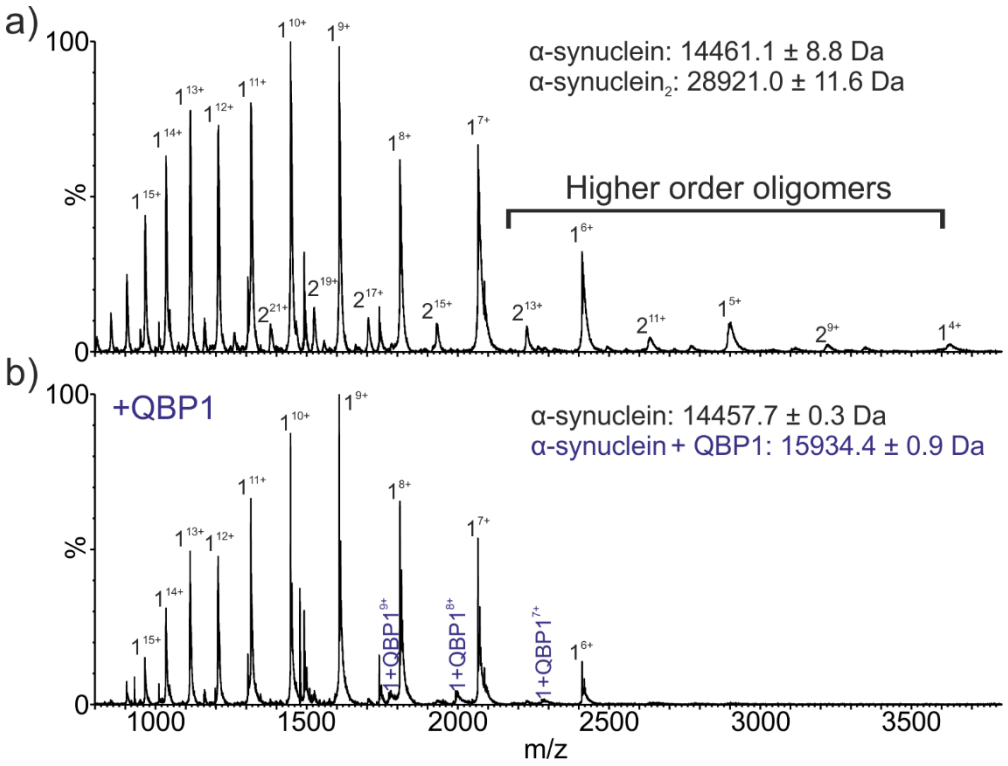


Figure 4.26: QBP1 forms a complex with the compact conformation of α-synuclein. nESI-MS of α-synuclein alone (a) reveals at least two CSDs. In addition dimer is observed throughout the spectrum while higher order oligomers are observed above 2000 m/z. The addition of QBP1 (b) results in the formation of low abundance ions for a 1:1 complex. These ions are observed only for the low z CSD. There is a reduction in the observation of higher order species in this spectrum however this may be related to signal to noise reduction in the presence of the peptide.

Upon the addition of QBP1 the formation of a 1:1 complex between α-synuclein and QBP1 is observed (Figure 4.26b). The peaks corresponding to this complex are low intensity and are observed only for the $M+4H^{4+}$ to $M+9H^{9+}$ CSD with complex 7+, 8+and 9+ observed. The complex $M+6H^{6+}$ may be too low intensity to observe. The lowest charge states of α-synuclein and the higher order species distribution are reduced in intensity in the samples with QBP1. This may be due to a change in the distribution of these species in solution; however, the ionisation efficiency α-synuclein was reduced in the presence of QBP1 and the reduction may be related to this phenomenon.

Given the observations in the experiments described above it seems likely that QBP1 does interact with α-synuclein. However, it appears that QBP1 does not inhibit aggregation as Hervás *et al.* hypothesised it might, based on their

observations of reductions in the amount of a proposed aggregation intermediate by atomic force microscopy (Hervas *et al.* 2012). This may be due to the experimental constraints on peptide concentration which prevent examination of the effect of an excess of QBP1 on α-synuclein aggregation. Alternatively, the interaction observed in the atomic force microscopy experiments (Hervas *et al.* 2012) and by MS here, may be an interaction that does not significantly modify the aggregation of α-synuclein.

4.11 Discussion

In this chapter has demonstrated that the QBP1 interacts with monomeric ataxin-3. The site of this interaction can be localised to residues 183-221 of ataxin-3, a surprising finding given the previous assumption that the polyQ domain is the site of action (Nagai *et al.* 2000). In addition, we show that ataxin-3 can interact with α-synuclein, a non-polyQ amyloid protein though this interaction appears to have no effect on the aggregation of the protein.

The native nESI-MS work presented in this chapter demonstrates that an interaction can occur between residues 183-221 of ataxin-3 and the peptide QBP1. The site of this interaction outside of the polyQ domain was unexpected as QBP1 was originally screened for interaction with polyQ fusion proteins and therefore no selective bias towards the unrelated sequence 183-221 would be expected (Nagai *et al.* 2000). Interestingly, subjecting ataxin-3 183-221 to a BlastP search against the human proteome reveals that this region has no significant homology to any other known protein in the SwissProt database. This finding suggests that residues 183-221 of ataxin-3 may form a unique epitope. Aligning the sequences of ataxin-3 182-221 with the glutathione-S-transferase Q62 construct against which QBP1 was originally selected shows there is no homology between the two sequences indicating that the interaction observed is not the result of aberrant (secondary) selection for sequences which bind to an epitope on GST.

While no previous publications indicate a role for the region 183-221 in the binding of QBP1, there has been interest in the role of the central flexible region (CFR, residues 183-291) in protein aggregation. Grandori and co-workers have shown, by means of the ThT assay and other methods, that the inclusion of the CFR in an ataxin-3 construct increases the rate of aggregation (Santambrogio *et al.* 2012). Interestingly, the co-incubation of ataxin-3 with a CFR peptide reduces the rate of ataxin-3 aggregation through what is believed to be competitive inhibition (Invernizzi *et al.* 2013). This second finding provides a precedent for binding to this region inhibiting aggregation of ataxin-3. However, the observed

change in polyQ-independent aggregation reported by Grandori and co-workers was not observed for the addition of QBP1 which inhibits the polyQ dependent stage only. Residues 183-221 are not predicted to have a high aggregation propensity either alone or as part of the CFR as estimated by TANGO and other amyloid prediction algorithms (Invernizzi *et al.* 2013; Santambrogio *et al.* 2012). However, AmylPred (Tsolis *et al.* 2013), a consensus prediction software, predicts that residues 175-184 and 195-200 of ataxin-3 may have some amyloidogenicity (Figure 4.27). It is interesting to note that all amyloid predictions algorithms used failed to detect the polyQ region as amyloidogenic.

Figure 4.27: Potential amyloidogenic regions of ataxin-3 identified by consensus prediction. The sequence of ataxin-3 78Q is shown. The consensus prediction is shown in the yellow bar. Residues which at least 5 of the 11 prediction algorithms scored as amyloidogenic are indicated with a #. The sequence was analysed using the Amylpred webserver (Tsolis *et al.* 2013). Full results can be found in Appendix III.

Previous work on the CFR has suggested that the UIMs may mediate aggregation in the full length ataxin-3 protein and that by interfering with this interaction the CFR peptide prevents aggregation (Invernizzi *et al.* 2013; Santambrogio *et al.* 2012). The MS data presented here show that QBP1 does not interact with the UIMs. Despite the lack of evidence of a direct interaction between QBP1 and the UIMs there may still be a role for the UIMs in the QBP1 mediated inhibition of aggregation. The first UIM1 lies immediately C-terminal to the binding site identified here and as such the binding of QBP1 to this region may disrupt an interaction with the neighbouring regions either through direct competition and steric hindrance or via an allosteric effect on conformation (a small scale change in conformation such as this may be invisible to IMS).

The finding that the QBP1 interaction is mediated by residues in the domains flanking the polyQ domain is consistent with the growing body of evidence that the flanking domains of amyloid proteins play an important role in modulating and controlling aggregation (Darnell *et al.* ; Lin *et al.* ; Masino *et al.* ; Saunders *et al.* ; Thual *et al.*). These flanking domains may represent a target for future therapeutic exploitation, especially in the case of polyQ proteins where they would provide a new method of ensuring specificity.

Investigation of the conformation of the various ataxin-3 constructs by IMS-MS reveals that the CCS of ataxin-3 monomer is unchanged upon binding QBP1. The resolution of IMS is limited, in particular in the case of largely disordered proteins such as ataxin-3 which populate a wide distribution of conformations. Some previously identified inhibitors of amyloid formation have been shown to alter the conformation (or conformational equilibrium) of the bound protein (Das *et al.* 2014; Kumar *et al.* 2016; Liu *et al.* 2011b; Young *et al.* 2015). The lack of conformational change upon binding shown here may be evidence for a directly competitive inhibition of aggregation. A further possibility is that QBP1 alters the conformation of higher order species that were not observed in these experiments. Previous work has hypothesised that QBP1 inhibits a key β-sheet transition which may occur in these higher order species (Nagai *et al.* 2007).

The interaction between ataxin-3 and the scrambled peptide, SP1 is interesting. The interaction observed appears similar to that observed for QBP1, with the formation of a 1:1 complex with compact forms for the monomer. These interactions are observed for the same ataxin-3 constructs as for QBP1 and once again the JD alone is not sufficient for interaction. This indicates that the interaction is likely to be mediated by the same binding site as for QBP1. However, when subjected to a ramp of trap voltages SP1 dissociates more easily that QBP1 indicating a lower affinity. The short length of the peptide and the frequent occurrence of the W residues which have been suggested to be key for inhibition may limit the variation possible between QBP1 and any scrambled form of the peptide. Thus the interaction being maintained to some degree in the scrambled form is not inconceivable. Ultimately, in order to full understand the interaction of SP1 with ataxin-3, a comprehensive set of mutations and truncations would be required which were beyond the scope and timescale of the current investigation. Despite this, the data presented here are in agreement with previous observations that SP1 does not inhibit the formation of fibrils (Nagai *et al.* 2000; Ren *et al.* 2001; Tomita *et al.* 2009).

In summary, the results presented in this chapter demonstrate that QBP1 interacts with ataxin-3 monomer at a novel site formed by residues 183-221. This interaction is maintained in ataxin-3 constructs which lack polyQ domain. Binding at this site does not alter the conformation of ataxin-3 monomer however comment cannot be made on the effect this interaction may have on higher order species. Given the suggested importance of the CFR in ataxin-3 aggregation (Invernizzi *et al.* 2013; Santambrogio *et al.* 2012) and the observation of QBP1 binding in this region future work should aim to more thoroughly address the function of the flanking domains in aggregation and the residues responsible for QBP1 binding, with the aim to exploit them as therapeutic targets.

5 Concluding remarks and future directions

5.1 Conclusions

Amyloid protein aggregation disorders are some of the most prevalent diseases in the developed world. Elucidating the mechanisms of amyloid aggregation modulators (inhibitors and accelerators) is vital in advancing our knowledge of protein aggregation and our ability to rationally design therapeutics to prevent amyloid diseases. Determining the site of interactions which alter the aggregation of amyloidogenic proteins is one step in this process. This thesis describes works towards characterising the kinetic effects and interactions of two modulators of amyloid aggregation, the peptide QBP1 and the protein YDL085CA.

Chapter 3 explores the interaction between YDL085CA and Aβ40. While orthologous proteins from other species have been shown to accelerate amyloid formation, YDL085CA was observed to inhibit amyloid formation under similar conditions. However, the effect seen was extremely buffer dependent as when YDL085CA and Aβ40 were incubated in a second buffer YDL085CA was observed to enhance aggregation, suggesting a role for ionic strength and specific ion effects in YDL085CA activity. Crosslinking followed by mass spectrometric analysis was used to map the binding interface of YDL085CA and Aβ40, highlighting the involvement of two helical regions of YDL085CA and the charged N-terminus of Aβ40.

The finding that YDL085CA shows such a strong buffer dependence in its activity highlights the importance of considering characteristics such as ionic strength and the Hofmeister effect (Kunz, Henle and Ninham 2004; Hofmeister 1888) when interrogating the structure and function of intrinsically disordered proteins and peptides (Liao *et al.* 2013; Wicky, Shammas and Clarke 2017). Had the analysis been conducted under a single set of conditions, a potentially important facet of the mechanism and function of YLD085CA may have been missed. The observed differences between YDL085CA and the activity of SERF1 in relation to Aβ40 aggregation may, in part, be explained by specific ion effects. However, given the similarity in the buffers employed and the evidence of an interaction

between Aβ40 and YDL085CA outside of conserved regions, it seems likely that the two proteins may have evolved different functions. This demonstrates the dangers of assigning functions to proteins based on sequence similarity alone.

Previously, QBP1 had been assumed to interact solely with the expanded polyQ domains of polyQ proteins and thus to inhibit their aggregation. While evidence exists for an interaction between QBP1 and a non-polyQ containing amyloid protein (Hervas *et al.* 2012), if and how this might inhibit aggregation had not been investigated nor had its relevance to polyQ proteins been considered. The work in Chapter 4 demonstrates, by means of native MS and supporting biophysical and biochemical techniques, that an interaction between QBP1 and the polyQ protein ataxin-3 is mediated by a region outside of the polyQ domain and persists in the absence of an expanded polyQ tract.

The interaction between QBP1 and ataxin-3 at a non-polyQ site was unexpected. This observation highlights the importance of studies on full length, disease causing proteins, rather than fusion proteins, when developing inhibitory molecules. The identification of this region as potentially important in the aggregation of ataxin-3 is not entirely without precedent (Invernizzi *et al.* 2013; Santambrogio *et al.* 2012). However, previous studies have attributed the aggregation modulating effects of this region to the UIMs.

These results together demonstrate that interactions distal to the sites traditionally associated with aggregation initiation can produce a powerful effect on the aggregation of an amyloid protein. These regions may represent alternative targets for therapeutic exploitation.

The data presented here also highlight the varied range of MS techniques which can be applied to the study of protein structure and inteactions. The diversity of techniques available in the field of MS has enabled the study of interactions which were extremely challenging to analyse initially (for example the difficulty of investigating the Aβ40:YDL085CA interaction by native MS was overcome by application of in-solution crosslinking techniques).

5.2 Future directions

Our understanding of the mechanisms of action of amyloid aggregation modulators still needs expanding, both in the context of the molecules described here and in a more general sense. Future investigations relating to the YLD085CA:Aβ40 interaction should focus on two broad lines of questioning: (1) how and why are SERF1 and YDL085CA different and (2) what effect do specific ions have on the interactions observed?

The first question may be approached by direct comparison of SERF1 and YDL085CA with regard to both their effect on the kinetics of amyloid protein aggregation and the sites of interaction between these proteins and amyloid proteins. In particular, comparing the involvement of the well-conserved regions and areas of greater sequence diversity in forming interactions with amyloid proteins and the relationship of the latter to activity could be of great interest.

Answering the second question would involve analysis of the effect of YDL085CA on amyloid protein aggregation in a wider range of buffers selected from different regions of the Hofmeister series. Both the effect on the activity of YDL085CA with relation to the aggregation kinetics of amyloid proteins and also the conformations of YDL085CA and Aβ40 may be affected, with the latter being likely to produce an influence on the former.

As described in Chapter 4, a scrambled sequence derived from the QBP1 peptide (SP1) does not prevent fibril formation by ataxin-3 but does bind to ataxin-3 in a similar manner to QBP1. A full study of this binding involving mutations of the peptide, with the aim of abrogating binding, would provide interesting insights into how the observed binding is mediated. In addition, investigation of the interactions of QBP1/SP1 with other polyQ proteins would determine whether the non-polyQ mediated interaction is a common occurrence.

More generally, it seems that there may be some value in revisiting previously published inhibitors of amyloid formation, in particular those whose mechanisms have not yet been elucidated, and investigating their effect in the context of full

length, endogenous amyloid proteins. While model systems such as fusion proteins and short peptides can be extremely useful, the use of such systems may be excluding inhibitors which target regions away from the amyloid core-forming sequence. These distal inhibitors may provide greater specificity in their effect than those targeting the amyloid core, which often produce their effect by non-specific binding (for example polyphenol compounds (Cao and Raleigh 2012; Ehrnhoefer *et al.* 2008; Palhano *et al.* 2013)). Increased specificity may in turn lead to reduced off target effects and increased efficacy.

Appendix I:

Appendix I contains figures and information relating to Chapter 2, which is concerned with the methods and materials employed in the investigations described in this thesis.

Appendix I.I Details of plasmids:

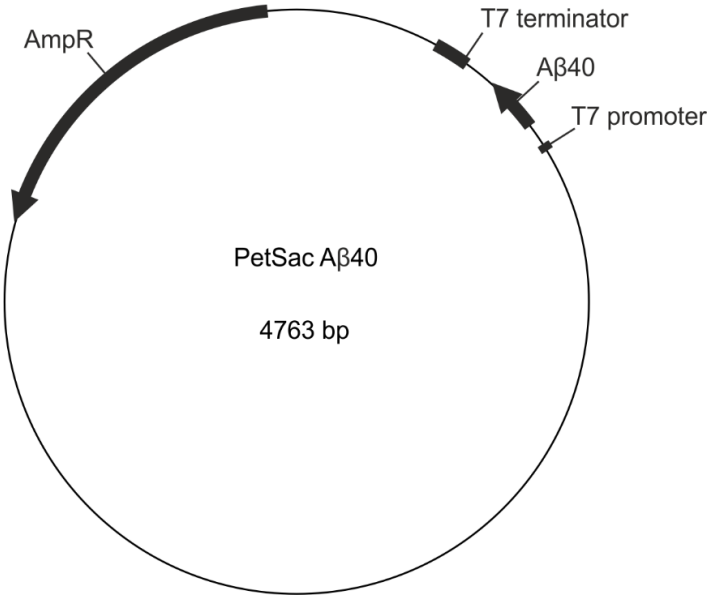


Figure AI.1: Plasmid map of PetSac Aβ40.

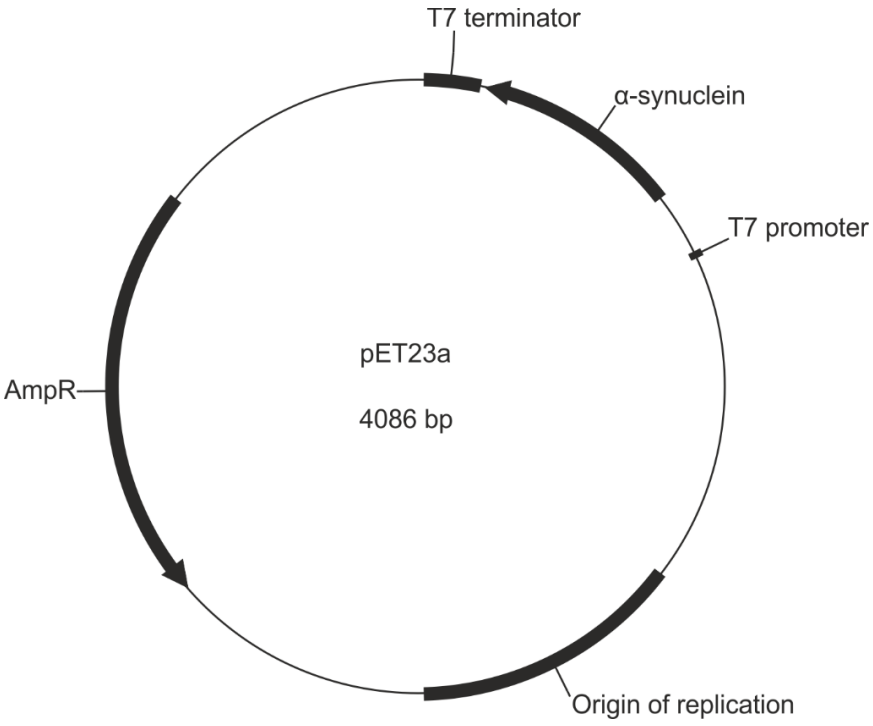


Figure AI.2: Plasmid map of pET23a α-synuclein.

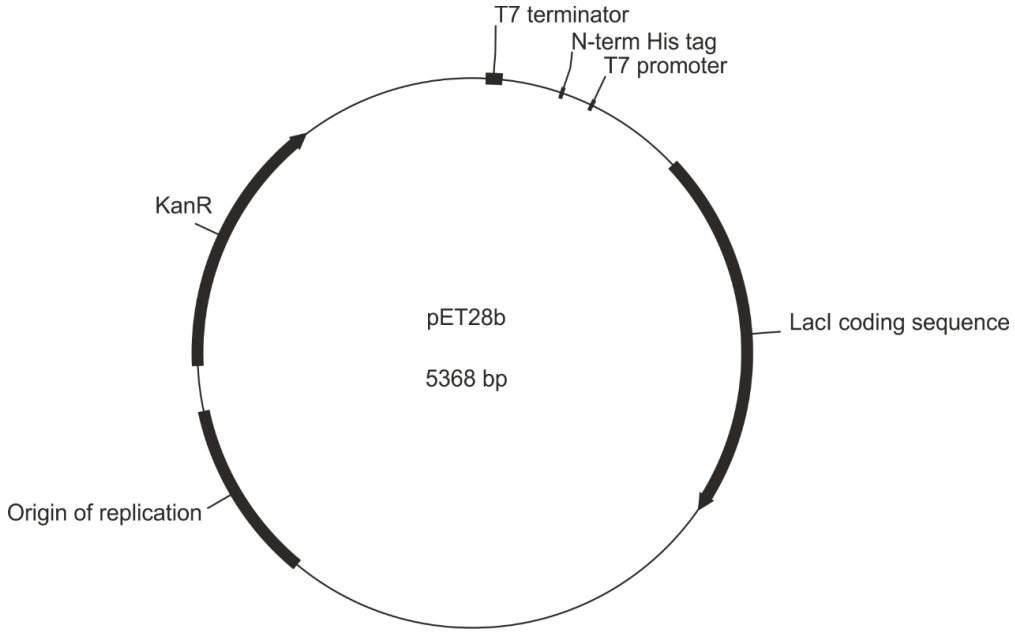


Figure AI.3: Plasmid map of pET28b used in the expression of the various YDL085CA proteins.

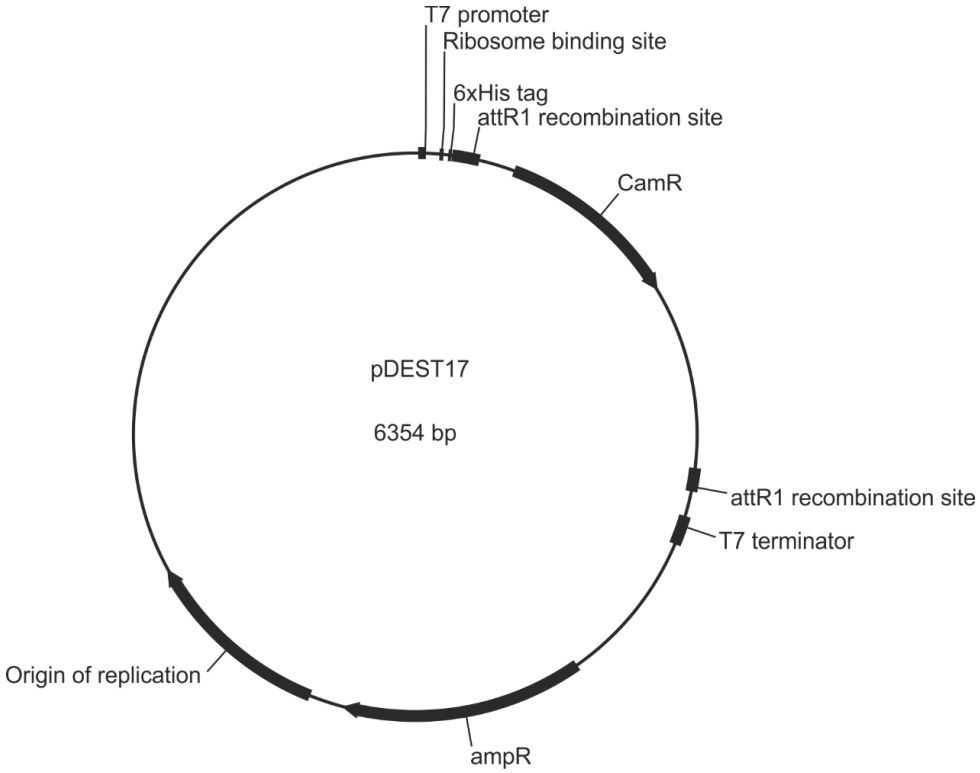


Figure AI.4: Plasmid map of pDEST17 used in the expression of the various ataxin-3 proteins. Note that the CamR gene is removed upon successful recombination.

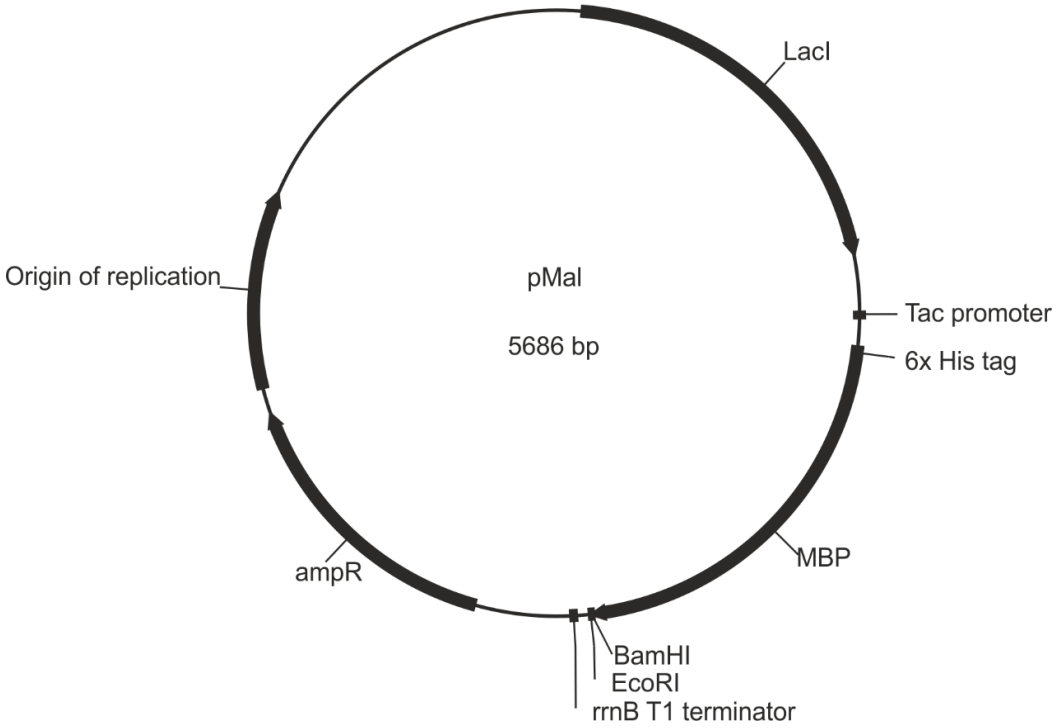


Figure AI.5: Plasmid map of pMAL Cx5 used in the expression of the MBP fusion protein. The initial plasmid was a gift from Dr David Brockwell (University of Leeds) and contains a N-terminal His tag and replaces the thrombin cleavage site at the C-terminus of MBP with a TEVp cleavage site.

Appendix I.2 DNA sequences:

Table AI.1: DNA sequences for the proteins used in this investigation. Note that sequences labelled with a * are part of a larger translated sequence which is subsequently processed.

Protein	Sequence	Chapter
mAβ40	ATGGACGCTGAATTCCGTCACGACTCTGGTTACGAAGT TCACCACCAGAAGCTGGTGTTCTTCGCTGAAGACGTGG GTTCTAACAAGGGTGCTATCATCGGTCTGATGGTTGGT GGCGTTGTG	3
α-synuclein	ATGGATGTCTTCATGAAAGGGCTGTCGAAAGCGAAAG AAGGTGTCGTAGCAGCTGCGGAAAAGACCAAACAAGG CGTAGCCGAAGCAGCCGGTAAAACGAAGGAAGGTGTG CTGTATGTCGGCAGCAAGACCAAAGAAGGCGTTGTCCA TGGGGTTGCGACTGTTGCGGAGAAAACGAAAGAGCAG GTGACAAACGTGGGAGGAGCCGTTGTGACCGGTGTGA CCGCAGTAGCCCAGAAAACGGTGGAAGGTGCTGGCTC CATTGCAGCGGCGACTGGCTTTGTGAAGAAAGACCAGT TAGGGAAAAACGAGGAAGGTGCTCCTCAAGAAGGCAT CTTGGAGGATATGCCGGTTGATCCGGATAATGAGGCCT ATGAGATGCCGAGTGAAGAAGGCTACCAGGACTACGA ACCAGAAGCG	3,4
YDL085CA(S)*	ATGGCTAGAGGTAATCAAAGAGACTTGGCAAGACAAA AAAACTTGAAAAAACAAAAGGACATGGCTAAGAACCA GAAAAAGAGTGGTGATCCTAAGAAAAGAATGGAGTCA GACGCCGAAATCTTGAGACAAAAGCAGGCCGCCGCAG ATGCTAGAAGAGAGGCTGAAAAGCTTGAGAAGTTAAA AGCCGAAAAGACGAGAAGATAA	3

YDL085CA*	ATGGCGCGCGGTAACCAGCGTGACCTGGCCCGTCAGA AAAATCTGAAAAAACAAAAAGATATGGCGAAGAATCA AAAAAAATCTGGTGATCCAAAAAAACGCATGGAGTCCG ACGCAGAAATTCTGCGCCAGAAACAGGCCGCCGCCGA GCACGTCGCGAAGCAGAAAAGCTGGAAAAACTGAAAG CGGAGAAAACCCGCCGTTAA	3
Ataxin-3 78Q	ATGTCGTACTACCATCACCATCACCATCACCTCGAAAAC CTGTATTTTCAGGGCATGGAGTCCATCTTCCACGAGAA ACAAGAAGGCTCACTTTGTGCTCAACATTGCCTGAATA ACTTATTGCAAGGAGAATATTTTAGCCCTGTGGAATTAT CCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGAT GAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTAT CGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGA CAGTGGTTTTTTCTCTATTCAGGTTATAAGCAATGCCTT GAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTC CAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAA AGATCATTTATATGCAATTATAAGGAACACTGGTTTACA GTTAGAAAATTAGGAAAACAGTGGTTTAACTTGAATTC TCTCTTGACGGGTCCAGAATTAATATCAGATACATATCT TGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTC TATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAG CTGACCAACTCCTGCAGATGATTAGGGTCCAACAGATG CATCGACCAAAACTGATTGGCGAAGAACTGGCTCAACT GAAAGAACAGCGTGTGCATAAGACCGACCTGGAACGT GTCCTGGAAGCAAATGACGGCAGCGGCATGCTGGATG AAGACGAAGAAGATCTGCAGCGTGCCCTGGCACTGTCT CGTCAGGAAATTGATATGGAAGACGAAGAAGCAGATC TGCGTCGCGCTATTCAGCTGTCAATGCAGGGCAGCTCT CGTAACATCTCGCAGGACATGACCCAGACGAGCGGTAC CAATCTGACGTCTGAAGAACTGCGCAAACGTCGCGAAG CATATTTTGAAAAACAGCAACAGAAGCAACAACAGCAG CAACAACAACAACAACAGCAGCAGCAGCAGCAACAAC AACAGCAACAACAACAGCAACAGCAGCAACAGCAACA GCAACAGCAGCAACAACAGCAACAGCAACAACAGCAA CAGCAACAACAGCAACAACAGCAACAACAGCAACAAC AGCAACAACAGCAACAACAGCAACAGCAACAACAGCA GCAGCAACAGCAGCAACAACAACAGAGAGATCTGTCA GGCCAGAGTTCCCATCCGTGTGAACGTCCGGCCACCTC AAGCGGTGCACTGGGTAGTGATCTGGGTGACGCCATG TCCGAAGAAGACATGCTGCAGGCAGCAGTGACGATGT CCCTGGAAACCGTGCGTAACGACCTGAAAACCGAAGG CAAAAAATAA	4

Ataxin-3 14Q	ATGTCGTACTACCATCACCATCACCATCACCTCGAAAAC CTGTATTTTCAGGGCATGGAGTCCATCTTCCACGAGAA ACAAGAAGGCTCACTTTGTGCTCAACATTGCCTGAATA ACTTATTGCAAGGAGAATATTTTAGCCCTGTGGAATTAT CCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGAT GAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTAT CGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGA CAGTGGTTTTTTCTCTATTCAGGTTATAAGCAATGCCTT GAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTC CAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAA AGATCATTTATATGCAATTATAAGGAACACTGGTTTACA GTTAGAAAATTAGGAAAACAGTGGTTTAACTTGAATTC TCTCTTGACGGGTCCAGAATTAATATCAGATACATATCT TGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTC TATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAG CTGACCAACTCCTACAGATGATTAGGGTCCAACAGATG CATCGACCAAAACTTATTGGAGAAGAATTAGCACAACT AAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACGA GTGTTAGAAGCAAATGATGGCTCAGGAATGTTAGACG AAGATGAGGAGGATTTGCAGAGGGCTCTGGCACTAAG TCGCCAAGAAATTGACATGGAAGATGAGGAAGCAGAT CTCCGCAGGGCTATTCAGCTAAGTATGCAAGGTAGTTC CAGAAACATATCTCAAGATATGACACAGACATCAGGTA CAAATCTTACTTCAGAAGAGCTTCGGAAGAGACGAGAA GCCTACTTTGAAAAACAGCAGCAAAAGCAGCAACAGCA GCAGCAGCAGCAGCAGCAGGGGGACCTATCAGGACAG AGTTCACATCCATGTGAAAGGCCAGCCACCAGTTCAGG AGCACTTGGGAGTGATCTAGGTGATGCTATGAGTGAA GAAGACATGCTTCAGGCAGCTGTGACCATGTCTTTAGA AACTGTCAGAAATGATTTGAAAACAGAAGGAAAAAAA TAAG	4

JDU1	ATGTCGTACTACCATCACCATCACCATCACCTCGAAAAC CTGTATTTTCAGGGCATGGAGTCCATCTTCCACGAGAA ACAAGAAGGCTCACTTTGTGCTCAACATTGCCTGAATA ACTTATTGCAAGGAGAATATTTTAGCCCTGTGGAATTAT CCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGAT GAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTAT CGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGA CAGTGGTTTTTTCTCTATTCAGGTTATAAGCAATGCCTT GAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTC CAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAA AGATCATTTATATGCAATTATAAGGAACACTGGTTTACA GTTAGAAAATTAGGAAAACAGTGGTTTAACTTGAATTC TCTCTTGACGGGTCCAGAATTAATATCAGATACATATCT TGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTC TATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAG CTGACCAACTCCTACAGATGATTAGGGTCCAACAGATG CATCGACCAAAACTTATTGGAGAAGAATTAGCACAACT AAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACGA GTGTTAGAAGCAAATGATGGCTCAGGAATGTTAGACG AAGATGAGGAGGATTTGCAGAGGGCTCTGGCACTAAG TCGCCAAGAAATTGACTGA	4
JD+	ATGTCGTACTACCATCACCATCACCATCACCTCGAAAAC CTGTATTTTCAGGGCATGGAGTCCATCTTCCACGAGAA ACAAGAAGGCTCACTTTGTGCTCAACATTGCCTGAATA ACTTATTGCAAGGAGAATATTTTAGCCCTGTGGAATTAT CCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGAT GAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTAT CGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGA CAGTGGTTTTTTCTCTATTCAGGTTATAAGCAATGCCTT GAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTC CAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAA AGATCATTTATATGCAATTATAAGGAACACTGGTTTACA GTTAGAAAATTAGGAAAACAGTGGTTTAACTTGAATTC TCTCTTGACGGGTCCAGAATTAATATCAGATACATATCT TGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTC TATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAG CTGACCAACTCCTACAGATGATTAGGGTCCAACAGATG CATCGACCAAAACTTATTGGAGAAGAATTAGCACAACT AAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACGA GTGTTAGAAGCAAATGATGGCTCAGGAATG	4

JD	ATGTCGTACTACCATCACCATCACCATCACCTCGAAAAC CTGTATTTTCAGGGCATGGAGTCCATCTTCCACGAGAA ACAAGAAGGCTCACTTTGTGCTCAACATTGCCTGAATA ACTTATTGCAAGGAGAATATTTTAGCCCTGTGGAATTAT CCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGAT GAGAATGGCAGAAGGAGGAGTTACTAGTGAAGATTAT CGCACGTTTTTACAGCAGCCTTCTGGAAATATGGATGA CAGTGGTTTTTTCTCTATTCAGGTTATAAGCAATGCCTT GAAAGTTTGGGGTTTAGAACTAATCCTGTTCAACAGTC CAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAA AGATCATTTATATGCAATTATAAGGAACACTGGTTTACA GTTAGAAAATTAGGAAAACAGTGGTTTAACTTGAATTC TCTCTTGACGGGTCCAGAATTAATATCAGATACATATCT TGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTC TATATTTGTCGTTAAGGGTGATCTGCCAGATTGCGAAG CTGACCAACTCCTACAGATGATTAGG	4

MBP+	ATGCATCACCATCACCATCACAAAATCGAAGAAGGTAA ACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACG GTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACC GGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGA AGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGGC CCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGC TACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGA CAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGG ATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCG ATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGAT CTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCC GGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGC GCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTG GCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGT ATGAAAACGGCAAGTACGACATTAAAGACGTGGGCGT GGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTG GTTGACCTGATTAAAAACAAACACATGAATGCAGACAC CGATTACTCCATCGCAGAAGCTGCCTTTAATAAAGGCG AAACAGCGATGACCATCAACGGCCCGTGGGCATGGTC CAACATCGACACCAGCAAAGTGAATTATGGTGTAACGG TACTGCCGACCTTCAAGGGTCAACCATCCAAACCGTTC GTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCC GAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATC TGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGA CAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAG GAAGAGTTGGTGAAAGATCCGCGTATTGCCGCCACTAT GGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATC CCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGC GGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATG AAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCGAAC AACAACAACAATAACAATAACAACAACCTCGGGATCGA GGGAAGGATTTCACATATGGAAAACCTGTACTTTCAGG GATCCGTCCAACAGATGCATCGACCAAAACTTATTGGA GAAGAATTAGCACAACTAAAAGAGCAAAGAGTCCATA AAACAGACCTGGAACGAGTGTTAGAAGCAAATGATGG CTCAGGAATGTAA	4

Appendix I.3 Protein sequences:

Table AI.2: Protein sequences

Protein	Sequence	Chapter
mAβ40	MDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV	3
Aβ40	DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV	3
α-synuclein	MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEPEA	3,4
YDL085CA(S)	SMARGNQRDLARQKNLKKQKDMAKNQKKSGDPKKRMESDAEILRQKQAAADARREAEKLEKLKAEKTRR	3
YDL085CA	MARGNQRDLARQKNLKKQKDMAKNQKKSGDPKKRMESDAEILRQKQAAADARREAEKLEKLKAEKTRR	3
Ataxin-3 78Q	SYYHHHHHHLENLYFQGMESIFHEKQEGSLCAQHCLNNLLQGEYFSPVELSSIAHQLDEEERMRMAEGGVTSEDYRTFLQQPSGNMDDSGFFSIQVISNALKVWGLELILFNSPEYQRLRIDPINERSFICNYKEHWFTVRKLGKQWFNLNSLLTGPELISDTYLALFLAQLQQEGYSIFVVKGDLPDCEADQLLQMIRVQQMHRPKLIGEELAQLKEQRVHKTDLERVLEANDGSGMLDEDEEDLQRALALSRQEIDMEDEEADLRRAIQLSMQGSSRNISQDMTQTSGTNLTSEELRKRREAYFEKQQQKQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQRDLSGQSSHPCERPATSSGALGSDLGDAMSEEDMLQAAVTMSLETVRNDLKTEGKK	4
Ataxin-3 14Q	SYYHHHHHHLENLYFQGMESIFHEKQEGSLCAQHCLNNLLQGEYFSPVELSSIAHQLDEEERMRMAEGGVTSEDYRTFLQQPSGNMDDSGFFSIQVISNALKVWGLELILFNSPEYQRLRIDPINERSFICNYKEHWFTVRKLGKQWFNLNSLLTGPELISDTYLALFLAQLQQEGYSIFVVKGDLPDCEADQLLQMIRVQQMHRPKLIGEELAQLKEQRVHKTDLERVLEANDGSGMLDEDEEDLQRALALSRQEIDMEDEEADLRRAIQLSMQGSSRNISQDMTQTSGTNLTSEELRKRREAYFEKQQQKQQQQQQQQQQGDLSGQSSHPCERPATSSGALGSDLGDAMSEEDMLQAAVTMSLETVRNDLKTEGKK	4
JDU1	SYYHHHHHHLENLYFQGMESIFHEKQEGSLCAQHCLNNLLQGEYFSPVELSSIAHQLDEEERMRMAEGGVTSEDYRTFLQQPSGNMDDSGFFSIQVISNALKVWGLELILFNSPEYQRLRIDPINERSFICNYKEHWFTVRKLGKQWFNLNSLLTGPELISDTYLALFLAQLQQEGYSIFVVKGDLPDCEADQLLQMIRVQQMHRPKLIGEELAQLKEQRVHKTDLERVLEANDGSGMLDEDEEDLQRALALSRQEID	4

JD+	SYYHHHHHHLENLYFQGMESIFHEKQEGSLCAQHCLNNL LQGEYFSPVELSSIAHQLDEEERMRMAEGGVTSEDYRTFL QQPSGNMDDSGFFSIQVISNALKVWGLELILFNSPEYQRL RIDPINERSFICNYKEHWFTVRKLGKQWFNLNSLLTGPELI SDTYLALFLAQLQQEGYSIFVVKGDLPDCEADQLLQMIRV QQMHRPKLIGEELAQLKEQRVHKTDLERVLEANDGSGM	4
JD	SYYHHHHHHLENLYFQGMESIFHEKQEGSLCAQHCLNNL LQGEYFSPVELSSIAHQLDEEERMRMAEGGVTSEDYRTFL QQPSGNMDDSGFFSIQVISNALKVWGLELILFNSPEYQRL RIDPINERSFICNYKEHWFTVRKLGKQWFNLNSLLTGPELI SDTYLALFLAQLQQEGYSIFVVKGDLPDCEADQLLQMIR	4
UIM1	LDEDEEDLQRALALSRQEID	4
UIM12	LDEDEEDLQRALALSRQEIDMEDEEADLRRAIQLSMQGSS RNI	4
MBP+	MHHHHHHKIEEGKLVIWINGDKGYNGLAEVGKKFEKDT GIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGY AQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVE ALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQ EPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAG LTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPW AWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAAS PNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEEL VKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVIN AASGRQTVDEALKDAQTNSSSNNNNNNNNNNLGIEGRI SHMENLYFQGSVQQMHRPKLIGEELAQLKEQRVHKTDLE RVLEANDGSGM	4
MBP+ (post-cleavage)	MHHHHHHKIEEGKLVIWINGDKGYNGLAEVGKKFEKDT GIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGY AQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVE ALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQ EPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAG LTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPW AWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAAS PNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEEL VKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVIN AASGRQTVDEALKDAQTNSSSNNNNNNNNNNLGIEGRI SHMENLYFQ	4
QBP1	SNWKWWPGIFD	4
SP1	WPIWSKGNDWF	4
SP1G	WPIWSKGNDWFG	4

Table AI.3: Predicted masses of the proteins and peptides used in this investigation.

Protein	Monoisotopic mass (Da)	Average mass (Da)	Chapter
mAβ40	4458.19	4461.06	3
Aβ40	4327.15	4329.86	3
α-synuclein	14451.22	14460.19	3,4
YDL085CA(S)	8031.36	8036.28	3
YDL085CA	7944.32	7949.20	3
Ataxin-3 78Q	51718.94	51751.02	4
Ataxin-3 14Q	43424.11	43451.51	4
JDU1	30025.77	30044.85	4
JD+	27686.65	27704.36	4
JD	23222.34	23237.25	4
UIM1	2357.13	2358.50	4
UIM12	4987.39	4990.43	4
MBP+	49176.87	49207.64	4
MBP+ (post-cleavage)	44596.39	44568.51	4
QBP1	1476.68	1477.64	4
SP1	1476.68	1477.64	4
SP1G	1533.70	1534.69	4

Appendix I.4 Further examples protein of purifications:

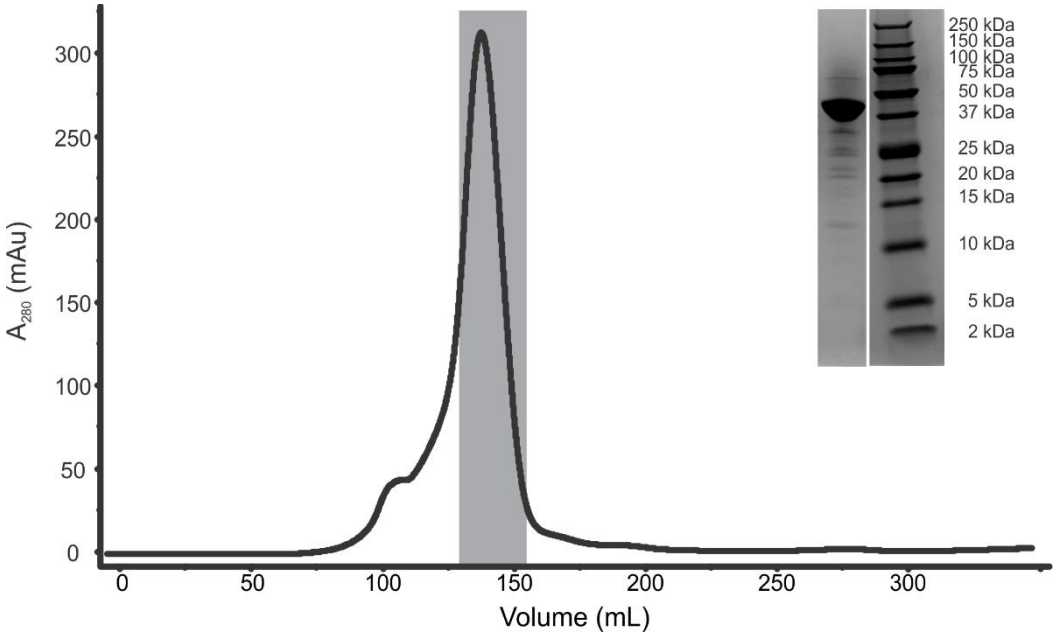


Figure AI.6: SEC elution profile for ataxin-3 14Q. The A280 trace for the SEC of ataxin-3 14Q is shown. The grey box indicates fractions collected as monomeric protein. Inset: SDS-PAGE of the collected fractions showing a band at approximately 43.4 kDa.

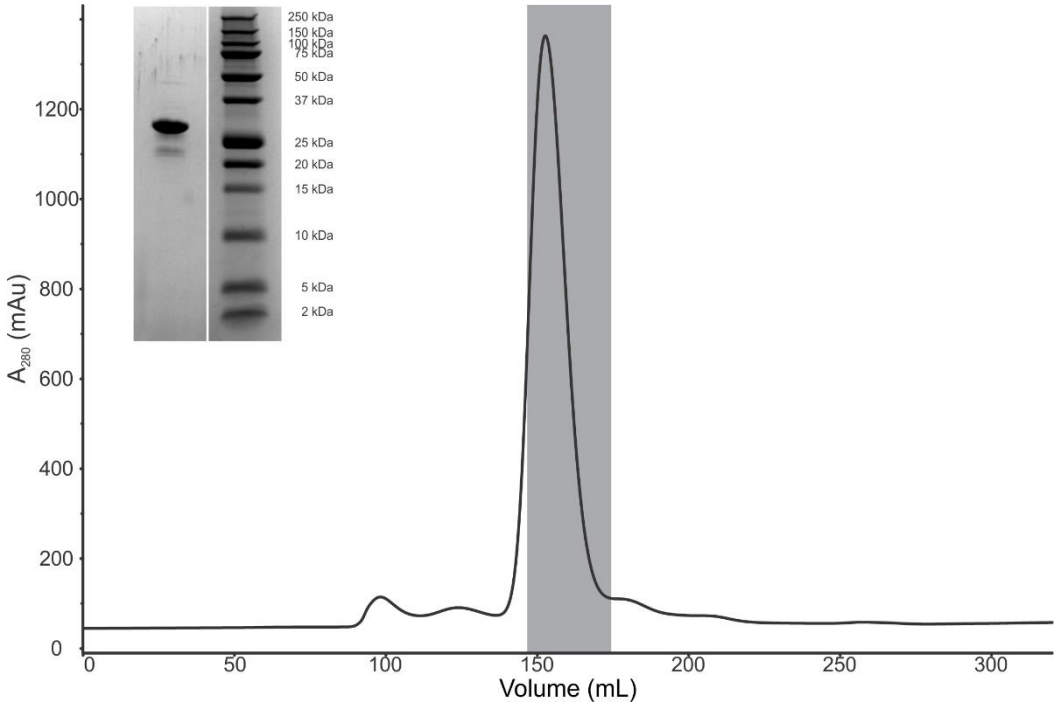


Figure AI.7: SEC elution profile for JDU1. The A280 trace for the SEC of JDU1 is shown. The grey box indicates fractions collected as monomeric protein. Inset: SDS-PAGE of the collected fractions showing a band at approximately 30.0 kDa.

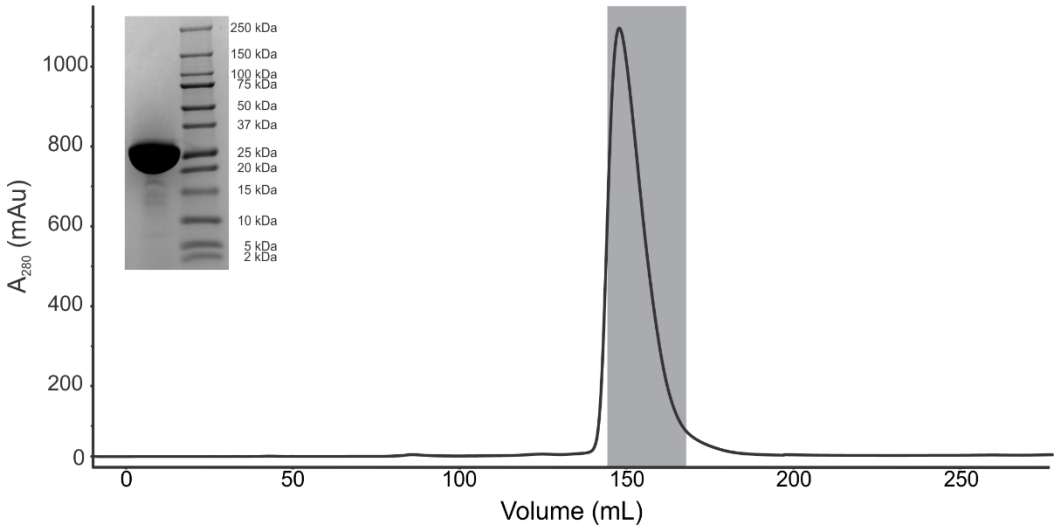


Figure AI.8: SEC elution profile for JD+. The A280 trace for the SEC of JD+ is shown. The grey box indicates fractions collected as monomeric protein. Inset: SDS-PAGE of the collected fractions showing a band at approximately 27.8 kDa.

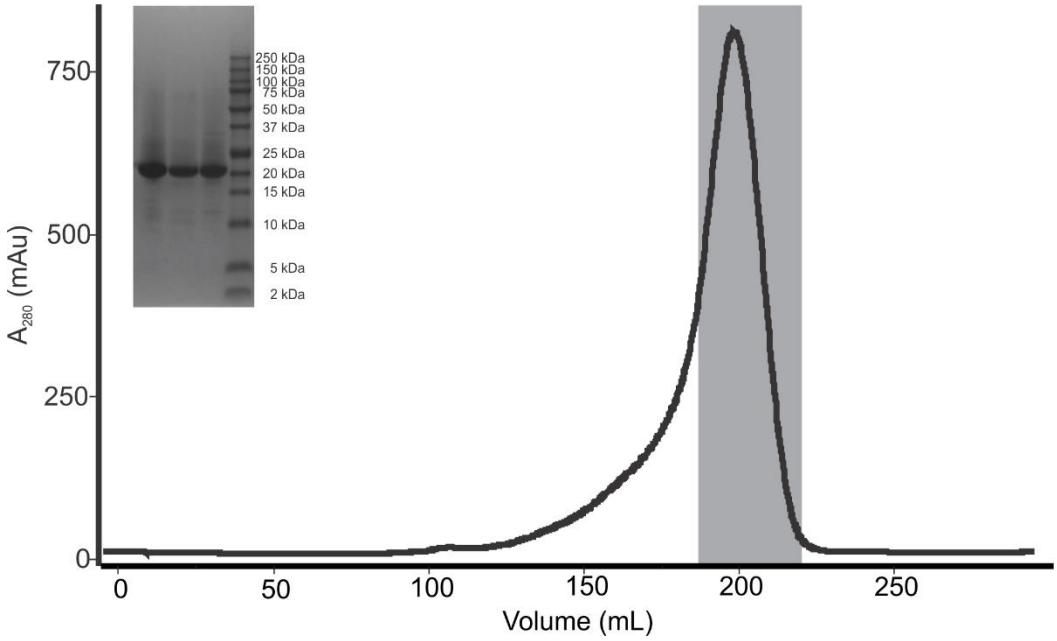


Figure AI.9: SEC elution profile for JD. The A280 trace for the SEC of JD is shown. The grey box indicates fractions collected as monomeric protein. Inset: SDS-PAGE of the collected fractions showing a band at approximately 23.2 kDa.

Appendix II:

Appendix II contains figures and information relating to Chapter 3, which is concerned with the investigation of the interaction between the YDL085CA and Aβ40 and the wider context of the effect of YDL085CA on amyloid protein aggregation.

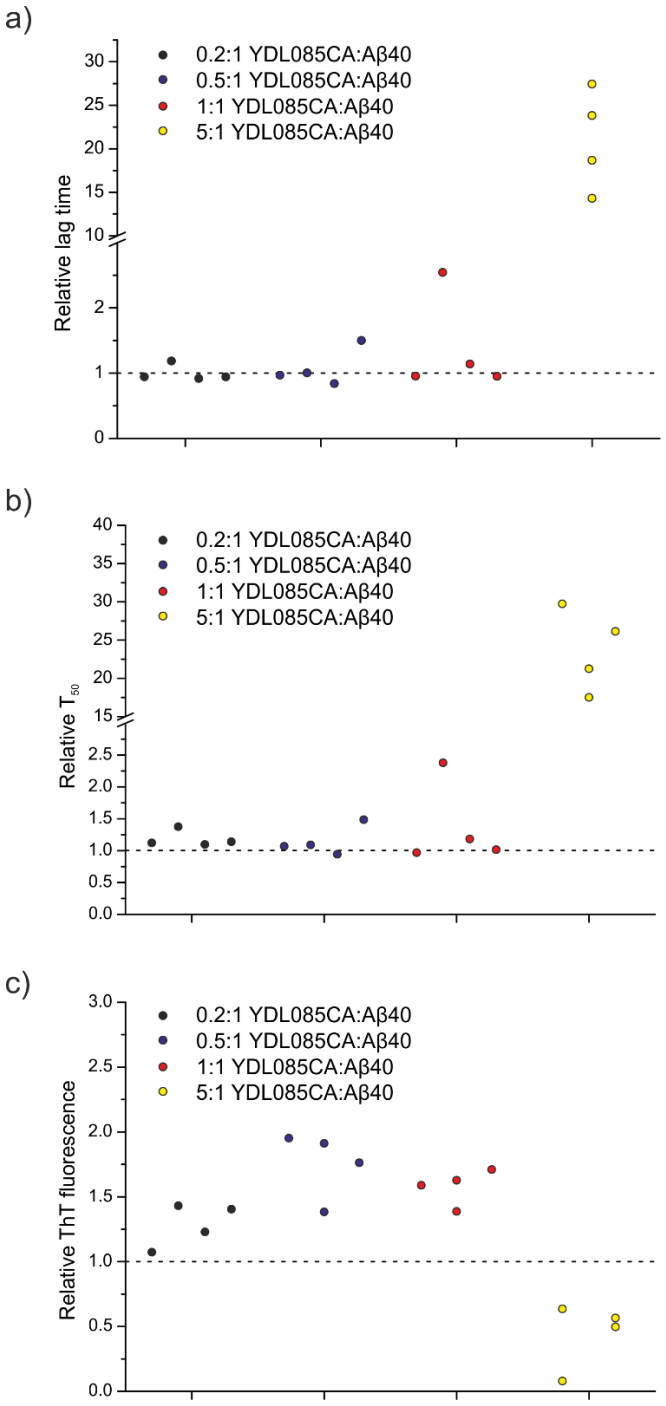


Figure A.II.1: Analysis of varying Aβ40 concentration upon Aβ40 aggregation in the presence of YDL085CA. (a) The lag time of aggregation, relative to Aβ40 alone, for four different concentrations of Aβ40 in the presence of 25 μM YDL085CA. (b) The T50 of Aβ40 aggregation for the same data set. (c) The relative maximum fluorescence value for the same data set. Note that the 5:1 samples had not reached completion (see Figure 3.12) and therefore the values may be taken to indicate an increase in T50 and a reduction in maximum fluorescence rather than as absolute values.

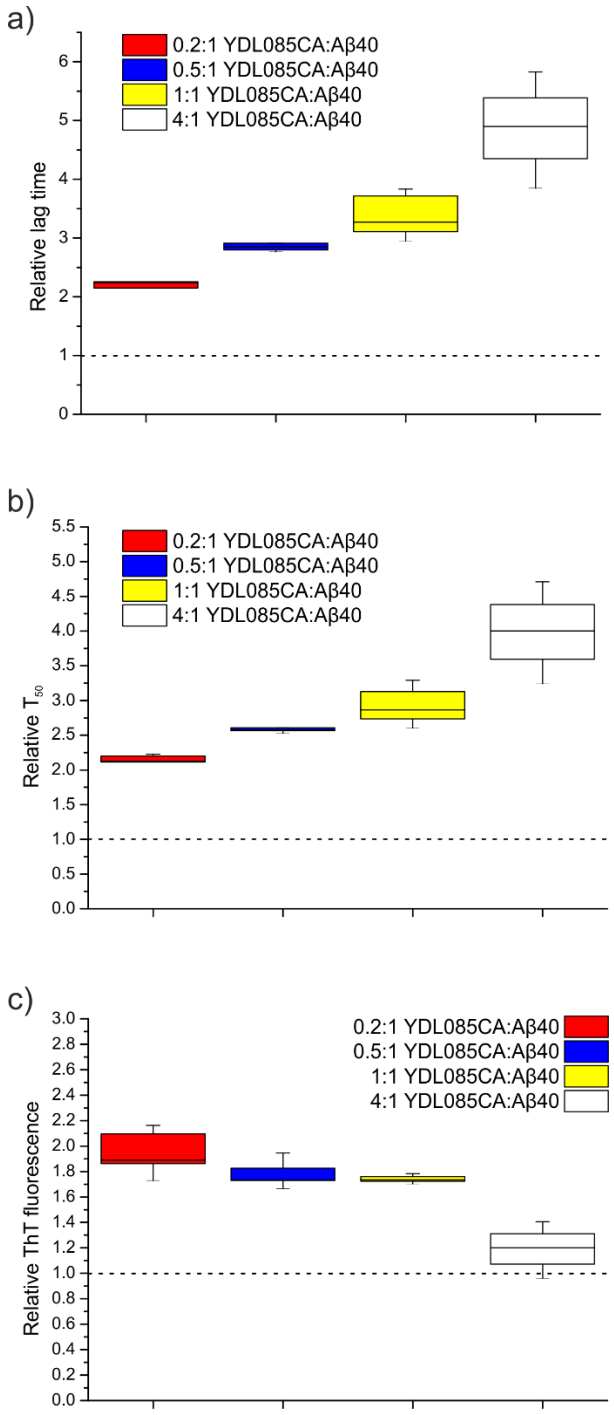


Figure A.II.2: Analysis of varying YDL085CA concentration upon aggregation of synthetic Aβ40. (a) The lag time of aggregation, relative to Aβ40 alone, for Aβ40 in the presence of four different concentrations of YDL085CA. (b) The T50 of Aβ40 aggregation for the same data set. (c) The relative maximum fluorescence value for the same data set.

Table A.II.1: Relative lag time, T50 and maximal fluorescence for ThT assay data presented in Figure 3.3.

YDL085CA:Aβ40		
Lag time	**T_{50}**	**Maximal fluorescence**
1.37	1.17	1.33
1.32	1.06	1.38
0.86	0.75	1.64

Table A.II.2: Relative lag time, T50 and maximal fluorescence for ThT assay data presented in Figure 3.8 and Figure 3.9. Values presented in the final row are p values relative to untreated wells.

YDL085CA:Aβ40											
0.2:1			0.5:1			1:1			4:1		
Lag time	T_{50}	Maximal flouresence	Lag time	T_{50}	Maximal flouresence	Lag time	T_{50}	Maximal flouresence	Lag time	T_{50}	Maximal flouresence
1.48	1.09	0.31	2.17	1.58	0.38	2.11	1.43	0.92	2.60	1.84	1.55
0.92	1.11	0.33	3.92	2.61	0.41	2.30	1.49	0.75	9.67	6.17	1.70
3.53	1.09	0.32	2.37	1.53	0.36	2.43	1.60	1.10	9.45	5.91	1.36
1.43	1.30	0.39	2.31	1.56	0.74	2.19	1.46	1.40	10.70	6.57	1.58
1.55	1.11	0.43	2.01	1.31	0.53	3.77	2.42	1.54	5.39	3.50	1.46
1.63	1.08	0.41	0.37	0.69	0.71	2.60	1.74	1.51	2.67	1.84	1.98
0.65	0.46	2.01	0.63	0.68	1.56	1.59	1.10	1.65	2.93	2.31	1.66
0.64	0.46	1.52	0.66	0.67	1.84	1.64	1.14	1.61	3.60	2.90	1.39
0.59	0.48	1.64	0.68	0.71	1.58	1.43	1.03	1.62	2.14	1.71	1.25
0.67	0.47	1.59	0.84	0.63	1.63	1.41	0.97	1.51	2.59	2.08	1.44
0.64	0.46	1.71	0.85	0.61	1.83	1.37	0.97	1.57	2.97	2.37	1.61
0.63	0.46	2.20	0.98	0.80	1.64	1.30	0.93	1.51	1.96	1.62	1.21
0.35	0.65	1.68	0.92	0.81	2.16	1.46	1.43	1.95	2.52	2.09	1.53
0.45	0.67	2.08	0.96	0.78	2.31	1.41	1.14	2.13	2.11	1.75	3.24
0.48	0.65	2.11	0.88	0.85	2.03	-0.69	1.21	1.84	2.25	1.61	1.85
0.74	0.57	1.75	1.16	0.81	1.83	1.49	1.44	1.81	3.07	2.17	1.74
0.75	0.57	1.82	0.98	0.86	2.31	1.76	1.25	1.58	1.32	1.47	1.54
0.75	0.56	2.21	1.21	1.02	1.73	1.73	1.23	2.17	3.12	2.51	1.53

1.09	0.91		1.13	0.94		0.97	0.86		5.09	3.32	
0.81	0.75		1.58	1.31		0.97	0.87		2.91	2.05	
0.77	0.69		0.83	0.74		1.68	1.37		2.91	2.08	
0.57	0.72		1.03	0.88		1.37	1.15		2.94	2.07	
0.83	0.71		1.33	1.09		1.29	1.09		2.39	1.89	
0.63	0.65					1.29	1.11		3.40	2.49	
p = 0.66621	p = 0.00005	p = 0.05761	p = 0.08811	p = 0.83134	p = 0.02007	p = 0.00077	p = 0.00080	p < 0.00001	p = 0.00002	p < 0.00001	p < 0.00001

Table A.II.3: Relative lag time, T50 and maximal fluorescence for ThT assay data presented in Figure 3.10 and Figure A.ii.i. The concentration of Aβ40 is indicated in brackets. Values presented in the final row are p values relative to untreated wells.

YDL085CA:Aβ40											
0.2:1 (100 µM)			0.5:1 (50 µM)			1:1 (25 µM)			5:1 (5 µM)		
Lag time	T_{50}	Maximal fluorescence	Lag time	T_{50}	Maximal fluorescence	Lag time	T_{50}	Maximal fluorescence	Lag time	T_{50}	Maximal fluorescence
0.92	1.10	1.23	0.84	0.94	1.38	1.14	1.18	1.63	14.31	17.49	0.57
1.19	1.38	1.43	1.01	1.09	1.91	2.54	2.38	1.39	18.67	21.24	0.50
0.94	1.14	1.40	1.50	1.49	1.95	0.95	1.02	1.59	23.81	26.12	0.08
0.94	1.12	1.07	0.97	1.07	1.76	0.95	0.97	1.71	27.43	29.73	0.63
p = 0.97533	p = 0.06632	p = 0.04244	p = 0.62505	p = 0.30226	p = 0.01011	p = 0.37926	p = 0.33130	p = 0.00351	p = 0.00606	p = 0.00351	p = 0.02114

Table A.II.4: Relative lag time, T50 and maximal fluorescence for ThT assay data presented in Figure 3.11 and Figure A.ii.ii. Values presented in the final row are p values relative to untreated wells.

YDL085CA:Aβ40											
0.2:1			0.5:1			1:1			4:1		
Lag time	T_{50}	Maximal flouresence	Lag time	T_{50}	Maximal flouresence	Lag time	T_{50}	Maximal flouresence	Lag time	T_{50}	Maximal flouresence
2.43	2.20	2.16	2.91	2.61	1.95	3.83	3.29	1.78	4.94	4.06	1.19
2.24	2.23	2.10	2.80	2.56	1.83	3.71	3.13	1.70	5.83	4.71	0.96
2.26	2.13	1.89	2.85	2.58	1.73	2.95	2.61	1.76	4.85	3.95	1.22
2.15	2.12	1.73	2.77	2.53	1.73	3.11	2.74	1.73	3.85	3.24	1.41
1.89	1.93	1.86	4.58	3.87	1.67	3.27	2.87	1.72			
p = 0.00017	p = 0.00003	p = 0.00028	p = 0.00340	p = 0.00213	p = 0.00009	p = 0.00016	p = 0.00011	p < 0.00001	p = 0.00243	p = 0.00217	p = 0.12690

Table A.II.5: Relative lag time, T50 and maximal fluorescence for ThT assay data presented in Figure 3.12 and Figure 3.13.

YDL085CA:Aβ40											
0.2:1			0.5:1			1:1			4:1		
Lag time	T_{50}	Maximal fluorescence	Lag time	T_{50}	Maximal fluorescence	Lag time	T_{50}	Maximal fluorescence	Lag time	T_{50}	Maximal fluorescence
0.34	0.25	1.53	0.50	0.43	1.29	0.27	0.22	1.40	0.52	0.42	2.67
0.24	0.21	1.56	0.50	0.44	1.62	0.25	0.19	1.17	0.53	0.48	3.08
0.27	0.22	1.69	0.52	0.43	1.38	0.27	0.23	1.43	0.51	0.50	3.19
0.29	0.27	1.71	0.45	0.39	1.43	0.29	0.23	0.91	0.49	0.40	3.30
0.29	0.25	1.72	0.47	0.33	1.41	0.32	0.24	1.11	0.47	0.41	2.90
0.40	0.28	1.51	0.21	0.21	1.35	0.32	0.21	0.97			
			0.20	0.17	1.64	0.48	0.41	2.04			
			0.26	0.21	1.87	0.48	0.41	1.97			
			0.25	0.21	1.64	0.46	0.40	1.80			
			0.39	0.28	1.57	0.45	0.36	1.90			
$p < 0.00001$	$p < 0.00001$	$p = 0.00002$	$p < 0.00001$	$p < 0.00001$	$p < 0.00001$	$p < 0.00001$	$p < 0.00001$	$p = 0.00710$	$p < 0.00001$	$p < 0.00001$	$p = 0.00005$

Table AII.6: Example cross linking data from a 1:1 complex of Aβ40 and YDL085CA. Selected columns from the Stavrox output file are shown. Not that the mass deviation each ion is for the corresponding $M+H^+$ ion compared to a theoretical mass.

Score	m/z	Charge	Deviation in ppm	Peptide 1	Protein 1	From	To	Peptide2	Protein 2	From	To	best linkage position peptide 1	best linkage position peptide 2	FDR
154	839.38	2	0.60	{MDAEFR]	>AB40	0	6	{MDAEFR]	>AB40	0	6	M0	M0	0.00
126	847.38	2	0.53	{MDAEFR]	>AB40	0	6	{mDAEFR]	>AB40	0	6	M0	M0	0.00
87	855.38	2	2.39	{mDAEFR]	>AB40	0	6	{mDAEFR]	>AB40	0	6	M0	M0	0.00
33	837.37	2	2.39	{MDAEFR]	>AB40	0	6	{MDAEFR]	>AB40	0	6	M0	M0	0.03
20	536.77	4	-2.66	{mDAEFR]	>AB40	0	6	[KQKDMAKNQK]	>YDL085CA	17	26	M0	K10	0.04
61	581.95	3	-1.02	{MDAEFR]	>AB40	0	6	[DMAKNQK]	>YDL085CA	20	26	M0	K4	0.03
190	623.32	3	2.23	{MDAEFR]	>AB40	0	6	[QKQAAADAR]	>YDL085CA	44	52	M0	K2	0.00
181	623.32	3	2.01	{MDAEFR]	>AB40	0	6	[QKQAAADAR]	>YDL085CA	44	52	M0	K2	0.00
120	628.65	3	2.15	{mDAEFR]	>AB40	0	6	[QKQAAADAR]	>YDL085CA	44	52	M0	K2	0.00
114	621.97	3	-0.40	{MDAEFR]	>AB40	0	6	[QKQAAADAR]	>YDL085CA	44	52	M0	K2	0.00
63	627.30	3	-1.19	{mDAEFR]	>AB40	0	6	[QKQAAADAR]	>YDL085CA	44	52	M0	K2	0.03
173	585.97	3	2.57	{MDAEFR]	>AB40	0	6	[EAEKLEK]	>YDL085CA	54	60	M0	K4	0.00
118	585.96	3	-0.57	{MDAEFR]	>AB40	0	6	[EAEKLEK]	>YDL085CA	54	60	M0	K4	0.00

Score	m/z	Charge	Deviation in ppm	Peptide 1	Protein 1	From	To	Peptide2	Protein 2	From	To	best linkage position peptide 1	best linkage position peptide 2	FDR
116	591.30	3	-0.06	{mDAEFR]	>AB40	0	6	[EAEKLEK]	>YDL085CA	54	60	M0	K4	0.00
106	666.36	3	-1.61	{MDAEFR]	>AB40	0	6	[EAEKLEKLK]	>YDL085CA	54	62	M0	K9	0.00
83	591.29	3	-1.92	{mDAEFR]	>AB40	0	6	[EAEKLEK]	>YDL085CA	54	60	M0	K4	0.00
63	500.02	4	-2.74	{MDAEFR]	>AB40	0	6	[EAEKLEKLK]	>YDL085CA	54	62	M0	K4	0.03
44	500.02	4	2.27	{MDAEFR]	>AB40	0	6	[EAEKLEKLK]	>YDL085CA	54	62	M0	K9	0.02
37	589.95	3	-1.92	{mDAEFR]	>AB40	0	6	[EAEKLEK]	>YDL085CA	54	60	M0	K4	0.02
133	643.81	2	-0.18	{MAR]	>YDL085CA	0	3	{MDAEFR]	>AB40	0	6	M0	M0	0.00
85	643.81	2	1.21	{MAR]	>YDL085CA	0	3	{MDAEFR]	>AB40	0	6	M0	M0	0.00
77	641.80	2	1.15	{MAR]	>YDL085CA	0	3	{MDAEFR]	>AB40	0	6	M0	M0	0.04
67	649.80	2	1.81	{MAR]	>YDL085CA	0	3	{mDAEFR]	>AB40	0	6	M0	M0	0.03
67	649.80	2	1.81	{mAR]	>YDL085CA	0	3	{MDAEFR]	>AB40	0	6	M0	M0	0.03
65	651.81	2	1.02	{MAR]	>YDL085CA	0	3	{mDAEFR]	>AB40	0	6	M0	M0	0.03
65	651.81	2	1.02	{mAR]	>YDL085CA	0	3	{MDAEFR]	>AB40	0	6	M0	M0	0.03
53	659.81	2	0.40	{mAR]	>YDL085CA	0	3	{mDAEFR]	>AB40	0	6	M0	M0	0.02
41	651.81	2	0.50	{MAR]	>YDL085CA	0	3	{mDAEFR]	>AB40	0	6	M0	M0	0.02
41	651.81	2	0.50	{mAR]	>YDL085CA	0	3	{MDAEFR]	>AB40	0	6	M0	M0	0.02
32	649.80	2	1.35	{MAR]	>YDL085CA	0	3	{mDAEFR]	>AB40	0	6	M0	M0	0.03

Score	m/z	Charge	Deviation in ppm	Peptide 1	Protein 1	From	To	Peptide2	Protein 2	From	To	best linkage position peptide 1	best linkage position peptide 2	FDR
32	649.80	2	1.35	{mAR]	>YDL085CA	0	3	{MDAEFR]	>AB40	0	6	M0	M0	0.03
23	651.81	2	1.73	{MAR]	>YDL085CA	0	3	{mDAEFR]	>AB40	0	6	M0	M0	0.04
23	651.81	2	1.73	{mAR]	>YDL085CA	0	3	{MDAEFR]	>AB40	0	6	M0	M0	0.04
27	519.27	3	-1.86	[QKNLK]	>YDL085CA	12	16	{mDAEFR]	>AB40	0	6	K2	M0	0.03
20	513.94	3	2.21	[QKNLK]	>YDL085CA	12	16	{MDAEFR]	>AB40	0	6	K2	M0	0.04
52	543.94	3	0.22	[QKDMAK]	>YDL085CA	18	23	{MDAEFR]	>AB40	0	6	K2	M0	0.02
79	481.58	3	-0.87	[NQKK]	>YDL085CA	24	27	{mDAEFR]	>AB40	0	6	K4	M0	0.00
32	476.25	3	0.76	[NQKK]	>YDL085CA	24	27	{MDAEFR]	>AB40	0	6	K4	M0	0.03
99	519.59	3	0.69	[KSGDPK]	>YDL085CA	27	32	{mDAEFR]	>AB40	0	6	K1	M0	0.00
94	514.26	3	2.19	[KSGDPK]	>YDL085CA	27	32	{MDAEFR]	>AB40	0	6	K1	M0	0.00
56	519.59	3	-1.70	[KSGDPK]	>YDL085CA	27	32	{mDAEFR]	>AB40	0	6	K1	M0	0.02
69	514.26	3	2.19	[SGDPKK]	>YDL085CA	28	33	{MDAEFR]	>AB40	0	6	S1	M0	0.03
32	519.59	3	0.69	[SGDPKK]	>YDL085CA	28	33	{mDAEFR]	>AB40	0	6	S1	M0	0.03
19	519.59	3	-1.70	[SGDPKK]	>YDL085CA	28	33	{mDAEFR]	>AB40	0	6	S1	M0	0.04
44	513.95	3	2.24	[LEKLK]	>YDL085CA	58	62	{MDAEFR]	>AB40	0	6	K5	M0	0.02
116	499.94	3	2.96	[LKAEK]	>YDL085CA	61	65	{MDAEFR]	>AB40	0	6	K2	M0	0.00
75	505.27	3	1.11	[LKAEK]	>YDL085CA	61	65	{mDAEFR]	>AB40	0	6	K2	M0	0.04
74	498.59	3	-2.00	[LKAEK]	>YDL085CA	61	65	{MDAEFR]	>AB40	0	6	K2	M0	0.03
65	505.27	3	1.65	[LKAEK]	>YDL085CA	61	65	{mDAEFR]	>AB40	0	6	K5	M0	0.03

Score	m/z	Charge	Deviation in ppm	Peptide 1	Protein 1	From	To	Peptide2	Protein 2	From	To	best linkage position peptide 1	best linkage position peptide 2	FDR
99	383.56	3	-1.44	{MAR]	>YDL085CA	0	3	[QKNLK]	>YDL085CA	12	16	M0	K2	0.00
50	582.84	2	-1.48	{mAR]	>YDL085CA	0	3	[QKNLK]	>YDL085CA	12	16	M0	K2	0.02
42	510.81	2	-0.99	{MAR]	>YDL085CA	0	3	[NLKK]	>YDL085CA	14	17	M0	K4	0.02
25	874.96	2	1.34	{MAR]	>YDL085CA	0	3	[KQKDmAKNQK]	>YDL085CA	17	26	M0	K7	0.03
140	413.56	3	0.08	{MAR]	>YDL085CA	0	3	[QKDMAK]	>YDL085CA	18	23	M0	K2	0.00
66	418.89	3	2.62	{mAR]	>YDL085CA	0	3	[QKDMAK]	>YDL085CA	18	23	M0	K2	0.03
23	418.89	3	-1.01	{mAR]	>YDL085CA	0	3	[QKDMAK]	>YDL085CA	18	23	M0	K2	0.04
20	418.89	3	-1.01	{MAR]	>YDL085CA	0	3	[QKDmAK]	>YDL085CA	18	23	M0	K2	0.04
153	492.94	3	2.45	{MAR]	>YDL085CA	0	3	[QKQAAADAR]	>YDL085CA	44	52	M0	K2	0.00
88	491.59	3	-0.90	{MAR]	>YDL085CA	0	3	[QKQAAADAR]	>YDL085CA	44	52	M0	K2	0.00
60	460.92	3	2.22	{mAR]	>YDL085CA	0	3	[EAEKLEK]	>YDL085CA	54	60	M0	K4	0.03
50	460.92	3	0.74	{mAR]	>YDL085CA	0	3	[EAEKLEK]	>YDL085CA	54	60	M0	K4	0.02
37	369.56	3	-0.66	{MAR]	>YDL085CA	0	3	[LKAEK]	>YDL085CA	61	65	M0	K2	0.02
46	659.06	3	-2.18	[QKNLK]	>YDL085CA	12	16	[NLKKQKDMAK]	>YDL085CA	14	23	K5	K4	0.02
29	624.01	3	1.08	[KQKDmAK]	>YDL085CA	17	23	[KQKDmAK]	>YDL085CA	17	23	K1	K1	0.03
70	536.28	3	-1.46	[DMAK]	>YDL085CA	20	23	[NQKKSGDPK]	>YDL085CA	24	32	K4	K3	0.03
34	484.26	4	-2.12	[DMAKNQK]	>YDL085CA	20	26	[QKQAAADAR]	>YDL085CA	44	52	K4	K2	0.02
62	539.64	3	2.72	[NQKK]	>YDL085CA	24	27	[QKQAAADAR]	>YDL085CA	44	52	K3	K2	0.03

Score	m/z	Charge	Deviation in ppm	Peptide 1	Protein 1	From	To	Peptide2	Protein 2	From	To	best linkage position peptide 1	best linkage position peptide 2	FDR
19	539.64	3	-1.73	[NQKK]	>YDL085CA	24	27	[QKQAAADAR]	>YDL085CA	44	52	K4	K2	0.04
51	498.27	3	-2.86	[KSGDPK]	>YDL085CA	27	32	[QKDMAK]	>YDL085CA	18	23	S2	K2	0.02
79	536.28	3	-1.46	[KSGDPK]	>YDL085CA	27	32	[DMAKNQK]	>YDL085CA	20	26	K1	K4	0.00
58	658.06	3	1.25	[KSGDPKKR]	>YDL085CA	27	34	[KSGDPKKR]	>YDL085CA	27	34	S2	K7	0.02
100	433.49	4	-1.14	[KSGDPK]	>YDL085CA	27	32	[QKQAAADAR]	>YDL085CA	44	52	K1	K2	0.00
71	577.65	3	-0.34	[KSGDPK]	>YDL085CA	27	32	[QKQAAADAR]	>YDL085CA	44	52	S2	K2	0.03
64	577.65	3	0.00	[KSGDPK]	>YDL085CA	27	32	[QKQAAADAR]	>YDL085CA	44	52	S2	K2	0.03
57	433.49	4	-2.60	[KSGDPK]	>YDL085CA	27	32	[QKQAAADAR]	>YDL085CA	44	52	K6	K2	0.02
52	577.65	3	-2.34	[KSGDPK]	>YDL085CA	27	32	[QKQAAADAR]	>YDL085CA	44	52	S2	K2	0.02
47	498.27	3	-2.86	[SGDPK]	>YDL085CA	28	32	[KQKDMAK]	>YDL085CA	17	23	S1	K1	0.02
86	536.28	3	-1.46	[SGDPK]	>YDL085CA	28	32	[DMAKNQKK]	>YDL085CA	20	27	S1	K7	0.00
18	536.28	3	-1.46	[SGDPKK]	>YDL085CA	28	33	[DMAKNQK]	>YDL085CA	20	26	K5	K4	0.06
76	433.49	4	-1.14	[SGDPKK]	>YDL085CA	28	33	[QKQAAADAR]	>YDL085CA	44	52	K5	K2	0.04
75	577.65	3	-0.34	[SGDPKK]	>YDL085CA	28	33	[QKQAAADAR]	>YDL085CA	44	52	K5	K2	0.04
48	433.49	4	-2.60	[SGDPKK]	>YDL085CA	28	33	[QKQAAADAR]	>YDL085CA	44	52	K6	K2	0.02

Score	m/z	Charge	Deviation in ppm	Peptide 1	Protein 1	From	To	Peptide2	Protein 2	From	To	best linkage position peptide 1	best linkage position peptide 2	FDR
32	577.65	3	0.00	[SGDPKK]	>YDL085CA	28	33	[QKQAAADAR]	>YDL085CA	44	52	S1	K2	0.03
22	577.65	3	-2.34	[SGDPKK]	>YDL085CA	28	33	[QKQAAADAR]	>YDL085CA	44	52	S1	K2	0.04
24	541.53	4	-2.89	[QKQAAADAR]	>YDL085CA	44	52	[MESDAEILR]	>YDL085CA	35	43	K2	S3	0.05
39	649.35	3	-0.94	[EAEKLEK]	>YDL085CA	54	60	[QKQAAADAR]	>YDL085CA	44	52	K4	K2	0.02
90	525.98	3	-0.27	[EAEK]	>YDL085CA	54	57	[LEKLKAEK]	>YDL085CA	58	65	K4	K3	0.00
31	524.63	3	0.50	[EAEK]	>YDL085CA	54	57	[LEKLKAEK]	>YDL085CA	58	65	K4	K5	0.03
131	525.98	3	-0.27	[LKAEK]	>YDL085CA	61	65	[EAEKLEK]	>YDL085CA	54	60	K2	K4	0.00
67	524.63	3	0.50	[LKAEK]	>YDL085CA	61	65	[EAEKLEK]	>YDL085CA	54	60	K2	K4	0.03
133	393.74	2	-0.17	[QKNLK]	>YDL085CA	12	16		dead-end			K2		0.00
50	438.73	2	0.68	[QKDMAK]	>YDL085CA	18	23		dead-end			K2		0.00
63	495.75	2	0.75	[DMAKNQK]	>YDL085CA	20	26		dead-end			K4		0.00
106	507.27	3	1.99	[KRmESDAEILR]	>YDL085CA	33	43		dead-end			K1		0.00
131	557.80	2	0.37	[QKQAAADAR]	>YDL085CA	44	52		dead-end			K2		0.00
67	501.77	2	0.87	[EAEKLEK]	>YDL085CA	54	60		dead-end			K4		0.00
64	524.63	3	0.50	[EAEKLEKLKAEK]	>YDL085CA	54	65		dead-end			K7		0.00
145	372.73	2	1.96	[LKAEK]	>YDL085CA	61	65		dead-end			K2		0.00

Score	m/z	Charge	Deviation in ppm	Peptide 1	Protein 1	From	To	Peptide2	Protein 2	From	To	best linkage position peptide 1	best linkage position peptide 2	FDR
113	493.77	2	-0.41	[KQKDMAK]	>YDL085CA	17	23		intrapeptidal			K1	K3	0.00
77	570.31	2	2.05	[NQKKSGDPK]	>YDL085CA	24	32		intrapeptidal			K4	S5	0.00
84	613.36	2	2.75	[EAEKLEKLK]	>YDL085CA	54	62		intrapeptidal			K4	K7	0.00
49	518.63	3	-2.43	[EAEKLEKLKAEK]	>YDL085CA	54	65		intrapeptidal			K7	K9	0.00

Appendix III:

Appendix III contains figures and information relating to Chapter 4, which is concerned with the investigation of the interaction between the peptide inhibitor of amyloid aggregation, QBP1, and the amyloid protein ataxin-3.

Table AIII.1: CCS of ataxin-3 78Q and ataxin-3 78Q:QBP1 complexes determined by nESI-IMS-MS.

	Ataxin-3 78Q		Ataxin-3 78Q + QBP1	
z	CCS (Å²)	SD (Å²)	CCS (Å²)	SD (Å²)
14	3508.93	22.99	3582.00	47.04
14	3563.57			
15	3775.87	21.62	3780.88	14.39
15			3698.69	4.28
16	3972.65	17.21	3969.91	6.37
17	4170.55	15.84	4186.54	3.99
18	4419.72	19.54	4419.51	19.54
19	4622.72	49.16	4614.87	48.09
20	4772.81	23.98	4772.58	23.98
20	4937.67	19.61	4903.95	36.11
21	4911.33		5010.95	61.98
21	5182.11	85.90	5216.07	41.22
21	5542.37	52.94	5402.66	

Table AIII.2: CCS of ataxin-3 14Q and ataxin-3 14Q:QBP1 complexes determined by nESI-IMS-MS.

	Ataxin-3 14Q		Ataxin-3 14Q + QBP1	
z	CCS (Å²)	SD (Å²)	CCS (Å²)	SD (Å²)
13	3254.7	20.04	3274.84	70.09
13	3368.5	65.05		
14	3539.43	74.45	3489.19	43.39
15	3731.69	56.70	3735.70	81.15
16	3969.11	43.97	3953.11	58.01
17	4265.97	54.79	4191.30	39.98
18	4319.92	82.64	4393.29	76.42
18	4531.99	43.19		
19			4457.07	
19	4848.49	41.79	4788.78	50.53
20	5166.29		5158.07	

Table AIII.3: CCS of JDU1 and JDU1:QBP1 complexes determined by nESI-IMS-MS.

	JDU1		JDU1 + QBP1	
z	CCS (Å²)	SD (Å²)	CCS (Å²)	SD (Å²)
9	2204.68	49.70	2229.06	
9	2444.11	23.64	2492.71	15.71
10	2432.55	68.94	2363.35	
10	2583.91	36.40	2569.04	47.10
11	2728.67	16.51	2580.39	
11	2826.89		2742.02	11.95
12	2999.80	47.77	2975.42	13.39
13	3436.18	35.30		
13	3133.57	69.71	3238.26	18.26
14	3100.68		3202.59	47.53
14	3583.39	112.74	3682.24	34.18

Table AIII.4: CCS of JD+ and JD+:QBP1 complexes determined by nESI-IMS-MS.

	JD+		JD+ + QBP1	
z	CCS (Å²)	SD (Å²)	CCS (Å²)	SD (Å²)
10	2418.16	24.96	2448.26	31.58
10	2645.30	18.43	2680.38	16.23
11	2671.92		2648.42	36.51
11	2788.29	24.08	2828.48	23.96
12	2985.72	29.29	2971.50	23.09
12			3050.23	
13	3277.13	11.25	3260.73	21.29
14	3736.23	63.59	3520.34	108.99
14	3228.36			
15	3517.36	90.84	3531.12	76.78
15	4023.13	19.87	4029.95	28.85

Table AIII.5: CCS of JD determined by nESI-IMS-MS.

	JD		JD + QBP1	
z	**CCS (Å²)**	**SD (Å²)**	**CCS (Å²)**	**SD (Å²)**
7	1808.09	58.91	-	-
7	1887.74	50.96	-	-
7	2076.46	30.41	-	-
8	1913.92		-	-
8	2206.51	46.73	-	-
9	2357.78	64.29	-	-
10	2564.02	83.44	-	-
11	2651.96		-	-
11	2970.21	73.51	-	-
12	2933.16	116.57	-	-
12	3301.47	80.56	-	-

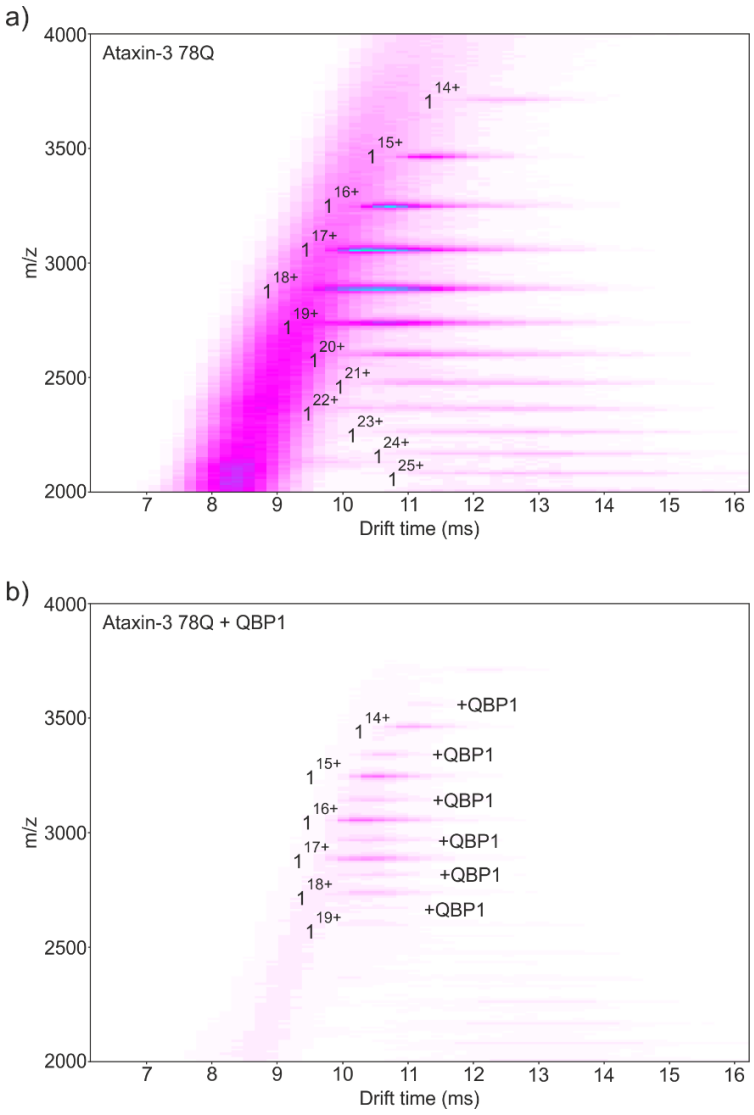


Figure AIII.1: IMS drift plots for ataxin-3 78Q (a) and ataxin-3 78Q + QBP1 (b). Plots are shown with drift time on the x-axis and m/z on the y-axis.

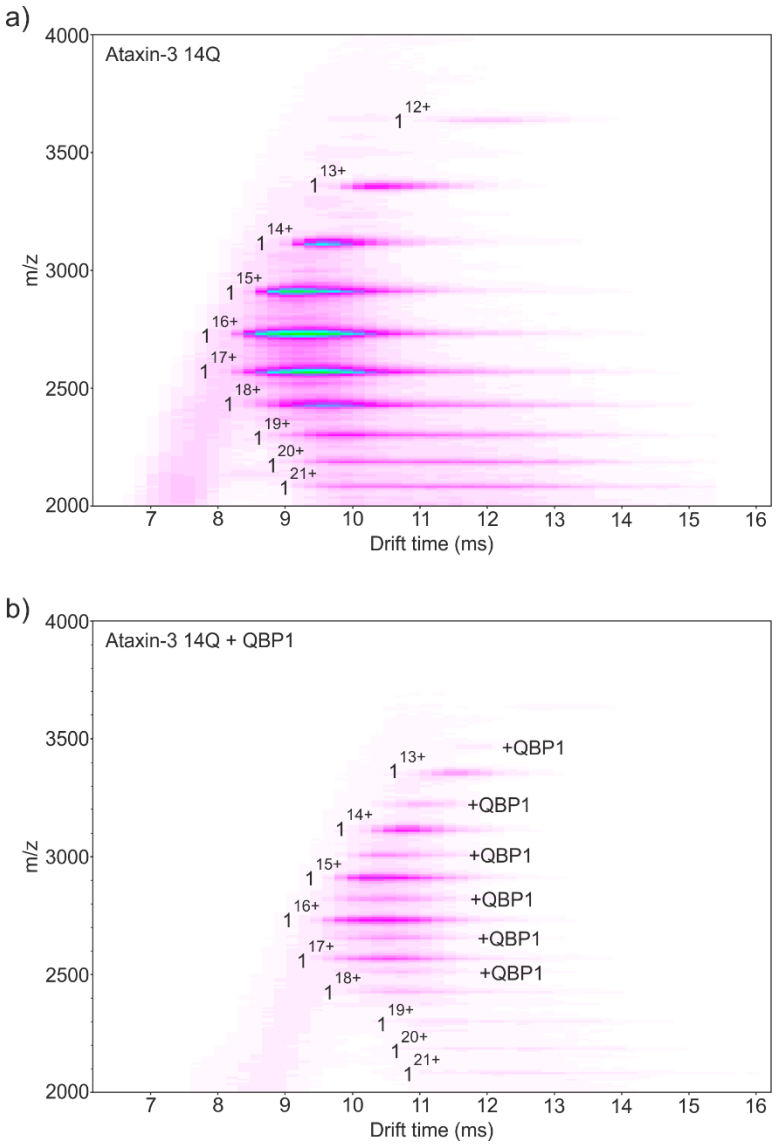


Figure AIII.2: IMS drift plots for ataxin-3 14Q (a) and ataxin-3 14Q + QBP1 (b). Plots are shown with drift time on the x-axis and m/z on the y-axis.

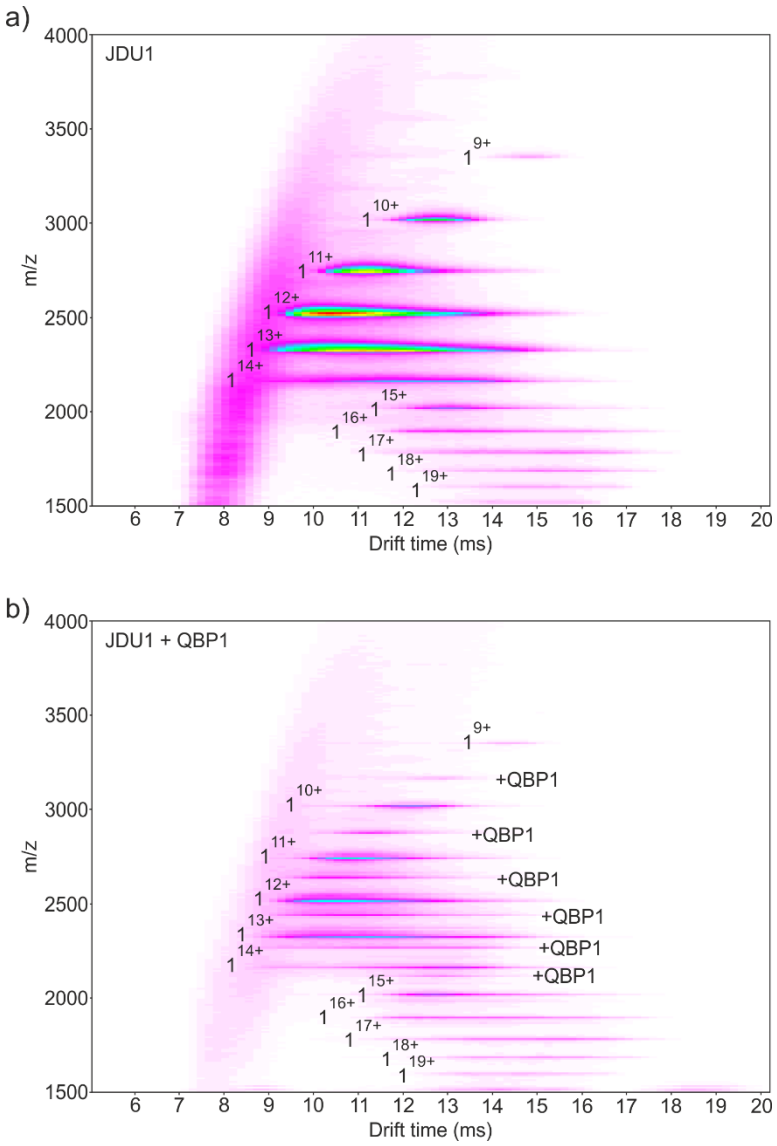


Figure AIII.3: IMS drift plots for JDU1 (a) and JDU1 + QBP1 (b). Plots are shown with drift time on the x-axis and m/z on the y-axis.

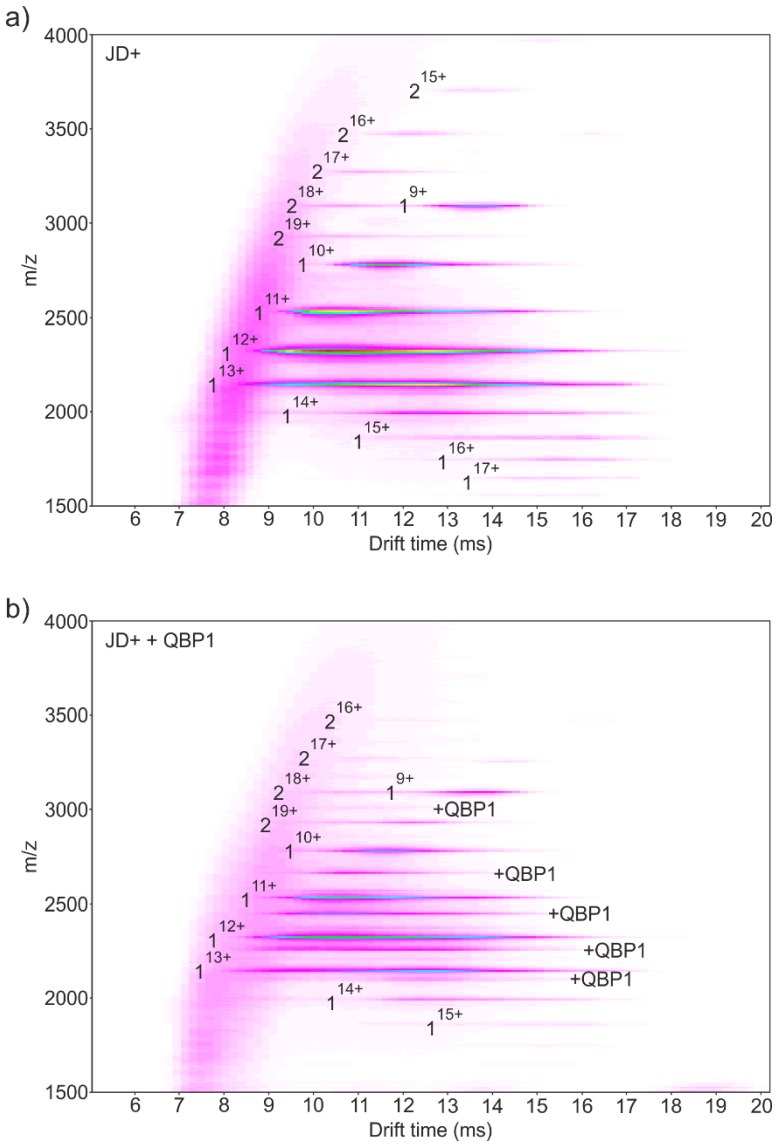


Figure AIII.4: IMS drift plots for JD+ (a) and JD+ + QBP1 (b). Plots are shown with drift time on the x-axis and m/z on the y-axis.

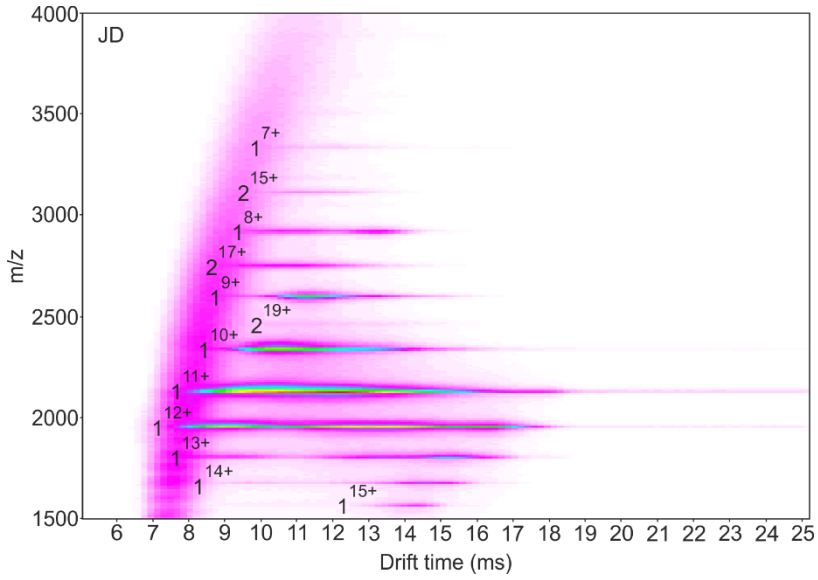


Figure AIII.5: IMS drift plot for JD. The plot shows drift time on the x-axis and m/z on the y-axis.

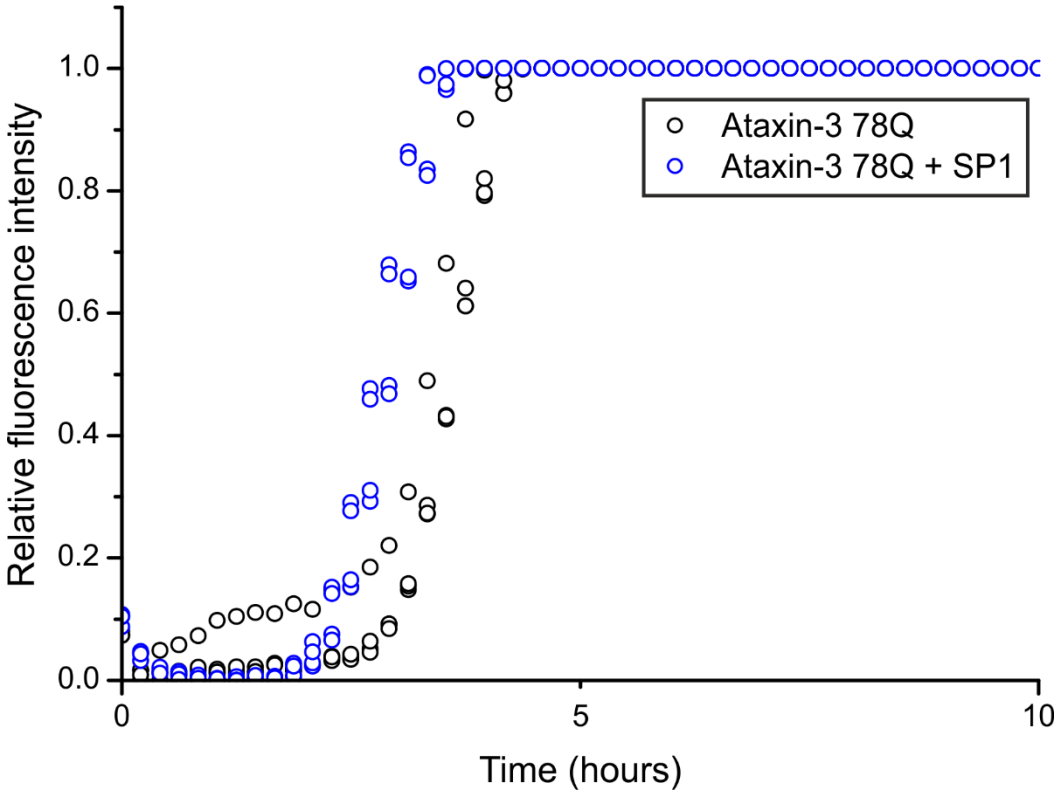
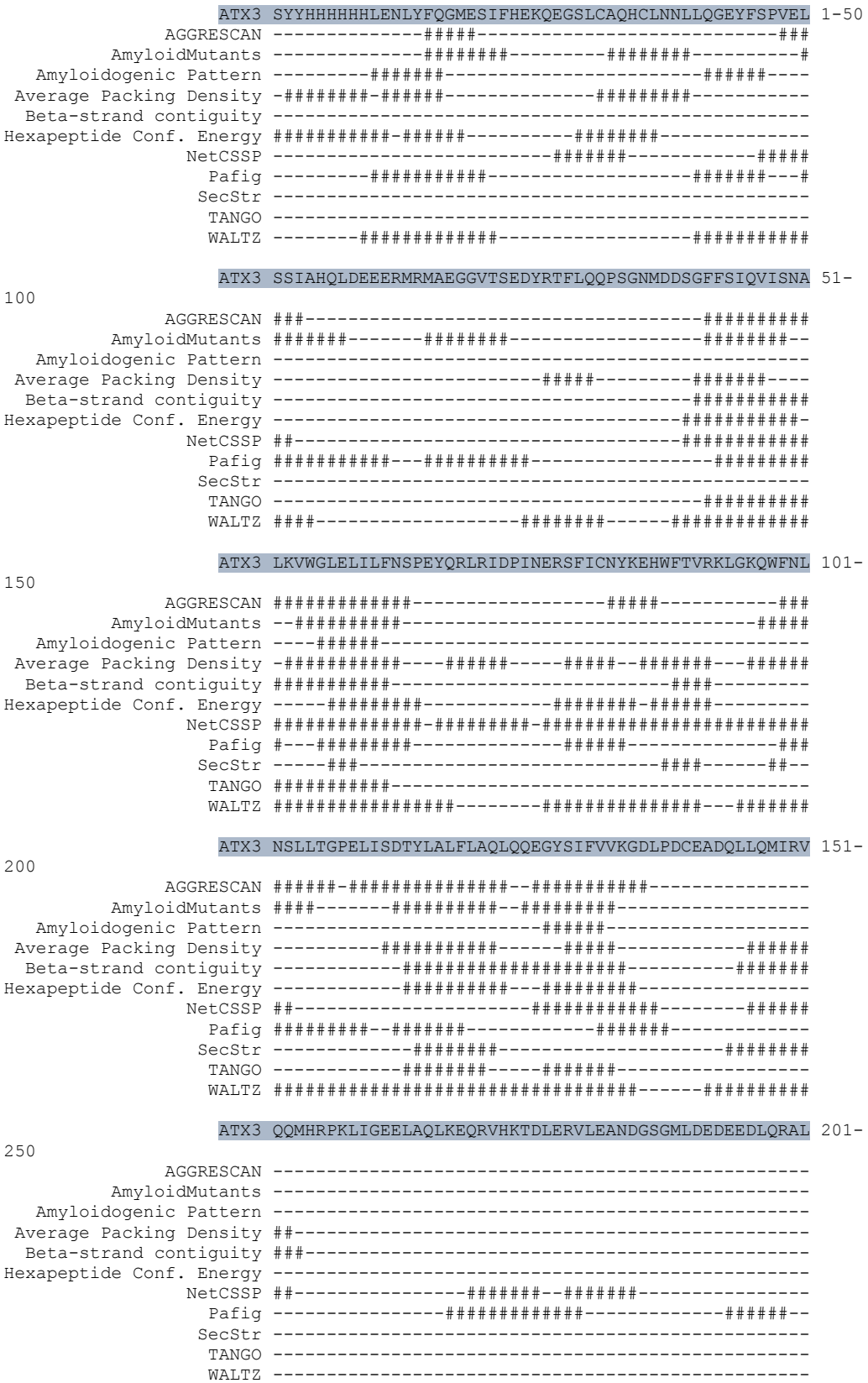


Figure AIII.6: Aggregation kinetics of ataxin-3 78Q in the presence of SP1. The kinetics of ataxin-3 78Q aggregation, measured by the ThT assay, for ataxin-3 78Q alone (black) and in the presence of a 5-fold excess of SP1 (blue). Not that the signal has reached maximum before five hours in all samples. Ataxin-3 was 10 µM in 250 mM ammonium bicarbonate, 1 mM DTT, 10 µM ThT, pH 7.8 and was incubated at 37 °C without shaking. SP1 was added to a final concentration of 50 µM.

The full output of the Amylpred 2.0 analysis described in Chapter 4 is provided below. Each algorithm employed is shown individually.



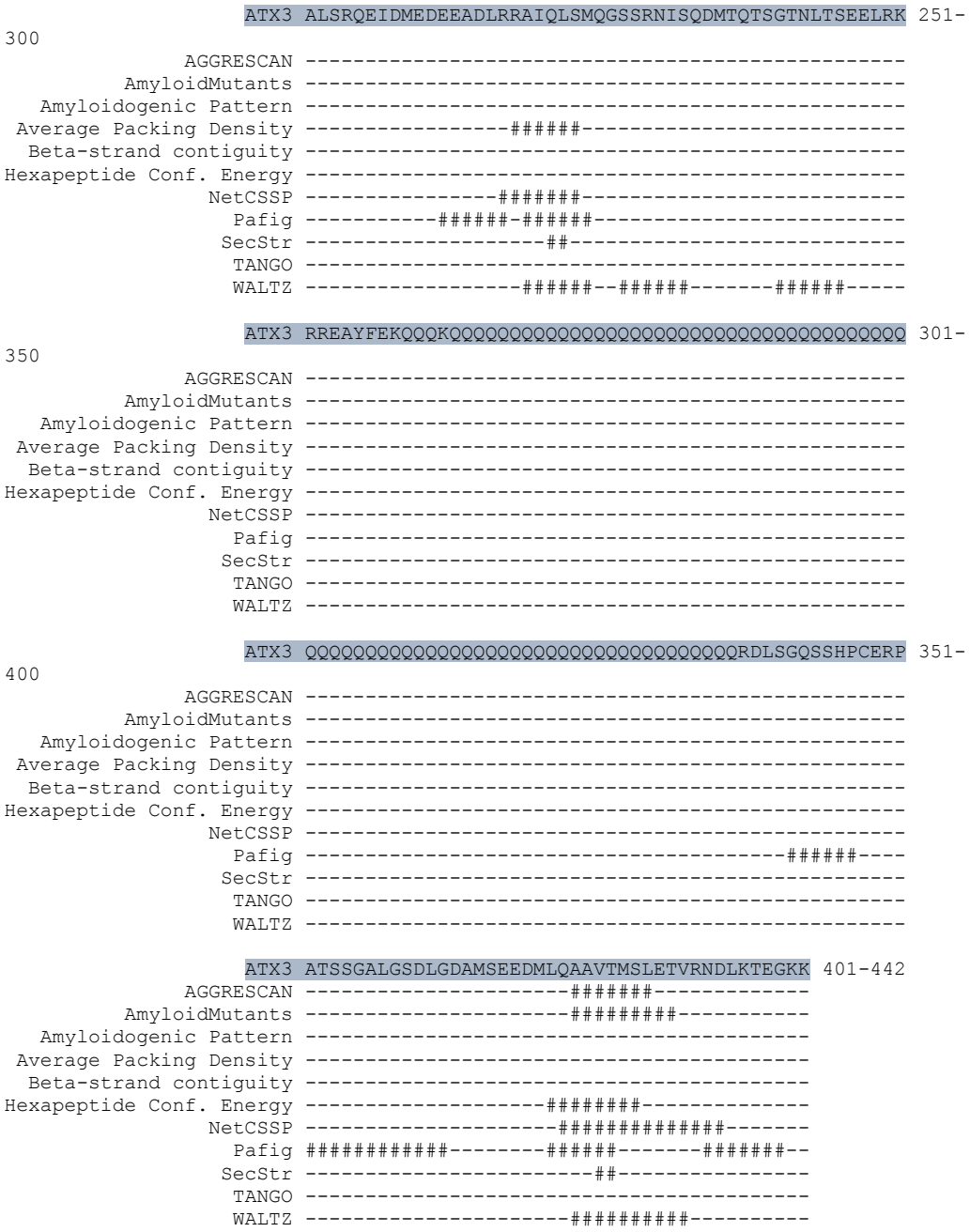


Figure AIII.7: Amylpred 2.0 analysis of the sequence of ataxin-3 78Q. The first line in each block indicates the region of ataxin-3 78Q. The 11 algorithms employed in the analysis are shown below the sequence. Where an algorithm predicts a potential amyloid forming region a # is shown. Figure prepared using Amylpred 2.0 (Tsolis *et al.* 2013).

References

ABELEIN, A., A. GRASLUND and J. DANIELSSON. 2015. Zinc as chaperone-mimicking agent for retardation of amyloid beta peptide fibril formation. *Proc Natl Acad Sci U S A,* **112**(17), pp.5407-12.

ADEGBUYIRO, A., F. SEDIGHI, A. W. T. PILKINGTON, S. GROOVER and J. LEGLEITER. 2017. Proteins Containing Expanded Polyglutamine Tracts and Neurodegenerative Disease. *Biochemistry,* **56**(9), pp.1199-1217.

ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS and D. J. LIPMAN. 1990. Basic local alignment search tool. *J Mol Biol,* **215**(3), pp.403-10.

AMIJEE, H., C. BATE, A. WILLIAMS, J. VIRDEE, R. JEGGO, D. SPANSWICK, D. I. SCOPES, J. M. TREHERNE, S. MAZZITELLI, R. CHAWNER, C. E. EYERS and A. J. DOIG. 2012. The N-methylated peptide SEN304 powerfully inhibits Abeta(1-42) toxicity by perturbing oligomer formation. *Biochemistry,* **51**(42), pp.8338-52.

ANDRADE, M. A., P. CHACON, J. J. MERELO and F. MORAN. 1993. Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng,* **6**(4), pp.383-90.

ANFINSEN, C. B., E. HABER, M. SELA and F. H. WHITE. 1961. THE KINETICS OF FORMATION OF NATIVE RIBONUCLEASE DURING OXIDATION OF THE REDUCED POLYPEPTIDE CHAIN. *Proc Natl Acad Sci U S A,* **47**(9), pp.1309-1314.

APICELLA, A., A. NATALELLO, A. M. FRANA, A. BASERGA, C. S. CASARI, C. E. BOTTANI, S. M. DOGLIA, P. TORTORA and M. E. REGONESI. 2012. Temperature profoundly affects ataxin-3 fibrillogenesis. *Biochimie,* **94**(4), pp.1026-31.

ARMEN, R. S., B. M. BERNARD, R. DAY, D. O. ALONSO and V. DAGGETT. 2005. Characterization of a possible amyloidogenic precursor in glutamine-repeat neurodegenerative diseases. *Proc Natl Acad Sci U S A,* **102**(38), pp.13433-8.

AROSIO, P., M. VENDRUSCOLO, C. M. DOBSON and T. P. KNOWLES. 2014. Chemical kinetics for drug discovery to combat protein aggregation diseases. *Trends Pharmacol Sci,* **35**(3), pp.127-35.

ASMUS, K. D. 1984. Pulse radiolysis methodology. *Methods Enzymol,* **105**, pp.167-78.

ASSARSSON, A., E. HELLSTRAND, C. CABALEIRO-LAGO and S. LINSE. 2014. Charge Dependent Retardation of Amyloid β Aggregation by Hydrophilic Proteins. *ACS Chem Neurosci,* **5**(4), pp.266-274.

AUER, S., F. MEERSMAN, C. M. DOBSON and M. VENDRUSCOLO. 2008. A Generic Mechanism of Emergence of Amyloid Protofilaments from Disordered Oligomeric Aggregates. *PLoS Computational Biology,* **4**(11), pe1000222.

BARBER, M., R. S. BORDOLI, R. D. SEDGWICK and A. N. TYLER. 1981. Fast atom bombardment of solids (F.A.B.): a new ion source for mass spectrometry. *Journal of the Chemical Society, Chemical Communications,* (7), pp.325-327.

BARR, J. D., L. SHI, D. H. RUSSELL, D. E. CLEMMER and A. E. HOLLIDAY. 2016. Following a Folding Transition with Capillary Electrophoresis and Ion Mobility Spectrometry. *Analytical Chemistry,* **88**(22), pp.10933-10939.

BAUER, P. O., A. GOSWAMI, H. K. WONG, M. OKUNO, M. KUROSAWA, M. YAMADA, H. MIYAZAKI, G. MATSUMOTO, Y. KINO, Y. NAGAI and N. NUKINA. 2010. Harnessing chaperone-mediated autophagy for the selective degradation of mutant huntingtin protein. *Nat Biotechnol,* **28**(3), pp.256-63.

BAZOTI, F. N., J. BERGQUIST, K. MARKIDES and A. TSARBOPOULOS. 2008. Localization of the noncovalent binding site between amyloid-beta-peptide and oleuropein using electrospray ionization FT-ICR mass spectrometry. *J Am Soc Mass Spectrom,* **19**(8), pp.1078-85.

BEESTON, H. S., J. R. AULT, S. D. PRINGLE, J. M. BROWN and A. E. ASHCROFT. 2015. Changes in protein structure monitored by use of gas-phase hydrogen/deuterium exchange. *Proteomics,* **15**(16), pp.2842-50.

BENESCH, J. L. and B. T. RUOTOLO. 2011. Mass spectrometry: come of age for structural and dynamical biology. *Curr Opin Struct Biol,* **21**(5), pp.641-9.

BENJAMIN, D. R., C. V. ROBINSON, J. P. HENDRICK, F. U. HARTL and C. M. DOBSON. 1998. Mass spectrometry of ribosomes and ribosomal subunits. *Proc Natl Acad Sci U S A,* **95**(13), pp.7391-5.

BENNHOLD, H. 1922. Eine specifische amyloid-farbung mit kongorot. *Munch. Med. Wochenschr,* **69**, pp.1537-1538.

BERKE, S. J., F. A. SCHMIED, E. R. BRUNT, L. M. ELLERBY and H. L. PAULSON. 2004. Caspase-mediated proteolysis of the polyglutamine disease protein ataxin-3. *J Neurochem,* **89**(4), pp.908-18.

BERNSTEIN, S. L., N. F. DUPUIS, N. D. LAZO, T. WYTTENBACH, M. M. CONDRON, G. BITAN, D. B. TEPLOW, J.-E. SHEA, B. T. RUOTOLO, C. V. ROBINSON and M. T. BOWERS. 2009a. Amyloid-β protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. *Nat Chem,* **1**(4), pp.326-331.

BERNSTEIN, S. L., N. F. DUPUIS, N. D. LAZO, T. WYTTENBACH, M. M. CONDRON, G. BITAN, D. B. TEPLOW, J. E. SHEA, B. T. RUOTOLO, C. V. ROBINSON and M. T. BOWERS. 2009b. Amyloid-beta protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. *Nat Chem,* **1**(4), pp.326-31.

BEVIVINO, A. E. and P. J. LOLL. 2001. An expanded glutamine repeat destabilizes native ataxin-3 structure and mediates formation of parallel β-fibrils. *Proceedings of the National Academy of Sciences,* **98**(21), pp.11955-11960.

BHATTACHARYYA, A., A. K. THAKUR, V. M. CHELLGREN, G. THIAGARAJAN, A. D. WILLIAMS, B. W. CHELLGREN, T. P. CREAMER and R. WETZEL. 2006. Oligoproline effects on polyglutamine conformation and aggregation. *J Mol Biol,* **355**(3), pp.524-35.

BIEMANN, K. 1988. Contributions of mass spectrometry to peptide and protein structure. *Biomed Environ Mass Spectrom,* **16**(1-12), pp.99-111.

BLEIHOLDER, C., S. CONTRERAS and M. T. BOWERS. 2013. A novel projection approximation algorithm for the fast and accurate computation of

molecular collision cross sections (IV). Application to polypeptides. *International Journal of Mass Spectrometry,* **354**, pp.275-280.

BLEIHOLDER, C., S. CONTRERAS, T. D. DO and M. T. BOWERS. 2013a. A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (II). Model parameterization and definition of empirical shape factors for proteins. *International Journal of Mass Spectrometry,* **345**, pp.89-96.

BLEIHOLDER, C., T. D. DO, C. WU, N. J. ECONOMOU, S. S. BERNSTEIN, S. K. BURATTO, J.-E. SHEA and M. T. BOWERS. 2013b. Ion Mobility Spectrometry Reveals the Mechanism of Amyloid Formation of Aβ(25–35) and Its Modulation by Inhibitors at the Molecular Level: Epigallocatechin Gallate and Scyllo-inositol. *Journal of the American Chemical Society,* **135**(45), pp.16926-16937.

BLEIHOLDER, C., T. WYTTENBACH and M. T. BOWERS. 2011. A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (I). Method. *International Journal of Mass Spectrometry,* **308**(1), pp.1-10.

BLOOM, G. S. 2014. Amyloid-beta and tau: the trigger and bullet in Alzheimer disease pathogenesis. *JAMA Neurol,* **71**(4), pp.505-8.

BONANOMI, M., C. VISENTIN, G. INVERNIZZI, P. TORTORA and M. E. REGONESI. 2015. The Toxic Effects of Pathogenic Ataxin-3 Variants in a Yeast Cellular Model. *PLoS One,* **10**(6), pe0129727.

BORYSIK, A. J., S. E. RADFORD and A. E. ASHCROFT. 2004. Co-populated conformational ensembles of beta2-microglobulin uncovered quantitatively by electrospray ionization mass spectrometry. *J Biol Chem,* **279**(26), pp.27069-77.

BRAUN, A. R., M. M. LACY, V. C. DUCAS, E. RHOADES and J. N. SACHS. 2014. alpha-Synuclein-induced membrane remodeling is driven by binding affinity, partition depth, and interleaflet order asymmetry. *J Am Chem Soc,* **136**(28), pp.9962-72.

BRIER, M. R., B. GORDON, K. FRIEDRICHSEN, J. MCCARTHY, A. STERN, J. CHRISTENSEN, C. OWEN, P. ALDEA, Y. SU, J. HASSENSTAB, N. J. CAIRNS, D. M. HOLTZMAN, A. M. FAGAN, J. C. MORRIS, T. L. S. BENZINGER and B. M. ANCES. 2016. Tau and Aβ imaging, CSF measures, and cognition in Alzheimer's disease. *Science Translational Medicine,* **8**(338), p338ra66.

BRODBELT, J. S. 2016. Ion Activation Methods for Peptides and Proteins. *Anal Chem,* **88**(1), pp.30-51.

BUCCIANTINI, M., E. GIANNONI, F. CHITI, F. BARONI, L. FORMIGLI, J. ZURDO, N. TADDEI, G. RAMPONI, C. M. DOBSON and M. STEFANI. 2002. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature,* **416**(6880), pp.507-11.

BUELL, A. K., C. M. DOBSON and T. P. KNOWLES. 2014. The physical chemistry of the amyloid phenomenon: thermodynamics and kinetics of filamentous protein aggregation. *Essays Biochem,* **56**, pp.11-39.

BURNETT, B., F. LI and R. N. PITTMAN. 2003. The polyglutamine neurodegenerative protein ataxin-3 binds polyubiquitylated proteins and has ubiquitin protease activity. *Hum Mol Genet,* **12**(23), pp.3195-205.

BUSH, A. I., W. H. PETTINGELL, G. MULTHAUP, M. D PARADIS, J. P. VONSATTEL, J. F. GUSELLA, K. BEYREUTHER, C. L. MASTERS and R. E. TANZI. 1994. Rapid induction of Alzheimer A beta amyloid formation by zinc. *Science,* **265**(5177), pp.1464-7.

BUSH, M. F., Z. HALL, K. GILES, J. HOYES, C. V. ROBINSON and B. T. RUOTOLO. 2010. Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology. *Anal Chem,* **82**(22), pp.9557-65.

BUTTERFIELD, D. A., A. M. SWOMLEY and R. SULTANA. 2013. Amyloid β-Peptide (1–42)-Induced Oxidative Stress in Alzheimer Disease: Importance in Disease Pathogenesis and Progression. *Antioxidants & Redox Signaling,* **19**(8), pp.823-835.

CAI, Z., B. ZHAO and A. RATKA. 2011. Oxidative stress and beta-amyloid protein in Alzheimer's disease. *Neuromolecular Med,* **13**(4), pp.223-50.

CAMERON, A. E. and D. F. EGGERS. 1948. An Ion ``Velocitron''. *Review of Scientific Instruments,* **19**(9), pp.605-607.

CAMERON, B. and G. E. LANDRETH. 2010. Inflammation, microglia, and Alzheimer's disease. *Neurobiol Dis,* **37**(3), pp.503-9.

CAO, J., J. E. BURKE and E. A. DENNIS. 2013. Using Hydrogen/Deuterium Exchange Mass Spectrometry to Define the Specific Interactions of the Phospholipase A2 Superfamily with Lipid Substrates, Inhibitors, and Membranes. *Journal of Biological Chemistry,* **288**(3), pp.1806-1813.

CAO, P. and D. P. RALEIGH. 2012. Analysis of the inhibition and remodeling of islet amyloid polypeptide amyloid fibers by flavanols. *Biochemistry,* **51**(13), pp.2670-83.

CARRASQUILLO, M. M., F. ZOU, V. S. PANKRATZ, S. L. WILCOX, L. MA, L. P. WALKER, S. G. YOUNKIN, C. S. YOUNKIN, L. H. YOUNKIN, G. D. BISCEGLIO, N. ERTEKIN-TANER, J. E. CROOK, D. W. DICKSON, R. C. PETERSEN, N. R. GRAFF-RADFORD and S. G. YOUNKIN. 2009. Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. *Nat Genet,* **41**(2), pp.192-8.

CARULLA, N., G. L. CADDY, D. R. HALL, J. ZURDO, M. GAIRI, M. FELIZ, E. GIRALT, C. V. ROBINSON and C. M. DOBSON. 2005. Molecular recycling within amyloid fibrils. *Nature,* **436**(7050), pp.554-8.

CASTILLO, G. M., W. LUKITO, T. N. WIGHT and A. D. SNOW. 1999. The sulfate moieties of glycosaminoglycans are critical for the enhancement of beta-amyloid protein fibril formation. *J Neurochem,* **72**(4), pp.1681-7.

CATTANEO, E., D. RIGAMONTI, D. GOFFREDO, C. ZUCCATO, F. SQUITIERI and S. SIPIONE. 2001. Loss of normal huntingtin function: new developments in Huntington's disease research. *Trends Neurosci,* **24**(3), pp.182-8.

CHAN, D. S., M. E. KAVANAGH, K. J. MCLEAN, A. W. MUNRO, D. MATAK-VINKOVIC, A. G. COYNE and C. ABELL. 2017. Effect of DMSO on Protein Structure and Interactions Assessed by Collision-Induced Dissociation and Unfolding. *Anal Chem*.

CHATANI, E., R. OHNISHI, T. KONUMA, K. SAKURAI, H. NAIKI and Y. GOTO. 2010. Pre-steady-state kinetic analysis of the elongation of amyloid fibrils of

beta(2)-microglobulin with tryptophan mutagenesis. *J Mol Biol,* **400**(5), pp.1057-66.

CHEN, X., S. QIN, S. CHEN, J. LI, L. LI, Z. WANG, Q. WANG, J. LIN, C. YANG and W. SHUI. 2015. A Ligand-observed Mass Spectrometry Approach Integrated into the Fragment Based Lead Discovery Pipeline. **5**, p8361.

CHITI, F. and C. M. DOBSON. 2006. Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem,* **75**, pp.333-66.

CHITI, F. and C. M. DOBSON. 2017. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu Rev Biochem,* **86**, pp.27-68.

CHITI, F., P. WEBSTER, N. TADDEI, A. CLARK, M. STEFANI, G. RAMPONI and C. M. DOBSON. 1999. Designing conditions for in vitro formation of amyloid protofilaments and fibrils. *Proc Natl Acad Sci U S A,* **96**(7), pp.3590-3594.

CHOW, M. K., A. M. ELLISDON, L. D. CABRITA and S. P. BOTTOMLEY. 2004. Polyglutamine expansion in ataxin-3 does not affect protein stability: implications for misfolding and disease. *J Biol Chem,* **279**(46), pp.47643-51.

CHOWDHURY, S. K., V. KATTA and B. T. CHAIT. 1990. Probing conformational changes in proteins by mass spectrometry. *Journal of the American Chemical Society,* **112**(24), pp.9012-9013.

CLOE, A. L., J. P. R. O. ORGEL, J. R. SACHLEBEN, R. TYCKO and S. C. MEREDITH. 2011. The Japanese Mutant Aβ (ΔE22-Aβ1−39) Forms Fibrils Instantaneously, with Low-Thioflavin T Fluorescence: Seeding of Wild-Type Aβ1−40 into Atypical Fibrils by ΔE22-Aβ1−39. *Biochemistry,* **50**(12), pp.2026-2039.

COELHO-CERQUEIRA, E., A. S. PINHEIRO and C. FOLLMER. 2014. Pitfalls associated with the use of Thioflavin-T to monitor anti-fibrillogenic activity. *Bioorg Med Chem Lett,* **24**(14), pp.3194-8.

COHEN, S. I., M. VENDRUSCOLO, C. M. DOBSON and T. P. KNOWLES. 2012. From macroscopic measurements to microscopic mechanisms of protein aggregation. *J Mol Biol,* **421**(2-3), pp.160-71.

COHEN, S. I. A., S. LINSE, L. M. LUHESHI, E. HELLSTRAND, D. A. WHITE, L. RAJAH, D. E. OTZEN, M. VENDRUSCOLO, C. M. DOBSON and T. P. J. KNOWLES. 2013. Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism. *Proceedings of the National Academy of Sciences,* **110**(24), pp.9758-9763.

COLE, H., M. PORRINI, R. MORRIS, T. SMITH, J. KALAPOTHAKIS, S. WEIDT, C. L. MACKAY, C. E. MACPHEE and P. E. BARRAN. 2015. Early stages of insulin fibrillogenesis examined with ion mobility mass spectrometry and molecular modelling. *Analyst,* **140**(20), pp.7000-11.

COLES, J. G., M. 1993. Orthogonal acceleration - a new direction for time-of-flight mass spectrometry: fast, sensitive mass analysis for continuous ion sources *Trends in Analytical Chemistry,* **12**(5).

COLLINS, C. J., B. SCHILLING, M. YOUNG, G. DOLLINGER and R. K. GUY. 2003. Isotopically labeled crosslinking reagents: resolution of mass degeneracy in the identification of crosslinked peptides. *Bioorg Med Chem Lett,* **13**(22), pp.4023-6.

COLVIN, M. T., R. SILVERS, Q. Z. NI, T. V. CAN, I. SERGEYEV, M. ROSAY, K. J. DONOVAN, B. MICHAEL, J. WALL, S. LINSE and R. G. GRIFFIN. 2016. Atomic Resolution Structure of Monomorphic Abeta42 Amyloid Fibrils. *J Am Chem Soc,* **138**(30), pp.9663-74.

COMPTON, L. A. and W. C. JOHNSON, JR. 1986. Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal Biochem,* **155**(1), pp.155-67.

CORDER, E. H., A. M. SAUNDERS, W. J. STRITTMATTER, D. E. SCHMECHEL, P. C. GASKELL, G. W. SMALL, A. D. ROSES, J. L. HAINES and M. A. PERICAK-VANCE. 1993. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science,* **261**(5123), pp.921-3.

COSTA MDO, C. and H. L. PAULSON. 2012. Toward understanding Machado-Joseph disease. *Prog Neurobiol,* **97**(2), pp.239-57.

CREMERS, C. M., D. KNOEFLER, S. GATES, N. MARTIN, J. U. DAHL, J. LEMPART, L. XIE, M. R. CHAPMAN, V. GALVAN, D. R. SOUTHWORTH and U. JAKOB. 2016. Polyphosphate: A Conserved Modifier of Amyloidogenic Processes. *Mol Cell,* **63**(5), pp.768-80.

CUFF, J. A. and G. J. BARTON. 1999. Evaluation and improvement of multiple sequence methods for protein secondary structure prediction. *Proteins,* **34**(4), pp.508-19.

CUKALEVSKI, R., X. YANG, G. MEISL, U. WEININGER, K. BERNFUR, B. FROHM, T. P. J. KNOWLES and S. LINSE. 2015. The A[small beta]40 and A[small beta]42 peptides self-assemble into separate homomolecular fibrils in binary mixtures but cross-react during primary nucleation. *Chemical Science,* **6**(7), pp.4215-4233.

CURRAIS, A., O. QUEHENBERGER, A. M ARMANDO, D. DAUGHERTY, P. MAHER and D. SCHUBERT. 2016. Amyloid proteotoxicity initiates an inflammatory response blocked by cannabinoids. **2**, p16012.

DARNELL, G., R. ORGEL JP FAU - PAHL, S. C. PAHL R FAU - MEREDITH and S. C. MEREDITH. 2007a. Flanking polyproline sequences inhibit beta-sheet structure in polyglutamine segments by inducing PPII-like helix structure. *J Mol Biol.,* **374**(3), pp.668-704.

DARNELL, G., J. P. ORGEL, R. PAHL and S. C. MEREDITH. 2007b. Flanking polyproline sequences inhibit beta-sheet structure in polyglutamine segments by inducing PPII-like helix structure. *J Mol Biol,* **374**(3), pp.688-704.

DARNELL, G. D., J. DERRYBERRY, J. W. KURUTZ and S. C. MEREDITH. 2009. Mechanism of cis-inhibition of polyQ fibrillation by polyP: PPII oligomers and the hydrophobic effect. *Biophys J,* **97**(8), pp.2295-305.

DAS, P., S.-G. KANG, S. TEMPLE and G. BELFORT. 2014. Interaction of Amyloid Inhibitor Proteins with Amyloid Beta Peptides: Insight from Molecular Dynamics Simulations. *PLoS One,* **9**(11), pe113041.

DAWSON, J. H. J. and M. GUILHAUS. 1989. Orthogonal-acceleration time-of-flight mass spectrometer. *Rapid Communications in Mass Spectrometry,* **3**(5), pp.155-159.

DEARBORN, A. D., J. S. WALL, N. CHENG, J. B. HEYMANN, A. V. KAJAVA, J. VARKEY, R. LANGEN and A. C. STEVEN. 2016. alpha-Synuclein Amyloid Fibrils with

Two Entwined, Asymmetrically Associated Protofibrils. *J Biol Chem,* **291**(5), pp.2310-8.

DELGADO, D. A., K. DOHERTY, Q. CHENG, H. KIM, D. XU, H. DONG, C. GREWER and W. QIANG. 2016. Distinct Membrane Disruption Pathways Are Induced by 40-Residue beta-Amyloid Peptides. *J Biol Chem,* **291**(23), pp.12233-44.

DEMPSTER, A. J. 1918. A new Method of Positive Ray Analysis. *Physical Review,* **11**(4), pp.316-325.

DEMURO, A., M. SMITH and I. PARKER. 2011. Single-channel Ca(2+) imaging implicates Abeta1-42 amyloid pores in Alzheimer's disease pathology. *J Cell Biol,* **195**(3), pp.515-24.

DEVINE, P. W. A., H. C. FISHER, A. N. CALABRESE, F. WHELAN, D. R. HIGAZI, J. R. POTTS, D. C. LOWE, S. E. RADFORD and A. E. ASHCROFT. 2017. Investigating the Structural Compaction of Biomolecules Upon Transition to the Gas-Phase Using ESI-TWIMS-MS. *J Am Soc Mass Spectrom*.

DILL, K. A. and H. S. CHAN. 1997. From Levinthal to pathways to funnels. *Nat Struct Biol,* **4**(1), pp.10-9.

DIVRY, P. F., M. 1927. Sur les propriétés optiques de l'amyloide. *C. R. Roc. Biol.,* **97**, pp.1808-1810.

DOBO, A. and I. A. KALTASHOV. 2001. Detection of multiple protein conformational ensembles in solution via deconvolution of charge-state distributions in ESI MS. *Anal Chem,* **73**(20), pp.4763-73.

DOLE, M., L. L. MACK, R. L. HINES, R. C. MOBLEY, L. D. FERGUSON and M. B. ALICE. 1968. Molecular Beams of Macroions. *The Journal of Chemical Physics,* **49**(5), pp.2240-2249.

DROZDETSKIY, A., C. COLE, J. PROCTER and G. J. BARTON. 2015. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res,* **43**(W1), pp.W389-94.

EANES, E. D. and G. G. GLENNER. 1968. X-ray diffraction studies on amyloid filaments. *J Histochem Cytochem,* **16**(11), pp.673-7.

EDEN, K., R. MORRIS, J. GILLAM, C. E. MACPHEE and R. J. ALLEN. 2015. Competition between primary nucleation and autocatalysis in amyloid fibril self-assembly. *Biophys J,* **108**(3), pp.632-43.

EFTEKHARZADEH, B., A. PIAI, G. CHIESA, D. MUNGIANU, J. GARCIA, R. PIERATTELLI, I. C. FELLI and X. SALVATELLA. 2016. Sequence Context Influences the Structure and Aggregation Behavior of a PolyQ Tract. *Biophys J,* **110**(11), pp.2361-2366.

EHRNHOEFER, D. E., J. BIESCHKE, A. BOEDDRICH, M. HERBST, L. MASINO, R. LURZ, S. ENGEMANN, A. PASTORE and E. E. WANKER. 2008. EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nat Struct Mol Biol,* **15**(6), pp.558-66.

EICHNER, T. and S. E. RADFORD. 2011. A diversity of assembly mechanisms of a generic amyloid fold. *Mol Cell,* **43**(1), pp.8-18.

ELLISDON, A. M., M. C. PEARCE and S. P. BOTTOMLEY. 2007. Mechanisms of ataxin-3 misfolding and fibril formation: kinetic analysis of a disease-associated polyglutamine protein. *J Mol Biol,* **368**(2), pp.595-605.

ELLISDON, A. M., B. THOMAS and S. P. BOTTOMLEY. 2006. The two-stage pathway of ataxin-3 fibrillogenesis involves a polyglutamine-independent step. *J Biol Chem,* **281**(25), pp.16888-96.

ENGEL, M. F. 2009. Membrane permeabilization by Islet Amyloid Polypeptide. *Chem Phys Lipids,* **160**(1), pp.1-10.

ENGEL, M. F., L. KHEMTEMOURIAN, C. C. KLEIJER, H. J. MEELDIJK, J. JACOBS, A. J. VERKLEIJ, B. DE KRUIJFF, J. A. KILLIAN and J. W. HOPPENER. 2008. Membrane damage by human islet amyloid polypeptide through fibril growth at the membrane. *Proc Natl Acad Sci U S A,* **105**(16), pp.6033-8.

EVERS, M. M., L. J. TOONEN and W. M. VAN ROON-MOM. 2014. Ataxin-3 protein and RNA toxicity in spinocerebellar ataxia type 3: current insights and emerging therapeutic strategies. *Mol Neurobiol,* **49**(3), pp.1513-31.

FALSONE, S. F., N. H. MEYER, E. SCHRANK, G. LEITINGER, C. L. PHAM, M. T. FODERO-TAVOLETTI, M. HOLMBERG, M. DULLE, B. SCICLUNA, B. GESSLBAUER, H. M. RUCKERT, G. E. WAGNER, D. A. MERLE, E. A. NOLLEN, A. J. KUNGL, A. F. HILL, R. CAPPAI and K. ZANGGER. 2012. SERF protein is a direct modifier of amyloid fiber assembly. *Cell Rep,* **2**(2), pp.358-71.

FAN, H.-C., L.-I. HO, C.-S. CHI, S.-J. CHEN, G.-S. PENG, T.-M. CHAN, S.-Z. LIN and H.-J. HARN. 2014. Polyglutamine (PolyQ) Diseases: Genetics to Treatments. *Cell Transplantation,* **23**(4-5), pp.441-458.

FEIGE, M. J., S. GROSCURTH, M. MARCINOWSKI, Z. T. YEW, V. TRUFFAULT, E. PACI, H. KESSLER and J. BUCHNER. 2008. The structure of a folding intermediate provides insight into differences in immunoglobulin amyloidogenicity. *Proceedings of the National Academy of Sciences,* **105**(36), pp.13373-13378.

FENN, J. B., M. MANN, C. K. MENG, S. F. WONG and C. M. WHITEHOUSE. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science,* **246**(4926), p64.

FERNANDEZ, C. O., W. HOYER, M. ZWECKSTETTER, E. A. JARES-ERIJMAN, V. SUBRAMANIAM, C. GRIESINGER and T. M. JOVIN. 2004. NMR of alpha-synuclein-polyamine complexes elucidates the mechanism and kinetics of induced aggregation. *EMBO J,* **23**(10), pp.2039-46.

FITZPATRICK, A. W., G. T. DEBELOUCHINA, M. J. BAYRO, D. K. CLARE, M. A. CAPORINI, V. S. BAJAJ, C. P. JARONIEC, L. WANG, V. LADIZHANSKY, S. A. MULLER, C. E. MACPHEE, C. A. WAUDBY, H. R. MOTT, A. DE SIMONE, T. P. KNOWLES, H. R. SAIBIL, M. VENDRUSCOLO, E. V. ORLOVA, R. G. GRIFFIN and C. M. DOBSON. 2013. Atomic structure and hierarchical assembly of a cross-beta amyloid fibril. *Proc Natl Acad Sci U S A,* **110**(14), pp.5468-73.

FITZPATRICK, A. W. P., B. FALCON, S. HE, A. G. MURZIN, G. MURSHUDOV, H. J. GARRINGER, R. A. CROWTHER, B. GHETTI, M. GOEDERT and S. H. W. SCHERES. 2017. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature,* **547**(7662), pp.185-190.

FOLCH, J., D. PETROV, M. ETTCHETO, S. ABAD, #XE1, NCHEZ-L, #XF3, E. PEZ, GARC, #XED, M. L. A, J. OLLOQUEQUI, C. BEAS-ZARATE, C. AULADELL and A. CAMINS. 2016. Current Research Therapeutic Strategies for Alzheimer’s Disease Treatment. *Neural Plasticity,* **2016**, p15.

FORMAN, M. D., R. F. STACK, P. S. MASTERS, C. R. HAUER and S. M. BAXTER. 1998. High level, context dependent misincorporation of lysine for arginine in Saccharomyces cerevisiae a1 homeodomain expressed in Escherichia coli. *Protein Sci,* **7**(2), pp.500-3.

FREDRICKSON, W. R. 1946. Proceedings of the American Physical Society. *Physical Review,* **69**(11-12), pp.674-674.

FRONTZEK, K., M. I. LUTZ, A. AGUZZI, G. G. KOVACS and H. BUDKA. 2016. Amyloid-beta pathology and cerebral amyloid angiopathy are frequent in iatrogenic Creutzfeldt-Jakob disease after dural grafting. *Swiss Med Wkly,* **146**, pw14287.

GADE MALMOS, K., L. M. BLANCAS-MEJIA, B. WEBER, J. BUCHNER, M. RAMIREZ-ALVARADO, H. NAIKI and D. OTZEN. 2017. ThT 101: a primer on the use of thioflavin T to investigate amyloid formation. *Amyloid,* **24**(1), pp.1-16.

GALANAKIS, P. A., F. N. BAZOTI, J. BERGQUIST, K. MARKIDES, G. A. SPYROULIAS and A. TSARBOPOULOS. 2011. Study of the interaction between the amyloid beta peptide (1-40) and antioxidant compounds by nuclear magnetic resonance spectroscopy. *Biopolymers,* **96**(3), pp.316-27.

GARAI, K. and C. FRIEDEN. 2013. Quantitative analysis of the time course of Abeta oligomerization and subsequent growth steps using tetramethylrhodamine-labeled Abeta. *Proc Natl Acad Sci U S A,* **110**(9), pp.3321-6.

GEDDES, A. J., K. D. PARKER, E. D. ATKINS and E. BEIGHTON. 1968. "Cross-beta" conformation in proteins. *J Mol Biol,* **32**(2), pp.343-58.

GILES, K., S. D. PRINGLE, K. R. WORTHINGTON, D. LITTLE, J. L. WILDGOOSE and R. H. BATEMAN. 2004. Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun Mass Spectrom,* **18**(20), pp.2401-14.

GILLAM, J. E. and C. E. MACPHEE. 2013. Modelling amyloid fibril formation kinetics: mechanisms of nucleation and growth. *J Phys Condens Matter,* **25**(37), p373101.

GIRI, M., A. SHAH, B. UPRETI and J. C. RAI. 2017. Unraveling the genes implicated in Alzheimer's disease. *Biomed Rep,* **7**(2), pp.105-114.

GLABE, C. G. 2008. Structural classification of toxic amyloid oligomers. *J Biol Chem,* **283**(44), pp.29639-43.

GLENNER, G. G. and C. W. WONG. 1984. Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem Biophys Res Commun,* **120**(3), pp.885-90.

GOEHLER, H., M. LALOWSKI, U. STELZL, S. WAELTER, M. STROEDICKE, U. WORM, A. DROEGE, K. S. LINDENBERG, M. KNOBLICH, C. HAENIG, M. HERBST, J. SUOPANKI, E. SCHERZINGER, C. ABRAHAM, B. BAUER, R. HASENBANK, A. FRITZSCHE, A. H. LUDEWIG, K. BUSSOW, S. H. COLEMAN, C. A. GUTEKUNST, B. G. LANDWEHRMEYER, H. LEHRACH and E. E. WANKER. 2004. A protein interaction network links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease. *Mol Cell,* **15**(6), pp.853-65.

GOLDSBURY, C. S., S. WIRTZ, S. A. MULLER, S. SUNDERJI, P. WICKI, U. AEBI and P. FREY. 2000. Studies on the in vitro assembly of a beta 1-40: implications

for the search for a beta fibril formation inhibitors. *J Struct Biol,* **130**(2-3), pp.217-31.

GOLDSCHMIDT, L., P. K. TENG, R. RIEK and D. EISENBERG. 2010. Identifying the amylome, proteins capable of forming amyloid-like fibrils. *Proc Natl Acad Sci U S A,* **107**(8), pp.3487-92.

GONG, Y., L. CHANG, K. L. VIOLA, P. N. LACOR, M. P. LAMBERT, C. E. FINCH, G. A. KRAFFT and W. L. KLEIN. 2003. Alzheimer's disease-affected brain: Presence of oligomeric Aβ ligands (ADDLs) suggests a molecular basis for reversible memory loss. *Proceedings of the National Academy of Sciences,* **100**(18), pp.10417-10422.

GOREVIC, P. D., E. M. CASTANO, R. SARMA and B. FRANGIONE. 1987. Ten to fourteen residue peptides of Alzheimer's disease protein are sufficient for amyloid fibril formation and its characteristic x-ray diffraction pattern. *Biochem Biophys Res Commun,* **147**(2), pp.854-62.

GOTI, D., S. M. KATZEN, J. MEZ, N. KURTIS, J. KILUK, L. BEN-HAIEM, N. A. JENKINS, N. G. COPELAND, A. KAKIZUKA, A. H. SHARP, C. A. ROSS, P. R. MOUTON and V. COLOMER. 2004. A mutant ataxin-3 putative-cleavage fragment in brains of Machado-Joseph disease patients and transgenic mice is cytotoxic above a critical concentration. *J Neurosci,* **24**(45), pp.10266-79.

GOTZE, M., J. PETTELKAU, S. SCHAKS, K. BOSSE, C. H. IHLING, F. KRAUTH, R. FRITZSCHE, U. KUHN and A. SINZ. 2012. StavroX--a software for analyzing crosslinked products in protein interaction studies. *J Am Soc Mass Spectrom,* **23**(1), pp.76-87.

GOURE, W. F., G. A. KRAFFT, J. JERECIC and F. HEFTI. 2014. Targeting the proper amyloid-beta neuronal toxins: a path forward for Alzheimer's disease immunotherapeutics. *Alzheimers Res Ther,* **6**(4), p42.

GREENFIELD, N. J. 2006. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc,* **1**(6), pp.2876-90.

GRIMM, M., T. ZIMNIAK, A. KAHRAMAN and F. HERZOG. 2015. xVis: a web server for the schematic visualization and interpretation of crosslink-derived spatial restraints. *Nucleic Acids Res,* **43**(W1), pp.W362-9.

GRIMSLEY, G. R. and C. N. PACE. 2004. Spectrophotometric determination of protein concentration. *Curr Protoc Protein Sci,* **Chapter 3**, pUnit 3 1.

GUIJARRO, J. I., M. SUNDE, J. A. JONES, I. D. CAMPBELL and C. M. DOBSON. 1998. Amyloid fibril formation by an SH3 domain. *Proc Natl Acad Sci U S A,* **95**(8), pp.4224-8.

GUILHAUS, M., D. SELBY and V. MLYNSKI. 2000. Orthogonal acceleration time-of-flight mass spectrometry. *Mass Spectrom Rev,* **19**(2), pp.65-107.

HAACKE, A., F. U. HARTL and P. BREUER. 2007. Calpain inhibition is sufficient to suppress aggregation of polyglutamine-expanded ataxin-3. *J Biol Chem,* **282**(26), pp.18851-6.

HABICHT, G., C. HAUPT, R. P. FRIEDRICH, P. HORTSCHANSKY, C. SACHSE, J. MEINHARDT, K. WIELIGMANN, G. P. GELLERMANN, M. BRODHUN, J. GOTZ, K. J. HALBHUBER, C. ROCKEN, U. HORN and M. FANDRICH. 2007. Directed selection of a conformational antibody domain that prevents mature amyloid fibril formation by stabilizing Abeta protofibrils. *Proc Natl Acad Sci U S A,* **104**(49), pp.19232-7.

HALL, Z., H. HERNANDEZ, J. A. MARSH, S. A. TEICHMANN and C. V. ROBINSON. 2013. The role of salt bridges, charge density, and subunit flexibility in determining disassembly routes of protein complexes. *Structure,* **21**(8), pp.1325-37.

HAMMARSTROM, P., R. L. WISEMAN, E. T. POWERS and J. W. KELLY. 2003. Prevention of transthyretin amyloid disease by changing protein misfolding energetics. *Science,* **299**(5607), pp.713-6.

HAN, L., S. J. HYUNG and B. T. RUOTOLO. 2012. Bound cations significantly stabilize the structure of multiprotein complexes in the gas phase. *Angew Chem Int Ed Engl,* **51**(23), pp.5692-5.

HAROLD, D., R. ABRAHAM, P. HOLLINGWORTH, R. SIMS, A. GERRISH, M. L. HAMSHERE, J. S. PAHWA, V. MOSKVINA, K. DOWZELL, A. WILLIAMS, N. JONES, C. THOMAS, A. STRETTON, A. R. MORGAN, S. LOVESTONE, J. POWELL, P. PROITSI, M. K. LUPTON, C. BRAYNE, D. C. RUBINSZTEIN, M. GILL, B. LAWLOR, A. LYNCH, K. MORGAN, K. S. BROWN, P. A. PASSMORE, D. CRAIG, B. MCGUINNESS, S. TODD, C. HOLMES, D. MANN, A. D. SMITH, S. LOVE, P. G. KEHOE, J. HARDY, S. MEAD, N. FOX, M. ROSSOR, J. COLLINGE, W. MAIER, F. JESSEN, B. SCHURMANN, R. HEUN, H. VAN DEN BUSSCHE, I. HEUSER, J. KORNHUBER, J. WILTFANG, M. DICHGANS, L. FROLICH, H. HAMPEL, M. HULL, D. RUJESCU, A. M. GOATE, J. S. KAUWE, C. CRUCHAGA, P. NOWOTNY, J. C. MORRIS, K. MAYO, K. SLEEGERS, K. BETTENS, S. ENGELBORGHS, P. P. DE DEYN, C. VAN BROECKHOVEN, G. LIVINGSTON, N. J. BASS, H. GURLING, A. MCQUILLIN, R. GWILLIAM, P. DELOUKAS, A. AL-CHALABI, C. E. SHAW, M. TSOLAKI, A. B. SINGLETON, R. GUERREIRO, T. W. MUHLEISEN, M. M. NOTHEN, S. MOEBUS, K. H. JOCKEL, N. KLOPP, H. E. WICHMANN, M. M. CARRASQUILLO, V. S. PANKRATZ, S. G. YOUNKIN, P. A. HOLMANS, M. O'DONOVAN, M. J. OWEN and J. WILLIAMS. 2009. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet,* **41**(10), pp.1088-93.

HAYES, J. J., L. KAM and T. D. TULLIUS. 1990. Footprinting protein-DNA complexes with gamma-rays. *Methods Enzymol,* **186**, pp.545-9.

HERVAS, R., J. OROZ, A. GALERA-PRAT, O. GONI, A. VALBUENA, A. M. VERA, A. GOMEZ-SICILIA, F. LOSADA-URZAIZ, V. N. UVERSKY, M. MENENDEZ, D. V. LAURENTS, M. BRUIX and M. CARRION-VAZQUEZ. 2012. Common features at the start of the neurodegeneration cascade. *PLoS Biol,* **10**(5), pe1001335.

HILTON, G. R. and J. L. P. BENESCH. 2012. Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *Journal of The Royal Society Interface,* **9**(70), p801.

HOFFMANN, W., G. VON HELDEN and K. PAGEL. 2017. Ion mobility-mass spectrometry and orthogonal gas-phase techniques to study amyloid formation and inhibition. *Curr Opin Struct Biol,* **46**, pp.7-15.

HOFMANN, H., D. NETTELS and B. SCHULER. 2013. Single-molecule spectroscopy of the unexpected collapse of an unfolded protein at low pH. *J Chem Phys,* **139**(12), p121930.

HOFMEISTER, F. 1888. Zur Lehre von der Wirkung der Salze. *Arch. Exp. Pathol. Pharmakol.,* **24**, pp.247-260.

HONG, Y. T., T. VEENITH, D. DEWAR, J. G. OUTTRIM, V. MANI, C. WILLIAMS, S. PIMLOTT, P. J. HUTCHINSON, A. TAVARES, R. CANALES, C. A. MATHIS, W. E. KLUNK, F. I. AIGBIRHIO, J. P. COLES, J. C. BARON, J. D. PICKARD, T. D. FRYER, W. STEWART and D. K. MENON. 2014. Amyloid imaging with carbon 11-labeled Pittsburgh compound B for traumatic brain injury. *JAMA Neurol,* **71**(1), pp.23-31.

HORTSCHANSKY, P., V. SCHROECKH, T. CHRISTOPEIT, G. ZANDOMENEGHI and M. FÄNDRICH. 2005. The aggregation kinetics of Alzheimer's β-amyloid peptide is controlled by stochastic nucleation. *Protein Sci,* **14**(7), pp.1753-1759.

HOYER, W., D. CHERNY, V. SUBRAMANIAM and T. M. JOVIN. 2004. Impact of the acidic C-terminal region comprising amino acids 109-140 on alpha-synuclein aggregation in vitro. *Biochemistry,* **43**(51), pp.16233-42.

HUBENER, J., J. J. WEBER, C. RICHTER, L. HONOLD, A. WEISS, F. MURAD, P. BREUER, U. WULLNER, P. BELLSTEDT, F. PAQUET-DURAND, J. TAKANO, T. C. SAIDO, O. RIESS and H. P. NGUYEN. 2013. Calpain-mediated ataxin-3 cleavage in the molecular pathogenesis of spinocerebellar ataxia type 3 (SCA3). *Hum Mol Genet,* **22**(3), pp.508-18.

HUNG, L. W., G. D. CICCOTOSTO, E. GIANNAKIS, D. J. TEW, K. PEREZ, C. L. MASTERS, R. CAPPAI, J. D. WADE and K. J. BARNHAM. 2008. Amyloid-beta peptide (Abeta) neurotoxicity is modulated by the rate of peptide aggregation: Abeta dimers and trimers correlate with neurotoxicity. *J Neurosci,* **28**(46), pp.11950-8.

IDA, N., T. HARTMANN, J. PANTEL, J. SCHRODER, R. ZERFASS, H. FORSTL, R. SANDBRINK, C. L. MASTERS and K. BEYREUTHER. 1996. Analysis of heterogeneous A4 peptides in human cerebrospinal fluid and blood by a newly developed sensitive Western blot assay. *J Biol Chem,* **271**(37), pp.22908-14.

INTERNATIONAL, A. S. D. 2015. World Alzheimer's reports. [online]. [Accessed 25.07.2017]. Available from: https://www.alz.co.uk/research/world-report-2016.

INVERNIZZI, G., M. LAMBRUGHI, M. E. REGONESI, P. TORTORA and E. PAPALEO. 2013. The conformational ensemble of the disordered and aggregation-protective 182-291 region of ataxin-3. *Biochim. Biophys. Acta.,* **1830**(11), pp.5236-47.

IRIBARNE, J. V. and B. A. THOMSON. 1976. On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics,* **64**(6), pp.2287-2294.

JAHN, T. R. and S. E. RADFORD. 2008. Folding versus aggregation: polypeptide conformations on competing pathways. *Arch Biochem Biophys,* **469**(1), pp.100-17.

JAIN, S. and J. B. UDGAONKAR. 2010. Salt-induced modulation of the pathway of amyloid fibril formation by the mouse prion protein. *Biochemistry,* **49**(35), pp.7615-24.

JAUNMUKTANE, Z., S. MEAD, M. ELLIS, J. D. WADSWORTH, A. J. NICOLL, J. KENNY, F. LAUNCHBURY, J. LINEHAN, A. RICHARD-LOENDT, A. S. WALKER, P. RUDGE, J. COLLINGE and S. BRANDNER. 2015. Evidence for human

transmission of amyloid-beta pathology and cerebral amyloid angiopathy. *Nature,* **525**(7568), pp.247-50.

JENNINGS, K. R. 1968. Collision-induced decompositions of aromatic molecular ions. *International Journal of Mass Spectrometry and Ion Physics,* **1**(3), pp.227-235.

JUCKER, M. and L. C. WALKER. 2011. Pathogenic protein seeding in Alzheimer disease and other neurodegenerative disorders. *Ann Neurol,* **70**(4), pp.532-40.

JUNG, J., K. XU, D. LESSING and N. M. BONINI. 2009. Preventing Ataxin-3 protein cleavage mitigates degeneration in a Drosophila model of SCA3. *Hum Mol Genet,* **18**(24), pp.4843-52.

KALKHOF, S. and A. SINZ. 2008. Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters. *Anal Bioanal Chem,* **392**(1-2), pp.305-12.

KALTASHOV, I. A. and A. MOHIMEN. 2005. Estimates of protein surface areas in solution by electrospray ionization mass spectrometry. *Anal Chem,* **77**(16), pp.5370-9.

KARAS, M., D. BACHMANN, U. BAHR and F. HILLENKAMP. 1987. Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International Journal of Mass Spectrometry and Ion Processes,* **78**, pp.53-68.

KARAS, M. and F. HILLENKAMP. 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry,* **60**(20), pp.2299-2301.

KAWAGUCHI, Y., T. OKAMOTO, M. TANIWAKI, M. AIZAWA, M. INOUE, S. KATAYAMA, H. KAWAKAMI, S. NAKAMURA, M. NISHIMURA, I. AKIGUCHI and ET AL. 1994. CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nat Genet,* **8**(3), pp.221-8.

KAYED, R., E. HEAD, F. SARSOZA, T. SAING, C. W. COTMAN, M. NECULA, L. MARGOL, J. WU, L. BREYDO, J. L. THOMPSON, S. RASOOL, T. GURLO, P. BUTLER and C. G. GLABE. 2007. Fibril specific, conformation dependent antibodies recognize a generic epitope common to amyloid fibrils and fibrillar oligomers that is absent in prefibrillar oligomers. *Mol Neurodegener,* **2**, p18.

KAYED, R., E. HEAD, J. L. THOMPSON, T. M. MCINTIRE, S. C. MILTON, C. W. COTMAN and C. G. GLABE. 2003. Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science,* **300**(5618), pp.486-9.

KAYED, R. and C. A. LASAGNA-REEVES. 2013. Molecular mechanisms of amyloid oligomers toxicity. *J Alzheimers Dis,* **33 Suppl 1**, pp.S67-78.

KEBARLE, P. 2000. A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *J Mass Spectrom,* **35**(7), pp.804-17.

KIM, Y. E., M. S. HIPP, A. BRACHER, M. HAYER-HARTL and F. U. HARTL. 2013. Molecular chaperone functions in protein folding and proteostasis. *Annu Rev Biochem,* **82**, pp.323-55.

KNAPMAN, T. W., J. T. BERRYMAN, I. CAMPUZANO, S. A. HARRIS and A. E. ASHCROFT. 2010. Considerations in experimental and theoretical collision

cross-section measurements of small molecules using travelling wave ion mobility spectrometry-mass spectrometry. *International Journal of Mass Spectrometry,* **298**(1), pp.17-23.

KNOWLES, T. P., M. VENDRUSCOLO and C. M. DOBSON. 2014. The amyloid state and its association with protein misfolding diseases. *Nat Rev Mol Cell Biol,* **15**(6), pp.384-96.

KONERMANN, L. 2007. A minimalist model for exploring conformational effects on the electrospray charge state distribution of proteins. *J Phys Chem B,* **111**(23), pp.6534-43.

KONERMANN, L., E. AHADI, A. D. RODRIGUEZ and S. VAHIDI. 2013. Unraveling the mechanism of electrospray ionization. *Anal Chem,* **85**(1), pp.2-9.

KONERMANN, L., X. TONG and Y. PAN. 2008. Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches. *J Mass Spectrom,* **43**(8), pp.1021-36.

KRUSEMARK, C. J., B. L. FREY, P. J. BELSHAW and L. M. SMITH. 2009. Modifying the Charge State Distribution of Proteins in Electrospray Ionization Mass Spectrometry by Chemical Derivatization. *J Am Soc Mass Spectrom,* **20**(9), pp.1617-1625.

KRZYWINSKI, M. and N. ALTMAN. 2014. Visualizing samples with box plots. *Nat Methods,* **11**(2), pp.119-20.

KUMAR, S., M. BIROL, D. E. SCHLAMADINGER, S. P. WOJCIK, E. RHOADES and A. D. MIRANKER. 2016. Foldamer-mediated manipulation of a pre-amyloid toxin. **7**, p11412.

KUNZ, W., J. HENLE and B. W. NINHAM. 2004. 'Zur Lehre von der Wirkung der Salze' (about the science of the effect of salts): Franz Hofmeister's historical papers. *Current Opinion in Colloid & Interface Science,* **9**(1), pp.19-37.

LA SPADA, A. 1999 *Spinal and Bulbar Muscular Atrophy.* [Accessed 01/09/17]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1333/.

LAI, A. L., E. M. CLERICO, M. E. BLACKBURN, N. A. PATEL, C. V. ROBINSON, P. P. BORBAT, J. H. FREED and L. M. GIERASCH. 2017. Key features of an Hsp70 chaperone allosteric landscape revealed by ion-mobility native mass spectrometry and double electron-electron resonance. *J Biol Chem,* **292**(21), pp.8773-8785.

LARNER, A. J. and M. DORAN. 2006. Clinical phenotypic heterogeneity of Alzheimer's disease associated with mutations of the presenilin-1 gene. *J Neurol,* **253**(2), pp.139-58.

LEBENDIKER, M. and T. DANIELI. 2011. Purification of proteins fused to maltose-binding protein. *Methods Mol Biol,* **681**, pp.281-93.

LEE, V. M., B. J. BALIN, L. OTVOS, JR. and J. Q. TROJANOWSKI. 1991. A68: a major subunit of paired helical filaments and derivatized forms of normal Tau. *Science,* **251**(4994), pp.675-8.

LEITNER, A., M. FAINI, F. STENGEL and R. AEBERSOLD. 2016. Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends Biochem Sci,* **41**(1), pp.20-32.

LENEY, A. C. and A. J. R. HECK. 2017. Native Mass Spectrometry: What is in the Name? *J Am Soc Mass Spectrom,* **28**(1), pp.5-13.

LESNÉ, S., M. T. KOH, L. KOTILINEK, R. KAYED, C. G. GLABE, A. YANG, M. GALLAGHER and K. H. ASHE. 2006. A specific amyloid-[beta] protein assembly in the brain impairs memory. *Nature,* **440**(7082), pp.352-357.

LEVINE, H., 3RD. 1999. Quantification of beta-sheet amyloid fibril structures with thioflavin T. *Methods Enzymol,* **309**, pp.274-84.

LEVINTHAL, C. 1968. Are there pathways for protein folding? *Extrait du Journal de Chimie Physique,* **65**(1), p44.

LEVINTHAL, C. 1969. How to Fold Graciously. *In: Proceedings of a Meeting held at Allerton House, Monticello, Illinoi,* Monticello, Illinoi. University of Illinious Press.

LEVITAN, K., D. CHEREAU, S. I. COHEN, T. P. KNOWLES, C. M. DOBSON, A. L. FINK, J. P. ANDERSON, J. M. GOLDSTEIN and G. L. MILLHAUSER. 2011. Conserved C-terminal charge exerts a profound influence on the aggregation rate of alpha-synuclein. *J Mol Biol,* **411**(2), pp.329-33.

LI, J., C. SANTAMBROGIO, S. BROCCA, G. ROSSETTI, P. CARLONI and R. GRANDORI. 2016. Conformational effects in protein electrospray-ionization mass spectrometry. *Mass Spectrom Rev,* **35**(1), pp.111-22.

LIAO, S., W. ZHANG, K. FAN, K. YE, X. ZHANG, J. ZHANG, C. XU and X. TU. 2013. Ionic strength-dependent conformations of a ubiquitin-like small archaeal modifier protein (SAMP2) from Haloferax volcanii. *Sci Rep,* **3**, p2136.

LIBEU, C. P., K. S. POKSAY, V. JOHN and D. E. BREDESEN. 2011. Structural and functional alterations in amyloid-beta precursor protein induced by amyloid-beta peptides. *J Alzheimers Dis,* **25**(3), pp.547-66.

LIMAN, J., S. DEEG, A. VOIGT, H. VOSSFELDT, C. P. DOHM, A. KARCH, J. WEISHAUPT, J. B. SCHULZ, M. BAHR and P. KERMER. 2014. CDK5 protects from caspase-induced Ataxin-3 cleavage and neurodegeneration. *J Neurochem,* **129**(6), pp.1013-23.

LIN, H., R. BHATIA and R. LAL. 2001. Amyloid beta protein forms ion channels: implications for Alzheimer's disease pathophysiology. *FASEB J,* **15**(13), pp.2433-44.

LIN, H. K., J. C. BOATZ, I. E. KRABBENDAM, R. KODALI, Z. HOU, R. WETZEL, A. M. DOLGA, M. A. POIRIER and P. C. A. VAN DER WEL. 2017. Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core. *Nat Commun.* .

LIU, B., A. MOLONEY, S. MEEHAN, K. MORRIS, S. E. THOMAS, L. C. SERPELL, R. HIDER, S. J. MARCINIAK, D. A. LOMAS and D. C. CROWTHER. 2011a. Iron promotes the toxicity of amyloid beta peptide by impeding its ordered aggregation. *J Biol Chem,* **286**(6), pp.4248-56.

LIU, H., X. LI, G. NING, S. ZHU, X. MA, X. LIU, C. LIU, M. HUANG, I. SCHMITT, U. WULLNER, Y. NIU, C. GUO, Q. WANG and T. S. TANG. 2016. The Machado-Joseph Disease Deubiquitinase Ataxin-3 Regulates the Stability and Apoptotic Function of p53. *PLoS Biol,* **14**(11), pe2000733.

LIU, Y., L. H. HO, J. A. CARVER and T. L. PUKALA. 2011b. Ion Mobility Mass Spectrometry Studies of the Inhibition of Alpha Synuclein Amyloid Fibril

Formation by (–)-Epigallocatechin-3-Gallate. *Australian Journal of Chemistry,* **64**(1), pp.36-40.

LIU, Y., J. S. SHARP, D. H. DO, R. A. KAHN, H. SCHWALBE, F. BUHR and J. H. PRESTEGARD. 2017. Mistakes in translation: Reflections on mechanism. *PLoS One,* **12**(6), pe0180566.

LONG, Z., B. TANG and H. JIANG. 2014. Alleviating neurodegeneration in Drosophila models of PolyQ diseases. *Cerebellum & Ataxias,* **1**, p9.

LOVEDAY, S. M., X. L. WANG, M. A. RAO, S. G. ANEMA and H. SINGH. 2012. β-Lactoglobulin nanofibrils: Effect of temperature on fibril formation kinetics, fibril morphology and the rheological properties of fibril dispersions. *Food Hydrocolloids,* **27**(1), pp.242-249.

LU, C., Y. GUO, J. YAN, Z. LUO, H. B. LUO, M. YAN, L. HUANG and X. LI. 2013a. Design, synthesis, and evaluation of multitarget-directed resveratrol derivatives for the treatment of Alzheimer's disease. *J Med Chem,* **56**(14), pp.5843-59.

LU, J. X., W. QIANG, W. M. YAU, C. D. SCHWIETERS, S. C. MEREDITH and R. TYCKO. 2013b. Molecular structure of beta-amyloid fibrils in Alzheimer's disease brain tissue. *Cell,* **154**(6), pp.1257-68.

LUO, S. and D. C. RUBINSZTEIN. 2009. Huntingtin promotes cell survival by preventing Pak2 cleavage. *J Cell Sci,* **122**(Pt 6), pp.875-85.

LUPTON, C. J., D. L. STEER, P. L. WINTRODE, S. P. BOTTOMLEY, V. A. HUGHES and A. M. ELLISDON. 2015. Enhanced molecular mobility of ordinarily structured regions drives polyglutamine disease. *J Biol Chem,* **290**(40), pp.24190-200.

MA, T., M. S. TAN, J. T. YU and L. TAN. 2014. Resveratrol as a therapeutic agent for Alzheimer's disease. *Biomed Res Int,* **2014**, p350516.

MADINE, J., E. JACK, P. G. STOCKLEY, S. E. RADFORD, L. C. SERPELL and D. A. MIDDLETON. 2008. Structural insights into the polymorphism of amyloid-like fibrils formed by region 20-29 of amylin revealed by solid-state NMR and X-ray fiber diffraction. *J Am Chem Soc,* **130**(45), pp.14990-5001.

MADLER, S., C. BICH, D. TOUBOUL and R. ZENOBI. 2009. Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *J Mass Spectrom,* **44**(5), pp.694-706.

MAHOOD, R. A. 2015. *Selection of RNA aptamers and their recognition of amyloid assemblies*. PhD thesis, University of Leeds.

MAKIN, O. S., E. ATKINS, P. SIKORSKI, J. JOHANSSON and L. C. SERPELL. 2005. Molecular basis for amyloid fibril formation and stability. *Proc Natl Acad Sci U S A,* **102**(2), pp.315-20.

MAKIN, O. S. and L. C. SERPELL. 2005. Structures for amyloid fibrils. *FEBS J,* **272**(23), pp.5950-61.

MAMYRIN, B. A. K., V. I.; SHMIKK, D. V.; ZAGULIN, V. A. 1973. The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Soviet Physics JETP,* **37**, p45.

MARCON, G., G. GIACCONE, C. CUPIDI, M. BALESTRIERI, C. A. BELTRAMI, N. FINATO, P. BERGONZI, S. SORBI, O. BUGIANI and F. TAGLIAVINI. 2004. Neuropathological and clinical phenotype of an Italian Alzheimer family

with M239V mutation of presenilin 2 gene. *J Neuropathol Exp Neurol,* **63**(3), pp.199-209.

MARKLUND, E. G., M. T. DEGIACOMI, C. V. ROBINSON, A. J. BALDWIN and J. L. BENESCH. 2015. Collision cross sections for structural proteomics. *Structure,* **23**(4), pp.791-9.

MARTÍNEZ, L. 2014. Introducing the Levinthal's Protein Folding Paradox and Its Solution. *Journal of Chemical Education,* **91**(11), pp.1918-1923.

MASINO, L., V. MUSI, R. P. MENON, P. FUSI, G. KELLY, T. A. FRENKIEL, Y. TROTTIER and A. PASTORE. 2003. Domain architecture of the polyglutamine protein ataxin-3: a globular domain followed by a flexible tail. *FEBS Lett,* **549**(1-3), pp.21-5.

MASINO, L., A. NICASTRO G FAU - DE SIMONE, L. DE SIMONE A FAU - CALDER, J. CALDER L FAU - MOLLOY, A. MOLLOY J FAU - PASTORE and A. PASTORE. 2011. The Josephin domain determines the morphological and mechanical properties of ataxin-3 fibrils. *Biophys J,* **100**(8), pp.2033-42.

MASON, E. A. and H. W. SCHAMP. 1958. Mobility of gaseous lons in weak electric fields. *Annals of Physics,* **4**(3), pp.233-270.

MASTERS, C. L., R. BATEMAN, K. BLENNOW, C. C. ROWE, R. A. SPERLING and J. L. CUMMINGS. 2015. Alzheimer's disease. *Nat Rev Dis Primers,* **1**, p15056.

MATSUI, T., M. INGELSSON, H. FUKUMOTO, K. RAMASAMY, H. KOWA, M. P. FROSCH, M. C. IRIZARRY and B. T. HYMAN. 2007. Expression of APP pathway mRNAs and proteins in Alzheimer's disease. *Brain Res,* **1161**, pp.116-23.

MCCLENDON, S., C. C. ROSPIGLIOSI and D. ELIEZER. 2009. Charge neutralization and collapse of the C-terminal tail of alpha-synuclein at low pH. *Protein Sci,* **18**(7), pp.1531-40.

MCMURRAY, C. T. 2010. Mechanisms of trinucleotide repeat instability during human development. *Nature reviews. Genetics,* **11**(11), pp.786-799.

MEINHARDT, J., C. SACHSE, P. HORTSCHANSKY, N. GRIGORIEFF and M. FANDRICH. 2009. Abeta(1-40) fibril polymorphism implies diverse interaction patterns in amyloid fibrils. *J Mol Biol,* **386**(3), pp.869-77.

MEISL, G., X. YANG, E. HELLSTRAND, B. FROHM, J. B. KIRKEGAARD, S. I. COHEN, C. M. DOBSON, S. LINSE and T. P. KNOWLES. 2014a. Differences in nucleation behavior underlie the contrasting aggregation kinetics of the Abeta40 and Abeta42 peptides. *Proc Natl Acad Sci U S A,* **111**(26), pp.9384-9.

MEISL, G., X. YANG, E. HELLSTRAND, B. FROHM, J. B. KIRKEGAARD, S. I. A. COHEN, C. M. DOBSON, S. LINSE and T. P. J. KNOWLES. 2014b. Differences in nucleation behavior underlie the contrasting aggregation kinetics of the Aβ40 and Aβ42 peptides. *Proceedings of the National Academy of Sciences,* **111**(26), pp.9384-9389.

MESLEH, M. F., J. M. HUNTER, A. A. SHVARTSBURG, G. C. SCHATZ and M. F. JARROLD. 1996. Structural Information from Ion Mobility Measurements: Effects of the Long-Range Potential. *The Journal of Physical Chemistry,* **100**(40), pp.16082-16086.

MIYATA, T., O. ODA, R. INAGI, Y. IIDA, N. ARAKI, N. YAMADA, S. HORIUCHI, N. TANIGUCHI, K. MAEDA and T. KINOSHITA. 1993. beta 2-Microglobulin

modified with advanced glycation end products is a major component of hemodialysis-associated amyloidosis. *J Clin Invest,* **92**(3), pp.1243-52.

MONDRAGON-RODRIGUEZ, S., G. PERRY, X. ZHU and J. BOEHM. 2012. Amyloid Beta and tau proteins as therapeutic targets for Alzheimer's disease treatment: rethinking the current strategy. *Int J Alzheimers Dis,* **2012**, p630182.

MOREL, B., L. VARELA, A. I. AZUAGA and F. CONEJERO-LARA. 2010. Environmental Conditions Affect the Kinetics of Nucleation of Amyloid Fibrils and Determine Their Morphology. *Biophys J,* **99**(11), pp.3801-3810.

MORENO-GONZALEZ, I. and C. SOTO. 2011. Misfolded Protein Aggregates: Mechanisms, Structures and Potential for Disease Transmission. *Seminars in cell & developmental biology,* **22**(5), pp.482-487.

MORTENSEN, D. N., A. C. SUSA and E. R. WILLIAMS. 2017. Collisional Cross-Sections with T-Wave Ion Mobility Spectrometry without Experimental Calibration. *J Am Soc Mass Spectrom,* **28**(7), pp.1282-1292.

MOULICK, R., R. R. GOLUGURI and J. B. UDGAONKAR. 2017. Role of Folding Intermediates in Initiating Aggregation of the Prion Protein. *The FASEB Journal,* **31**(1 Supplement), pp.763.6-763.6.

MUCHOWSKI, P. J., K. NING, C. D'SOUZA-SCHOREY and S. FIELDS. 2002. Requirement of an intact microtubule cytoskeleton for aggregation and inclusion body formation by a mutant huntingtin fragment. *Proceedings of the National Academy of Sciences,* **99**(2), pp.727-732.

MULLER-SPATH, S., A. SORANNO, V. HIRSCHFELD, H. HOFMANN, S. RUEGGER, L. REYMOND, D. NETTELS and B. SCHULER. 2010. From the Cover: Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc Natl Acad Sci U S A,* **107**(33), pp.14609-14.

MULLER, D. R., P. SCHINDLER, H. TOWBIN, U. WIRTH, H. VOSHOL, S. HOVING and M. O. STEINMETZ. 2001. Isotope-tagged cross-linking reagents. A new tool in mass spectrometric protein interaction analysis. *Anal Chem,* **73**(9), pp.1927-34.

MUNSON, M. S. B. and F. H. FIELD. 1966. Chemical Ionization Mass Spectrometry. I. General Introduction. *Journal of the American Chemical Society,* **88**(12), pp.2621-2630.

MURAKAMI, K., H. ELMLUND, N. KALISMAN, D. A. BUSHNELL, C. M. ADAMS, M. AZUBEL, D. ELMLUND, Y. LEVI-KALISMAN, X. LIU, B. J. GIBBONS, M. LEVITT and R. D. KORNBERG. 2013. Architecture of an RNA polymerase II transcription pre-initiation complex. *Science,* **342**(6159), p1238724.

MURPHY, M. P. and H. LEVINE, 3RD. 2010. Alzheimer's disease and the amyloid-beta peptide. *J Alzheimers Dis,* **19**(1), pp.311-23.

NAGAI, Y., N. FUJIKAKE, K. OHNO, H. HIGASHIYAMA, H. A. POPIEL, J. RAHADIAN, M. YAMAGUCHI, W. J. STRITTMATTER, J. R. BURKE and T. TODA. 2003. Prevention of polyglutamine oligomerization and neurodegeneration by the peptide inhibitor QBP1 in Drosophila. *Hum Mol Genet,* **12**(11), pp.1253-9.

NAGAI, Y., T. INUI, H. A. POPIEL, N. FUJIKAKE, K. HASEGAWA, Y. URADE, Y. GOTO, H. NAIKI and T. TODA. 2007. A toxic monomeric conformer of the polyglutamine protein. *Nat Struct Mol Biol,* **14**(4), pp.332-40.

NAGAI, Y., T. TUCKER, H. REN, D. J. KENAN, B. S. HENDERSON, J. D. KEENE, W. J. STRITTMATTER and J. R. BURKE. 2000. Inhibition of polyglutamine protein aggregation and cell death by novel peptides identified by phage display screening. *J Biol Chem,* **275**(14), pp.10437-42.

NAIKI, H., K. HIGUCHI, M. HOSOKAWA and T. TAKEDA. 1989. Fluorometric determination of amyloid fibrils in vitro using the fluorescent dye, thioflavin T1. *Anal Biochem,* **177**(2), pp.244-9.

NCBI. 2016. *ENTREZ Gene ATXN3*. [Accessed 14/08/17]. Available from: https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=ShowDetailView&TermToSearch=4287.

NEUDECKER, P., P. ROBUSTELLI, A. CAVALLI, P. WALSH, P. LUNDSTRÖM, A. ZARRINE-AFSAR, S. SHARPE, M. VENDRUSCOLO and L. E. KAY. 2012. Structure of an Intermediate State in Protein Folding and Aggregation. *Science,* **336**(6079), p362.

NEVES-CARVALHO, A., E. LOGARINHO, A. FREITAS, S. DUARTE-SILVA, C. COSTA MDO, A. SILVA-FERNANDES, M. MARTINS, S. C. SERRA, A. T. LOPES, H. L. PAULSON, P. HEUTINK, J. B. RELVAS and P. MACIEL. 2015. Dominant negative effect of polyglutamine expansion perturbs normal function of ataxin-3 in neuronal cells. *Hum Mol Genet,* **24**(1), pp.100-17.

NGUYEN-HUYNH, N. T., G. SHAROV, C. POTEL, P. FICHTER, S. TROWITZSCH, I. BERGER, V. LAMOUR, P. SCHULTZ, N. POTIER and E. LEIZE-WAGNER. 2015. Chemical cross-linking and mass spectrometry to determine the subunit interaction network in a recombinant human SAGA HAT subcomplex. *Protein Sci,* **24**(8), pp.1232-46.

NICASTRO, G., L. MASINO, V. ESPOSITO, R. P. MENON, A. DE SIMONE, F. FRATERNALI and A. PASTORE. 2009. Josephin domain of ataxin-3 contains two distinct ubiquitin-binding sites. *Biopolymers,* **91**(12), pp.1203-14.

NICASTRO, G., R. P. MENON, L. MASINO, P. P. KNOWLES, N. Q. MCDONALD and A. PASTORE. 2005. The solution structure of the Josephin domain of ataxin-3: structural determinants for molecular recognition. *Proc Natl Acad Sci U S A,* **102**(30), pp.10493-8.

NIELSEN, J. T., M. BJERRING, M. D. JEPPESEN, R. O. PEDERSEN, J. M. PEDERSEN, K. L. HEIN, T. VOSEGAARD, T. SKRYDSTRUP, D. E. OTZEN and N. C. NIELSEN. 2009. Unique identification of supramolecular structures in amyloid fibrils by solid-state NMR spectroscopy. *Angew Chem Int Ed Engl,* **48**(12), pp.2118-21.

NIU, S., B. C. KIM, C. A. FIERKE and B. T. RUOTOLO. 2016. Ion mobility-mass spectrometry reveals evidence of specific complex formation between human histone deacetylase 8 and poly-r(C)-binding protein 1. *International Journal of Mass Spectrometry*.

NUGENT, E., C. F. KAMINSKI and G. S. KAMINSKI SCHIERLE. 2017. Super-resolution imaging of alpha-synuclein polymorphisms and their potential role in neurodegeneration. *Integrative Biology,* **9**(3), pp.206-210.

NUZZO, M. T., M. FIOCCHETTI, P. TOTTA, M. A. MELONE, A. CARDINALE, F. R. FUSCO, S. GUSTINCICH, F. PERSICHETTI, P. ASCENZI and M. MARINO. 2016. Huntingtin polyQ Mutation Impairs the 17beta-

Estradiol/Neuroglobin Pathway Devoted to Neuron Survival. *Mol Neurobiol*.

O'BRIEN, R. J. and P. C. WONG. 2011. Amyloid precursor protein processing and Alzheimer's disease. *Annu Rev Neurosci,* **34**, pp.185-204.

OKADA, T., K. IKEDA, M. WAKABAYASHI, M. OGAWA and K. MATSUZAKI. 2008. Formation of toxic Abeta(1-40) fibrils on GM1 ganglioside-containing membranes mimicking lipid rafts: polymorphisms in Abeta(1-40) fibrils. *J Mol Biol,* **382**(4), pp.1066-74.

OKAMOTO, Y., Y. NAGAI, N. FUJIKAKE, H. AKIKO POPIEL, T. YOSHIOKA, T. TODA and T. INUI. 2009. Surface plasmon resonance characterization of specific binding of polyglutamine aggregation inhibitors to the expanded polyglutamine stretch. *Biochem Biophys Res Commun,* **378**(3), pp.634-9.

OLZSCHA, H., S. M. SCHERMANN, A. C. WOERNER, S. PINKERT, M. H. HECHT, G. G. TARTAGLIA, M. VENDRUSCOLO, M. HAYER-HARTL, F. U. HARTL and R. M. VABULAS. 2011. Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell,* **144**(1), pp.67-78.

OTZEN, D. 2010. Functional amyloid: turning swords into plowshares. *Prion,* **4**(4), pp.256-64.

PALHANO, F. L., J. LEE, N. P. GRIMSTER and J. W. KELLY. 2013. Toward the molecular mechanism(s) by which EGCG treatment remodels mature amyloid fibrils. *J Am Chem Soc,* **135**(20), pp.7503-10.

PARAVASTU, A. K., R. D. LEAPMAN, W. M. YAU and R. TYCKO. 2008. Molecular structural basis for polymorphism in Alzheimer's beta-amyloid fibrils. *Proc Natl Acad Sci U S A,* **105**(47), pp.18349-54.

PAUL, W. and H. STEINWEDEL. 1953. *Notizen: Ein neues Massenspektrometer ohne Magnetfeld*. *Zeitschrift für Naturforschung A.*

PAULING, L. and R. B. COREY. 1951. Configurations of Polypeptide Chains With Favored Orientations Around Single Bonds: Two New Pleated Sheets. *Proc Natl Acad Sci U S A,* **37**(11), pp.729-740.

PAULSON, H. L., S. S. DAS, P. B. CRINO, M. K. PEREZ, S. C. PATEL, D. GOTSDINER, K. H. FISCHBECK and R. N. PITTMAN. 1997. Machado-Joseph disease gene product is a cytoplasmic protein widely expressed in brain. *Ann Neurol,* **41**(4), pp.453-62.

PAUWELS, K., T. L. WILLIAMS, K. L. MORRIS, W. JONCKHEERE, A. VANDERSTEEN, G. KELLY, J. SCHYMKOWITZ, F. ROUSSEAU, A. PASTORE, L. C. SERPELL and K. BROERSEN. 2012. Structural basis for increased toxicity of pathological abeta42:abeta40 ratios in Alzheimer disease. *J Biol Chem,* **287**(8), pp.5650-60.

PEARSON, K. M., L. K. PANNELL and H. M. FALES. 2002. Intramolecular cross-linking experiments on cytochrome c and ribonuclease A using an isotope multiplet method. *Rapid Commun Mass Spectrom,* **16**(3), pp.149-59.

PELLARIN, R., P. SCHUETZ, E. GUARNERA and A. CAFLISCH. 2010. Amyloid Fibril Polymorphism Is under Kinetic Control. *Journal of the American Chemical Society,* **132**(42), pp.14960-14970.

PERTINHEZ, T. A., M. BOUCHARD, E. J. TOMLINSON, R. WAIN, S. J. FERGUSON, C. M. DOBSON and L. J. SMITH. 2001. Amyloid fibril formation by a helical cytochrome. *FEBS Lett,* **495**(3), pp.184-186.

PETKOVA, A. T., R. D. LEAPMAN, Z. GUO, W. M. YAU, M. P. MATTSON and R. TYCKO. 2005. Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils. *Science,* **307**(5707), pp.262-5.

PETKOVA, A. T., W. M. YAU and R. TYCKO. 2006. Experimental constraints on quaternary structure in Alzheimer's beta-amyloid fibrils. *Biochemistry,* **45**(2), pp.498-512.

PFEIFFER, A., M. S. LUIJSTERBURG, K. ACS, W. W. WIEGANT, A. HELFRICHT, L. K. HERZOG, M. MINOIA, C. BOTTCHER, F. A. SALOMONS, H. VAN ATTIKUM and N. P. DANTUMA. 2017. Ataxin-3 consolidates the MDC1-dependent DNA double-strand break response by counteracting the SUMO-targeted ubiquitin ligase RNF4. *EMBO J,* **36**(8), pp.1066-1083.

PINOTSI, D., C. H. MICHEL, A. K. BUELL, R. F. LAINE, P. MAHOU, C. M. DOBSON, C. F. KAMINSKI and G. S. KAMINSKI SCHIERLE. 2016. Nanoscopic insights into seeding mechanisms and toxicity of α-synuclein species in neurons. *Proceedings of the National Academy of Sciences,* **113**(14), pp.3815-3819.

POLITIS, A., C. SCHMIDT, E. TJIOE, A. M. SANDERCOCK, K. LASKER, Y. GORDIYENKO, D. RUSSEL, A. SALI and C. V. ROBINSON. 2015. Topological models of heteromeric protein assemblies from mass spectrometry: application to the yeast eIF3:eIF5 complex. *Chem Biol,* **22**(1), pp.117-28.

POPIEL, H. A., Y. NAGAI, N. FUJIKAKE and T. TODA. 2007. Protein transduction domain-mediated delivery of QBP1 suppresses polyglutamine-induced neurodegeneration in vivo. *Mol Ther,* **15**(2), pp.303-9.

POPIEL, H. A., Y. NAGAI, N. FUJIKAKE and T. TODA. 2009. Delivery of the aggregate inhibitor peptide QBP1 into the mouse brain using PTDs and its therapeutic effect on polyglutamine disease mice. *Neurosci Lett,* **449**(2), pp.87-92.

POPIEL, H. A., T. TAKEUCHI, J. R. BURKE, W. J. STRITTMATTER, T. TODA, K. WADA and Y. NAGAI. 2013. Inhibition of protein misfolding/aggregation using polyglutamine binding peptide QBP1 as a therapy for the polyglutamine diseases. *Neurotherapeutics,* **10**(3), pp.440-6.

POPIEL, H. A., T. TAKEUCHI, H. FUJITA, K. YAMAMOTO, C. ITO, H. YAMANE, S. MURAMATSU, T. TODA, K. WADA and Y. NAGAI. 2012. Hsp40 gene therapy exerts therapeutic effects on polyglutamine disease mice via a non-cell autonomous mechanism. *PLoS One,* **7**(11), pe51069.

PRESTON, G. W., S. E. RADFORD, A. E. ASHCROFT and A. J. WILSON. 2012. Covalent cross-linking within supramolecular peptide structures. *Anal Chem,* **84**(15), pp.6790-7.

PRILLER, C., T. BAUER, G. MITTEREGGER, B. KREBS, H. A. KRETZSCHMAR and J. HERMS. 2006. Synapse formation and function is modulated by the amyloid precursor protein. *J Neurosci,* **26**(27), pp.7212-21.

PRINGLE, S. D., K. GILES, J. L. WILDGOOSE, J. P. WILLIAMS, S. E. SLADE, K. THALASSINOS, R. H. BATEMAN, M. T. BOWERS and J. H. SCRIVENS. 2007. An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *International Journal of Mass Spectrometry,* **261**(1), pp.1-12.

The Proteomics Protocols Handbook. 2005. 1 ed.

PROVENCHER, S. W. and J. GLOCKNER. 1981. Estimation of globular protein secondary structure from circular dichroism. *Biochemistry,* **20**(1), pp.33-7.

PRUSINER, S. B. 1991. Molecular biology of prion diseases. *Science,* **252**(5012), pp.1515-22.

QI-TAKAHARA, Y., M. MORISHIMA-KAWASHIMA, Y. TANIMURA, G. DOLIOS, N. HIROTANI, Y. HORIKOSHI, F. KAMETANI, M. MAEDA, T. C. SAIDO, R. WANG and Y. IHARA. 2005. Longer forms of amyloid beta protein: implications for the mechanism of intramembrane cleavage by gamma-secretase. *J Neurosci,* **25**(2), pp.436-45.

QIANG, W., W. M. YAU, J. X. LU, J. COLLINGE and R. TYCKO. 2017. Structural variation in amyloid-beta fibrils from Alzheimer's disease clinical subtypes. *Nature,* **541**(7636), pp.217-221.

QIANG, W., W. M. YAU, Y. LUO, M. P. MATTSON and R. TYCKO. 2012. Antiparallel beta-sheet architecture in Iowa-mutant beta-amyloid fibrils. *Proc Natl Acad Sci U S A,* **109**(12), pp.4443-8.

QUIST, A., I. DOUDEVSKI, H. LIN, R. AZIMOVA, D. NG, B. FRANGIONE, B. KAGAN, J. GHISO and R. LAL. 2005. Amyloid ion channels: a common structural link for protein-misfolding disease. *Proc Natl Acad Sci U S A,* **102**(30), pp.10427-32.

RAMOS-MARTIN, F., R. HERVAS, M. CARRION-VAZQUEZ and D. V. LAURENTS. 2014. NMR spectroscopy reveals a preferred conformation with a defined hydrophobic cluster for polyglutamine binding peptide 1. *Arch Biochem Biophys,* **558**, pp.104-10.

RAND, K. D., S. D. PRINGLE, J. P. MURPHY, 3RD, K. E. FADGEN, J. BROWN and J. R. ENGEN. 2009. Gas-phase hydrogen/deuterium exchange in a traveling wave ion guide for the examination of protein conformations. *Anal Chem,* **81**(24), pp.10019-28.

RANJAN, P. and A. KUMAR. 2017. Perturbation in Long-Range Contacts Modulates the Kinetics of Amyloid Formation in alpha-Synuclein Familial Mutants. *ACS Chem Neurosci*.

RASIA, R. M., C. W. BERTONCINI, D. MARSH, W. HOYER, D. CHERNY, M. ZWECKSTETTER, C. GRIESINGER, T. M. JOVIN and C. O. FERNANDEZ. 2005. Structural characterization of copper(II) binding to alpha-synuclein: Insights into the bioinorganic chemistry of Parkinson's disease. *Proc Natl Acad Sci U S A,* **102**(12), pp.4294-9.

RAYLEIGH, L. 1882. On the equilibrium of liquid conducting masses charged with electricity. *Philosophical Magazine,* **14**(87), pp.184-186.

REAUME, A. G., J. L. ELLIOTT, E. K. HOFFMAN, N. W. KOWALL, R. J. FERRANTE, D. F. SIWEK, H. M. WILCOX, D. G. FLOOD, M. F. BEAL, R. H. BROWN, JR., R. W. SCOTT and W. D. SNIDER. 1996. Motor neurons in Cu/Zn superoxide dismutase-deficient mice develop normally but exhibit enhanced cell death after axonal injury. *Nat Genet,* **13**(1), pp.43-7.

REN, H., Y. NAGAI, T. TUCKER, W. J. STRITTMATTER and J. R. BURKE. 2001. Amino acid sequence requirements of peptides that inhibit polyglutamine-protein aggregation and cell death. *Biochem Biophys Res Commun,* **288**(3), pp.703-10.

RICHARD, T., P. POUPARD, M. NASSRA, Y. PAPASTAMOULIS, M. L. IGLESIAS, S. KRISA, P. WAFFO-TEGUO, J. M. MERILLON and J. P. MONTI. 2011. Protective effect of epsilon-viniferin on beta-amyloid peptide aggregation investigated by electrospray ionization mass spectrometry. *Bioorg Med Chem,* **19**(10), pp.3152-5.

ROCKABRAND, E., N. SLEPKO, A. PANTALONE, V. N. NUKALA, A. KAZANTSEV, J. L. MARSH, P. G. SULLIVAN, J. S. STEFFAN, S. L. SENSI and L. M. THOMPSON. 2007. The first 17 amino acids of Huntingtin modulate its sub-cellular localization, aggregation and effects on calcium homeostasis. *Hum Mol Genet,* **16**(1), pp.61-77.

ROEPSTORFF, P. and J. FOHLMAN. 1984. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom,* **11**(11), p601.

ROHER, A. E., J. D. LOWENSON, S. CLARKE, C. WOLKOW, R. WANG, R. J. COTTER, I. M. REARDON, H. A. ZURCHER-NEELY, R. L. HEINRIKSON, M. J. BALL and ET AL. 1993a. Structural alterations in the peptide backbone of beta-amyloid core protein may account for its deposition and stability in Alzheimer's disease. *J Biol Chem,* **268**(5), pp.3072-83.

ROHER, A. E., J. D. LOWENSON, S. CLARKE, A. S. WOODS, R. J. COTTER, E. GOWING and M. J. BALL. 1993b. beta-Amyloid-(1-42) is a major component of cerebrovascular amyloid deposits: implications for the pathology of Alzheimer disease. *Proc Natl Acad Sci U S A,* **90**(22), pp.10836-40.

ROSATI, S., Y. YANG, A. BARENDREGT and A. J. HECK. 2014. Detailed mass analysis of structural heterogeneity in monoclonal antibodies using native mass spectrometry. *Nat Protoc,* **9**(4), pp.967-76.

ROSENBERG, I. M. 1996. Tracking the Target Protein. *Protein analysis and purification : benchtop techniques.* 2 ed. Boston: Birkhäuser, p.49.

RUBIN, S. M., S. Y. LEE, E. J. RUIZ, A. PINES and D. E. WEMMER. 2002. Detection and characterization of xenon-binding sites in proteins by 129Xe NMR spectroscopy. *J Mol Biol,* **322**(2), pp.425-40.

RUOTOLO, B. T., J. L. BENESCH, A. M. SANDERCOCK, S. J. HYUNG and C. V. ROBINSON. 2008. Ion mobility-mass spectrometry analysis of large protein complexes. *Nat Protoc,* **3**(7), pp.1139-52.

RUOTOLO, B. T., K. GILES, I. CAMPUZANO, A. M. SANDERCOCK, R. H. BATEMAN and C. V. ROBINSON. 2005. Evidence for macromolecular protein rings in the absence of bulk water. *Science,* **310**(5754), pp.1658-61.

RUOTOLO, B. T., S. J. HYUNG, P. M. ROBINSON, K. GILES, R. H. BATEMAN and C. V. ROBINSON. 2007. Ion mobility-mass spectrometry reveals long-lived, unfolded intermediates in the dissociation of protein complexes. *Angew Chem Int Ed Engl,* **46**(42), pp.8001-4.

SACCHETTINI, J. C. and J. W. KELLY. 2002. Therapeutic strategies for human amyloid diseases. *Nat Rev Drug Discov,* **1**(4), pp.267-275.

SAIKI, M., S. HONDA, K. KAWASAKI, D. ZHOU, A. KAITO, T. KONAKAHARA and H. MORII. 2005. Higher-order molecular packing in amyloid-like fibrils constructed with linear arrangements of hydrophobic and hydrogen-bonding side-chains. *J Mol Biol,* **348**(4), pp.983-98.

SANTAMBROGIO, C., A. M. FRANA, A. NATALELLO, E. PAPALEO, M. E. REGONESI, S. M. DOGLIA, P. TORTORA, G. INVERNIZZI and R. GRANDORI. 2012. The role of the central flexible region on the aggregation and conformational properties of human ataxin-3. *FEBS J.,* **279**(3), pp.451-63.

SARELL, C. J., S. R. WILKINSON and J. H. VILES. 2010. Substoichiometric levels of Cu2+ ions accelerate the kinetics of fiber formation and promote cell toxicity of amyloid-{beta} from Alzheimer disease. *J Biol Chem,* **285**(53), pp.41533-40.

SAUNDERS, H. M., G. D, R. M, D. Y, R. AL and S. P. BOTTOMLEY. 2011. Flanking domain stability modulates the aggregation kinetics of a polyglutamine disease protein. *Protein Sci,* **20**(10), pp.1675-81.

SCARFF, C. A., B. ALMEIDA, J. FRAGA, S. MACEDO-RIBEIRO, S. E. RADFORD and A. E. ASHCROFT. 2015. Examination of Ataxin-3 (atx-3) Aggregation by Structural Mass Spectrometry Techniques: A Rationale for Expedited Aggregation upon Polyglutamine (polyQ) Expansion. *Mol Cell Proteomics,* **14**(5), pp.1241-53.

SCARFF, C. A., A. SICORELLO, R. J. TOME, S. MACEDO-RIBEIRO, A. E. ASHCROFT and S. E. RADFORD. 2013. A tale of a tail: Structural insights into the conformational properties of the polyglutamine protein ataxin-3. *Int J Mass Spectrom,* **345-347**, pp.63-70.

SCHARF, J. M., M. G. ENDRIZZI, A. WETTER, S. HUANG, T. G. THOMPSON, K. ZERRES, W. F. DIETRICH, B. WIRTH and L. M. KUNKEL. 1998. Identification of a candidate modifying gene for spinal muscular atrophy by comparative genomics. *Nat Genet,* **20**(1), pp.83-6.

SCHIFFRIN, B., A. N. CALABRESE, P. W. A. DEVINE, S. A. HARRIS, A. E. ASHCROFT, D. J. BROCKWELL and S. E. RADFORD. 2016. Skp is a multivalent chaperone of outer-membrane proteins. *Nat Struct Mol Biol,* **23**(9), pp.786-793.

SCHILLING, B., R. H. ROW, B. W. GIBSON, X. GUO and M. M. YOUNG. 2003. MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides. *J Am Soc Mass Spectrom,* **14**(8), pp.834-50.

SCHMIDT, C. and C. V. ROBINSON. 2014a. A comparative cross-linking strategy to probe conformational changes in protein complexes. *Nat. Protocols,* **9**(9), pp.2224-2236.

SCHMIDT, C. and C. V. ROBINSON. 2014b. A comparative cross-linking strategy to probe conformational changes in protein complexes. *Nat Protoc,* **9**(9), pp.2224-36.

SCHONBRUN, J. and K. A. DILL. 2003. Fast protein folding kinetics. *Proceedings of the National Academy of Sciences,* **100**(22), pp.12678-12682.

SCHWAMBORN, K., M. KRIEGSMANN and W. WEICHERT. 2017. MALDI imaging mass spectrometry - From bench to bedside. *Biochim Biophys Acta,* **1865**(7), pp.776-783.

SCIENTIFIC, T. 2012. *Instructions - DSS and BS3 Crosslinkers*. Online: Thermo Scientific.

SCLAVI, B., S. WOODSON, M. SULLIVAN, M. CHANCE and M. BRENOWITZ. 1998. Following the folding of RNA with time-resolved synchrotron X-ray footprinting. *Methods Enzymol,* **295**, pp.379-402.

SCOPES, R. K. 1974. Measurement of protein by spectrophotometry at 205 nm. *Anal Biochem,* **59**(1), pp.277-82.

SELIVANOVA, O. M., A. V. GLYAKINA, E. Y. GORBUNOVA, L. G. MUSTAEVA, M. Y. SUVORINA, E. I. GRIGORASHVILI, A. D. NIKULIN, N. V. DOVIDCHENKO, V. V. REKSTINA, T. S. KALEBINA, A. K. SURIN and O. V. GALZITSKAYA. 2016. Structural model of amyloid fibrils for amyloidogenic peptide from Bgl2p-glucantransferase of S. cerevisiae cell wall and its modifying analog. New morphology of amyloid fibrils. *Biochim Biophys Acta,* **1864**(11), pp.1489-99.

SELKOE, D. J. and J. HARDY. 2016. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol Med,* **8**(6), pp.595-608.

SENGUPTA, U., A. N. NILSON and R. KAYED. 2016. The Role of Amyloid-β Oligomers in Toxicity, Propagation, and Immunotherapy. *EBioMedicine,* **6**, pp.42-9.

SEO, J., W. HOFFMANN, S. WARNKE, M. T. BOWERS, K. PAGEL and G. VON HELDEN. 2016. Retention of Native Protein Structures in the Absence of Solvent: A Coupled Ion Mobility and Spectroscopic Study. *Angew Chem Int Ed Engl,* **55**(45), pp.14173-14176.

SEO, J., W. HOFFMANN, S. WARNKE, X. HUANG, S. GEWINNER, W. SCHOLLKOPF, M. T. BOWERS, G. VON HELDEN and K. PAGEL. 2017. An infrared spectroscopy approach to follow beta-sheet formation in peptide amyloid assemblies. *Nat Chem,* **9**(1), pp.39-44.

SERVAGE, K. A., J. A. SILVEIRA, K. L. FORT, D. E. CLEMMER and D. H. RUSSELL. 2015. Water-Mediated Dimerization of Ubiquitin Ions Captured by Cryogenic Ion Mobility-Mass Spectrometry. *The Journal of Physical Chemistry Letters,* **6**(24), pp.4947-4951.

SEURING, C., J. VERASDONCK, P. RINGLER, R. CADALBERT, H. STAHLBERG, A. BÖCKMANN, B. H. MEIER and R. RIEK. 2017. Amyloid Fibril Polymorphism: Almost Identical on the Atomic Level, Mesoscopically Very Different. *The Journal of Physical Chemistry B,* **121**(8), pp.1783-1792.

SHARP, J. S., J. M. BECKER and R. L. HETTICH. 2004. Analysis of protein solvent accessible surfaces by photochemical oxidation and mass spectrometry. *Anal Chem,* **76**(3), pp.672-83.

SHEA, Y. F., L. W. CHU, A. O. CHAN, J. HA, Y. LI and Y. Q. SONG. 2016. A systematic review of familial Alzheimer's disease: Differences in presentation of clinical features among three mutated genes and potential ethnic differences. *J Formos Med Assoc,* **115**(2), pp.67-75.

SHEKHTMAN, A. and D. COWBURN. 2002. A ubiquitin-interacting motif from Hrs binds to and occludes the ubiquitin surface necessary for polyubiquitination in monoubiquitinated proteins. *Biochem Biophys Res Commun,* **296**(5), pp.1222-7.

SHELIMOV, K. B. and M. F. JARROLD. 1995. Ion Mobility Studies of PdCn+ Clusters: Where Are the Fullerenes? *The Journal of Physical Chemistry,* **99**(50), pp.17677-17679.

SHEPHERD, D. A., K. HOLMES, D. J. ROWLANDS, N. J. STONEHOUSE and A. E. ASHCROFT. 2013. Using ion mobility spectrometry-mass spectrometry to

decipher the conformational and assembly characteristics of the hepatitis B capsid protein. *Biophys J,* **105**(5), pp.1258-67.

SHI, J.-M., J. PEI, E.-Q. LIU and L. ZHANG. 2017. Bis(sulfosuccinimidyl) suberate (BS3) crosslinking analysis of the behavior of amyloid-β peptide in solution and in phospholipid membranes. *PLoS One,* **12**(3), pe0173871.

SHIMMYO, Y., T. KIHARA, A. AKAIKE, T. NIIDOME and H. SUGIMOTO. 2008. Multifunction of myricetin on A beta: neuroprotection via a conformational change of A beta and reduction of A beta via the interference of secretases. *J Neurosci Res,* **86**(2), pp.368-77.

SHVARTSBURG, A. A. and M. F. JARROLD. 1996. An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chemical Physics Letters,* **261**(1), pp.86-91.

SHVARTSBURG, A. A. and R. D. SMITH. 2008. Fundamentals of traveling wave ion mobility spectrometry. *Anal Chem,* **80**(24), pp.9689-99.

SILVERS, R., M. T. COLVIN, K. K. FREDERICK, A. C. JACAVONE, S. L. LINDQUIST, S. LINSE and R. G. GRIFFIN. 2017. Aggregation and Fibril Structure of AbetaM01-42 and Abeta1-42. *Biochemistry*.

SIMOES, A. T., N. GONCALVES, A. KOEPPEN, N. DEGLON, S. KUGLER, C. B. DUARTE and L. PEREIRA DE ALMEIDA. 2012. Calpastatin-mediated inhibition of calpains in the mouse brain prevents mutant ataxin 3 proteolysis, nuclear localization and aggregation, relieving Machado-Joseph disease. *Brain,* **135**(Pt 8), pp.2428-39.

SIMOES, A. T., N. GONCALVES, R. J. NOBRE, C. B. DUARTE and L. PEREIRA DE ALMEIDA. 2014. Calpain inhibition reduces ataxin-3 cleavage alleviating neuropathology and motor impairments in mouse models of Machado-Joseph disease. *Hum Mol Genet,* **23**(18), pp.4932-44.

SINHA, S., D. H. LOPES, Z. DU, E. S. PANG, A. SHANMUGAM, A. LOMAKIN, P. TALBIERSKY, A. TENNSTAEDT, K. MCDANIEL, R. BAKSHI, P. Y. KUO, M. EHRMANN, G. B. BENEDEK, J. A. LOO, F. G. KLARNER, T. SCHRADER, C. WANG and G. BITAN. 2011. Lysine-specific molecular tweezers are broad-spectrum inhibitors of assembly and toxicity of amyloid proteins. *J Am Chem Soc,* **133**(42), pp.16958-69.

SIPE, J. D., M. D. BENSON, J. N. BUXBAUM, S. IKEDA, G. MERLINI, M. J. SARAIVA, P. WESTERMARK and A. NOMENCLATURE COMMITTEE OF THE INTERNATIONAL SOCIETY OF. 2012. Amyloid fibril protein nomenclature: 2012 recommendations from the Nomenclature Committee of the International Society of Amyloidosis. *Amyloid,* **19**(4), pp.167-70.

SIPE, J. D., M. D. BENSON, J. N. BUXBAUM, S. I. IKEDA, G. MERLINI, M. J. SARAIVA and P. WESTERMARK. 2016. Amyloid fibril proteins and amyloidosis: chemical identification and clinical classification International Society of Amyloidosis 2016 Nomenclature Guidelines. *Amyloid,* **23**(4), pp.209-213.

SITTLER, A., S. WALTER, N. WEDEMEYER, R. HASENBANK, E. SCHERZINGER, H. EICKHOFF, G. P. BATES, H. LEHRACH and E. E. WANKER. 1998. SH3GL3 associates with the Huntingtin exon 1 protein and promotes the formation of polygln-containing protein aggregates. *Mol Cell,* **2**(4), pp.427-36.

SIVANESAM, K., I. SHU, K. N. L. HUGGINS, M. TATAREK-NOSSOL, A. KAPURNIOTU and N. H. ANDERSEN. 2016. Peptide Inhibitors of the amyloidogenesis of IAPP: verification of the hairpin-binding geometry hypothesis. *FEBS Lett,* **590**(16), pp.2575-2583.

SMITH, D. P., T. W. KNAPMAN, I. CAMPUZANO, R. W. MALHAM, J. T. BERRYMAN, S. E. RADFORD and A. E. ASHCROFT. 2009. Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies. *Eur J Mass Spectrom (Chichester),* **15**(2), pp.113-30.

SMITH, D. P., S. E. RADFORD and A. E. ASHCROFT. 2010. Elongated oligomers in beta2-microglobulin amyloid assembly revealed by ion mobility spectrometry-mass spectrometry. *Proc Natl Acad Sci U S A,* **107**(15), pp.6794-8.

SMITH, D. P., L. A. WOODS, S. E. RADFORD and A. E. ASHCROFT. 2011. Structure and dynamics of oligomeric intermediates in beta2-microglobulin self-assembly. *Biophys J,* **101**(5), pp.1238-47.

SMITH, P. K., R. I. KROHN, G. T. HERMANSON, A. K. MALLIA, F. H. GARTNER, M. D. PROVENZANO, E. K. FUJIMOTO, N. M. GOEKE, B. J. OLSON and D. C. KLENK. 1985. Measurement of protein using bicinchoninic acid. *Anal Biochem,* **150**(1), pp.76-85.

SNEIDERIS, T., D. DARGUZIS, A. BOTYRIUTE, M. GRIGALIUNAS, R. WINTER and V. SMIRNOVAS. 2015. pH-Driven Polymorphism of Insulin Amyloid-Like Fibrils. *PLoS One,* **10**(8), pe0136602.

SNIJDER, J., R. J. ROSE, D. VEESLER, J. E. JOHNSON and A. J. R. HECK. 2013. Studying 18 Mega Dalton Virus Assemblies with Native Mass Spectrometry(). *Angew Chem Int Ed Engl,* **52**(14), pp.4020-4023.

SONG, A. X., C. J. ZHOU, Y. PENG, X. C. GAO, Z. R. ZHOU, Q. S. FU, J. HONG, D. H. LIN and H. Y. HU. 2010. Structural transformation of the tandem ubiquitin-interacting motifs in ataxin-3 and their cooperative interactions with ubiquitin chains. *PLoS One,* **5**(10), pe13202.

SOPER-HOPPER, M. T., J. D. ESCHWEILER and B. T. RUOTOLO. 2017. Ion Mobility-Mass Spectrometry Reveals a Dipeptide That Acts as a Molecular Chaperone for Amyloid β. *ACS Chemical Biology,* **12**(4), pp.1113-1120.

SREERAMA, N. and R. W. WOODY. 1993. A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal Biochem,* **209**(1), pp.32-44.

STATISTICS, O. F. N. 2015. *Deaths registered in England and Wales (Series DR): 2015.*

STEWART, K. L., E. HUGHES, E. A. YATES, G. R. AKIEN, T. Y. HUANG, M. A. LIMA, T. R. RUDD, M. GUERRINI, S. C. HUNG, S. E. RADFORD and D. A. MIDDLETON. 2016. Atomic Details of the Interactions of Glycosaminoglycans with Amyloid-beta Fibrils. *J Am Chem Soc,* **138**(27), pp.8328-31.

STEWART, K. L. and S. E. RADFORD. 2017. Amyloid plaques beyond Abeta: a survey of the diverse modulators of amyloid aggregation. *Biophys Rev*.

SULTANA, R., M. PERLUIGI and D. A. BUTTERFIELD. 2009. Oxidatively modified proteins in Alzheimer's disease (AD), mild cognitive impairment and animal models of AD: role of Abeta in pathogenesis. *Acta Neuropathol,* **118**(1), pp.131-50.

SYKA, J. E. P., J. J. COON, M. J. SCHROEDER, J. SHABANOWITZ and D. F. HUNT. 2004. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A,* **101**(26), pp.9528-9533.

TAIT, D., M. RICCIO, A. SITTLER, E. SCHERZINGER, S. SANTI, A. OGNIBENE, N. M. MARALDI, H. LEHRACH and E. E. WANKER. 1998. Ataxin-3 is transported into the nucleus and associates with the nuclear matrix. *Hum Mol Genet,* **7**(6), pp.991-7.

TAKAMOTO, K. and M. R. CHANCE. 2006. Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes. *Annu Rev Biophys Biomol Struct,* **35**, pp.251-76.

TAO, K., J. WANG, P. ZHOU, C. WANG, H. XU, X. ZHAO and J. R. LU. 2011. Self-assembly of short abeta(16-22) peptides: effect of terminal capping and the role of electrostatic interaction. *Langmuir,* **27**(6), pp.2723-30.

TAYLOR, R. C. and A. DILLIN. 2011. Aging as an Event of Proteostasis Collapse. *Cold Spring Harb Perspect Biol,* **3**(5).

TESTA, L., S. BROCCA and R. GRANDORI. 2011. Charge-surface correlation in electrospray ionization of folded and unfolded proteins. *Anal Chem,* **83**(17), pp.6459-63.

TESTA, L., S. BROCCA, C. SANTAMBROGIO, A. D'URZO, J. HABCHI, S. LONGHI, V. N. UVERSKY and R. GRANDORI. 2013. Extracting structural information from charge-state distributions of intrinsically disordered proteins by non-denaturing electrospray-ionization mass spectrometry. *Intrinsically Disord Proteins,* **1**(1), pe25068.

THAKUR, A. K., M. JAYARAMAN, R. MISHRA, M. THAKUR, V. M. CHELLGREN, I. J. BYEON, D. H. ANJUM, R. KODALI, T. P. CREAMER, J. F. CONWAY, A. M. GRONENBORN and R. WETZEL. 2009. Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism. *Nat Struct Mol Biol,* **16**(4), pp.380-9.

THOMSON, J. J. 1897. Cathode Rays. *Philosophical Magazine,* **44**(293).

THUAL, C., B. L, K. AA., W. S, B. J, C. C. and R. MELKI. 2001. Stability, folding, dimerization, and assembly properties of the yeast prion Ure2p. *Biochemistry,* **40**(6), pp.1764-73.

TIPPING, K. W., T. K. KARAMANOS, T. JAKHRIA, M. G. IADANZA, S. C. GOODCHILD, R. TUMA, N. A. RANSON, E. W. HEWITT and S. E. RADFORD. 2015. pH-induced molecular shedding drives the formation of amyloid fibril-derived oligomers. *Proc Natl Acad Sci U S A,* **112**(18), pp.5691-6.

TJERNBERG, L. O., D. J. CALLAWAY, A. TJERNBERG, S. HAHNE, C. LILLIEHOOK, L. TERENIUS, J. THYBERG and C. NORDSTEDT. 1999. A molecular model of Alzheimer amyloid beta-peptide fibril formation. *J Biol Chem,* **274**(18), pp.12619-25.

TOLMACHEV, A. V., M. E. CLOWERS BH FAU - BELOV, R. D. BELOV ME FAU - SMITH and R. D. SMITH. 2009. Coulombic effects in ion mobility spectrometry. *Anal Chem,* **81**(12), pp.4778-87.

TOMIC, J. L., A. PENSALFINI, E. HEAD and C. G. GLABE. 2009. Soluble fibrillar oligomer levels are elevated in Alzheimer's disease brain and correlate with cognitive dysfunction. *Neurobiol Dis,* **35**(3), pp.352-8.

TOMITA, K., H. A. POPIEL, Y. NAGAI, T. TODA, Y. YOSHIMITSU, H. OHNO, S. OISHI and N. FUJII. 2009. Structure-activity relationship study on polyglutamine binding peptide QBP1. *Bioorg Med Chem,* **17**(3), pp.1259-63.

TSOLIS, A. C., N. C. PAPANDREOU, V. A. ICONOMIDOU and S. J. HAMODRAKAS. 2013. A Consensus Method for the Prediction of 'Aggregation-Prone' Peptides in Globular Proteins. *PLoS One,* **8**(1), pe54175.

TYCKO, R. 2014. Physical and structural basis for polymorphism in amyloid fibrils. *Protein Science,* **23**(11), pp.1528-1539.

URBANC, B., L. CRUZ, S. YUN, S. V. BULDYREV, G. BITAN, D. B. TEPLOW and H. E. STANLEY. 2004. In silico study of amyloid beta-protein folding and oligomerization. *Proc Natl Acad Sci U S A,* **101**(50), pp.17345-50.

V. BARWICK, J. L., T. MALLET, B. STEIN, K. WEBB. 2006. *Best practice guide for generating mass spectra* [online]. Online: LGC ltd. Available from: http://www.rsc.org/images/MS2new_tcm18-102519.pdf.

VALLE-DELGADO, J. J., M. ALFONSO-PRIETO, N. S. DE GROOT, S. VENTURA, J. SAMITIER, C. ROVIRA and X. FERNANDEZ-BUSQUETS. 2010. Modulation of Abeta42 fibrillogenesis by glycosaminoglycan structure. *FASEB J,* **24**(11), pp.4250-61.

VALLEIX, S., J. D. GILLMORE, F. BRIDOUX, P. P. MANGIONE, A. DOGAN, B. NEDELEC, M. BOIMARD, G. TOUCHARD, J. M. GOUJON, C. LACOMBE, P. LOZERON, D. ADAMS, C. LACROIX, T. MAISONOBE, V. PLANTE-BORDENEUVE, J. A. VRANA, J. D. THEIS, S. GIORGETTI, R. PORCARI, S. RICAGNO, M. BOLOGNESI, M. STOPPINI, M. DELPECH, M. B. PEPYS, P. N. HAWKINS and V. BELLOTTI. 2012. Hereditary systemic amyloidosis due to Asp76Asn variant beta2-microglobulin. *N Engl J Med,* **366**(24), pp.2276-83.

VAN BROECKHOVEN, C. L. 1995. Molecular genetics of Alzheimer disease: identification of genes and gene mutations. *Eur Neurol,* **35**(1), pp.8-19.

VAN DEN HEUVEL, C., E. THORNTON and R. VINK. 2007. Traumatic brain injury and Alzheimer's disease: a review. *Prog Brain Res,* **161**, pp.303-16.

VAN HAM, T. J., M. A. HOLMBERG, A. T. VAN DER GOOT, E. TEULING, M. GARCIA-ARENCIBIA, H. E. KIM, D. DU, K. L. THIJSSEN, M. WIERSMA, R. BURGGRAAFF, P. VAN BERGEIJK, J. VAN RHEENEN, G. JERRE VAN VELUW, R. M. HOFSTRA, D. C. RUBINSZTEIN and E. A. NOLLEN. 2010. Identification of MOAG-4/SERF as a regulator of age-related proteotoxicity. *Cell,* **142**(4), pp.601-12.

VASSAR, P. S. and C. F. CULLING. 1959. Fluorescent stains, with special reference to amyloid and connective tissues. *Arch Pathol,* **68**, pp.487-98.

VERMA, M., A. VATS and V. TANEJA. 2015. Toxic species in amyloid disorders: Oligomers or mature fibrils. *Ann Indian Acad Neurol,* **18**(2), pp.138-45.

VIVEKANANDAN, S., J. R. BRENDER, S. Y. LEE and A. RAMAMOORTHY. 2011. A Partially Folded Structure of Amyloid-Beta(1-40) in an Aqueous Environment. *Biochem Biophys Res Commun,* **411**(2), pp.312-316.

VUGMEYSTER, L., M. A. CLARK, I. B. FALCONER, D. OSTROVSKY, D. GANTZ, W. QIANG and G. L. HOATSON. 2016. Flexibility and Solvation of Amyloid-beta Hydrophobic Core. *J Biol Chem,* **291**(35), pp.18484-95.

WALSH, D. M., D. M. HARTLEY, Y. KUSUMOTO, Y. FEZOUI, M. M. CONDRON, A. LOMAKIN, G. B. BENEDEK, D. J. SELKOE and D. B. TEPLOW. 1999. Amyloid beta-protein fibrillogenesis. Structure and biological activity of protofibrillar intermediates. *J Biol Chem,* **274**(36), pp.25945-52.

WALSH, D. M., E. THULIN, A. M. MINOGUE, N. GUSTAVSSON, E. PANG, D. B. TEPLOW and S. LINSE. 2009. A facile method for expression and purification of the Alzheimer's disease-associated amyloid beta-peptide. *FEBS J,* **276**(5), pp.1266-81.

WALTI, M. A., F. RAVOTTI, H. ARAI, C. G. GLABE, J. S. WALL, A. BOCKMANN, P. GUNTERT, B. H. MEIER and R. RIEK. 2016. Atomic-resolution structure of a disease-relevant Abeta(1-42) amyloid fibril. *Proc Natl Acad Sci U S A,* **113**(34), pp.E4976-84.

WANG, L. and M. R. CHANCE. 2011. Structural Mass Spectrometry of Proteins Using Hydroxyl Radical Based Protein Footprinting. *Analytical Chemistry,* **83**(19), pp.7234-7241.

WANG, L., A. I. ILITCHEV, M. J. GIAMMONA, F. LI, S. K. BURATTO and M. T. BOWERS. 2016a. Human Islet Amyloid Polypeptide Assembly: The Key Role of the 8-20 Fragment. *J Phys Chem B,* **120**(46), pp.11905-11911.

WANG, Q., L. LI and Y. YE. 2006. Regulation of retrotranslocation by p97-associated deubiquitinating enzyme ataxin-3. *J Cell Biol,* **174**(7), pp.963-71.

WANG, Y. and E. MANDELKOW. 2016. Tau in physiology and pathology. *Nat Rev Neurosci,* **17**(1), pp.22-35.

WANG, Y., X. YANG, W. MA, J. LI, Q. ZHANG, S. XIA, H. WANG, C. ZHANG, X. XU and J. ZHENG. 2016b. Clinical features and genetic diagnosis of hereditary spinocerebellar ataxia 3. *Mol Med Rep,* **14**(4), pp.3731-4.

WEGGEN, S. and D. BEHER. 2012. Molecular consequences of amyloid precursor protein and presenilin mutations causing autosomal-dominant Alzheimer's disease. *Alzheimers Res Ther,* **4**(2), p9.

WELANDER, H., J. FRANBERG, C. GRAFF, E. SUNDSTROM, B. WINBLAD and L. O. TJERNBERG. 2009. Abeta43 is more frequent than Abeta40 in amyloid plaque cores from Alzheimer disease brains. *J Neurochem,* **110**(2), pp.697-706.

WELLINGTON, C. L., L. M. ELLERBY, A. S. HACKAM, R. L. MARGOLIS, M. A. TRIFIRO, R. SINGARAJA, K. MCCUTCHEON, G. S. SALVESEN, S. S. PROPP, M. BROMM, K. J. ROWLAND, T. ZHANG, D. RASPER, S. ROY, N. THORNBERRY, L. PINSKY, A. KAKIZUKA, C. A. ROSS, D. W. NICHOLSON, D. E. BREDESEN and M. R. HAYDEN. 1998. Caspase cleavage of gene products associated with triplet expansion disorders generates truncated fragments containing the polyglutamine tract. *J Biol Chem,* **273**(15), pp.9158-67.

WESTERMARK, G. T., K. H. JOHNSON and P. WESTERMARK. 1999. Staining methods for identification of amyloid in tissue. *Methods Enzymol,* **309**, pp.3-25.

WESTERMARK, P., A. ANDERSSON and G. T. WESTERMARK. 2011. Islet amyloid polypeptide, islet amyloid, and diabetes mellitus. *Physiol Rev,* **91**(3), pp.795-826.

WHITMORE, L. and B. A. WALLACE. 2004. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res,* **32**(Web Server issue), pp.W668-73.

WICKY, B. I. M., S. L. SHAMMAS and J. CLARKE. 2017. Affinity of IDPs to their targets is modulated by ion-specific changes in kinetics and residual structure. *Proc Natl Acad Sci U S A*.

WILEY, W. C. and I. H. MCLAREN. 1955. Time-of-Flight Mass Spectrometer with Improved Resolution. *Review of Scientific Instruments,* **26**(12), pp.1150-1157.

WILM, M. and M. MANN. 1996. Analytical Properties of the Nanoelectrospray Ion Source. *Analytical Chemistry,* **68**(1), pp.1-8.

WILM, M. S. and M. MANN. 1994. Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *International Journal of Mass Spectrometry and Ion Processes,* **136**(2), pp.167-180.

WINBORN, B. J., S. M. TRAVIS, S. V. TODI, K. M. SCAGLIONE, P. XU, A. J. WILLIAMS, R. E. COHEN, J. PENG and H. L. PAULSON. 2008. The deubiquitinating enzyme ataxin-3, a polyglutamine disease protein, edits Lys63 linkages in mixed linkage ubiquitin chains. *J Biol Chem,* **283**(39), pp.26436-43.

WINKLHOFER, K. F., J. TATZELT and C. HAASS. 2008. The two faces of protein misfolding: gain- and loss-of-function in neurodegenerative diseases. *EMBO J,* **27**(2), pp.336-349.

WRIGHT, M. A., F. A. APRILE, P. AROSIO, M. VENDRUSCOLO, C. M. DOBSON and T. P. KNOWLES. 2015. Biophysical approaches for the study of interactions between molecular chaperones and protein aggregates. *Chem Commun (Camb),* **51**(77), pp.14425-34.

XIAO, Y., B. MA, D. MCELHENY, S. PARTHASARATHY, F. LONG, M. HOSHI, R. NUSSINOV and Y. ISHII. 2015. Abeta(1-42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. *Nat Struct Mol Biol,* **22**(6), pp.499-505.

XU, G. and M. R. CHANCE. 2007. Hydroxyl radical-mediated modification of proteins as probes for structural proteomics. *Chem Rev,* **107**(8), pp.3514-43.

YAMASHITA, M. and J. B. FENN. 1984. Electrospray ion source. Another variation on the free-jet theme. *The Journal of Physical Chemistry,* **88**(20), pp.4451-4459.

YAN, S. D., X. CHEN, J. FU, M. CHEN, H. ZHU, A. ROHER, T. SLATTERY, L. ZHAO, M. NAGASHIMA, J. MORSER, A. MIGHELI, P. NAWROTH, D. STERN and A. M. SCHMIDT. 1996. RAGE and amyloid-beta peptide neurotoxicity in Alzheimer's disease. *Nature,* **382**(6593), pp.685-91.

YAN, S. D., H. ZHU, A. ZHU, A. GOLABEK, H. DU, A. ROHER, J. YU, C. SOTO, A. M. SCHMIDT, D. STERN and M. KINDY. 2000. Receptor-dependent cell stress and amyloid accumulation in systemic amyloidosis. *Nat Med,* **6**(6), pp.643-51.

YANG, F., G. P. LIM, A. N. BEGUM, O. J. UBEDA, M. R. SIMMONS, S. S. AMBEGAOKAR, P. P. CHEN, R. KAYED, C. G. GLABE, S. A. FRAUTSCHY and G. M. COLE. 2005. Curcumin inhibits formation of amyloid beta oligomers

and fibrils, binds plaques, and reduces amyloid in vivo. *J Biol Chem,* **280**(7), pp.5892-901.

YEN, H. Y., J. T. S. HOPPER, I. LIKO, T. M. ALLISON, Y. ZHU, D. WANG, M. STEGMANN, S. MOHAMMED, B. WU and C. V. ROBINSON. 2017. Ligand binding to a G protein-coupled receptor captured in a mass spectrometer. *Sci Adv,* **3**(6), pe1701016.

YIANNOPOULOU, K. G. and S. G. PAPAGEORGIOU. 2013. Current and future treatments for Alzheimer's disease. *Therapeutic Advances in Neurological Disorders,* **6**(1), pp.19-33.

YOSHIMURA, Y., M. A. HOLMBERG, P. KUKIC, C. B. ANDERSEN, A. MATA-CABANA, S. F. FALSONE, M. VENDRUSCOLO, E. A. A. NOLLEN and F. A. A. MULDER. 2017. MOAG-4 promotes the aggregation of alpha-synuclein by competing with self-protective electrostatic interactions. *J Biol Chem,* **292**(20), pp.8269-8278.

YOUNG, L. M., P. CAO, D. P. RALEIGH, A. E. ASHCROFT and S. E. RADFORD. 2014. Ion mobility spectrometry-mass spectrometry defines the oligomeric intermediates in amylin amyloid formation and the mode of action of inhibitors. *J Am Chem Soc,* **136**(2), pp.660-70.

YOUNG, L. M., J. C. SAUNDERS, R. A. MAHOOD, C. H. REVILL, R. J. FOSTER, A. E. ASHCROFT and S. E. RADFORD. 2016. ESI-IMS-MS: A method for rapid analysis of protein aggregation and its inhibition by small molecules. *Methods,* **95**, pp.62-9.

YOUNG, L. M., J. C. SAUNDERS, R. A. MAHOOD, C. H. REVILL, R. J. FOSTER, L. H. TU, D. P. RALEIGH, S. E. RADFORD and A. E. ASHCROFT. 2015. Screening and classifying small-molecule inhibitors of amyloid formation using ion mobility spectrometry-mass spectrometry. *Nat Chem,* **7**(1), pp.73-81.

ZHANG, M., B. REN, Y. LIU, G. LIANG, Y. SUN, L. XU and J. ZHENG. 2017. Membrane Interactions of hIAPP Monomer and Oligomer with Lipid Membranes by Molecular Dynamics Simulations. *ACS Chem Neurosci*.

ZHANG, N., Y. GORDIYENKO, N. JOLY, E. LAWTON, C. V. ROBINSON and M. BUCK. 2014. Subunit dynamics and nucleotide-dependent asymmetry of an AAA(+) transcription complex. *J Mol Biol,* **426**(1), pp.71-83.

ZHAO, H., E. K. TUOMINEN and P. K. KINNUNEN. 2004. Formation of amyloid fibers triggered by phosphatidylserine-containing membranes. *Biochemistry,* **43**(32), pp.10302-7.

ZHEMKOV, V. A., A. A. KULMINSKAYA, I. B. BEZPROZVANNY and M. KIM. 2016. The 2.2-Angstrom resolution crystal structure of the carboxy-terminal region of ataxin-3. *FEBS Open Bio,* **6**(3), pp.168-78.

ZIDOVETZKI, R. 2009. *Wheel* [online]. [Accessed 14.09.17]. Available from: http://rzlab.ucr.edu/scripts/wheel/wheel.cgi?sequence=ABCDEFGHIJLKMNOP&submit=Submit.

ZUBAREV, R. A., N. L. KELLEHER and F. W. MCLAFFERTY. 1998. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *Journal of the American Chemical Society,* **120**(13), pp.3265-3266.

ZWANZIG, R., A. SZABO and B. BAGCHI. 1992. Levinthal's paradox. *Proc Natl Acad Sci U S A,* **89**(1), pp.20-2.