INVESTIGATION OF THE MULTI-DOMAIN AGGREGATION MECHANISM OF ATAXIN-3

A thesis submitted for the degree of Doctor of Philosophy

Helen Michelle Saunders

B.Sc.(Hons), B.Com., Dip.Lang(Chi)

Department of Biochemistry and Molecular Biology

Monash University

May 2010
Notice 1
Under the Copyright Act 1968, this thesis must be used only under the normal conditions of scholarly fair dealing. In particular no results or conclusions should be extracted from it, nor should it be copied or closely paraphrased in whole or in part without the written consent of the author. Proper written acknowledgement should be made for any assistance obtained from this thesis.

Notice 2
I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.
ERRATA

p 2, section 1.1.2, line 2: “hereditary and transmissible” for “hereditary as and transmissible”
p 2, section 1.1.2, line 4: “include” for “includes”
p 34, last line: “protein-containing” for “protein containing”
p 57, line 2: “investigation of the aggregation” for “investigation the aggregation”
p 73, 2nd paragraph, line 6: “proportion” for “proportions”
p 92, last paragraph, line 1: “Although HX is commonly” for “Although commonly”
p 93, 2nd paragraph, line 1: “sequence-specific” for “sequence specific”
p 93, 2nd paragraph, line 2: “side chains” for “side chain”
p 121, 2nd paragraph, line 4: “have been characterized” for “has been characterized”

ADDENDUM

p 23, 2nd paragraph, line 5, ‘can cause’ for ‘causes’
p 23, 2nd paragraph, add to end ‘However in a recent transgenic mouse model, it has been demonstrated that wild type ataxin-3 does not abrogate the effects of pathogenic-length ataxin-3 (300).’
p 40, 2.1.3, add the following to the oligonucleotide list:

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>5’ – 3’ sequence</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>L169Qfwd</td>
<td>GTTAAGGGTGATCAGCAGATTGCGAA</td>
<td>60 ºC</td>
</tr>
<tr>
<td>L169Qrev</td>
<td>TTCGCAATCTGGGATCACCCTTAAC</td>
<td>60 ºC</td>
</tr>
<tr>
<td>L169Ifwd</td>
<td>GTTAAGGGTGATCCGATTGCGAA</td>
<td>58 ºC</td>
</tr>
<tr>
<td>L169Irev</td>
<td>TTCGCAATCTGGATTATCACCCTTAAC</td>
<td>58 ºC</td>
</tr>
<tr>
<td>L169Tfwd</td>
<td>GTTAAGGGTGATCCGATTGCGAA</td>
<td>60 ºC</td>
</tr>
<tr>
<td>L169Trev</td>
<td>TTCGCAATCTGGGATATCACCCTTAAC</td>
<td>60 ºC</td>
</tr>
<tr>
<td>Q24Gfwd</td>
<td>AATAACTTATTGGGAGGAATTATT</td>
<td>51 ºC</td>
</tr>
<tr>
<td>Q24Grev</td>
<td>AAAATATTCTCCTCCGATTT</td>
<td>51 ºC</td>
</tr>
</tbody>
</table>

p 68, line 2, add ‘In the representative membrane shown (Fig 4 A), ataxin-3(Q64) showed a lower signal at 65 hrs compared to 48 hrs. This may be due to the sample dropping out of solution, however as observed in Fig 4 B, was not the norm over the 3 independent experiments.’
p 70, 2nd paragraph, line 3, ‘are separable’ for ‘are not independent’
p 70, 2nd paragraph, line 4, ‘wherein’ for ‘whereby’
p 71, sentence 2, add ‘and that aggregation is not necessarily impacted by changes in stability.’
p 73, 4th paragraph, line 3, add to end ‘In addition, the catalytic activity of the Josephin domain may contribute to the resistance of aggregation within the cell, as demonstrated in Drosophila (Warrick et al., 2005).’
p 81, Figure 2 legend, line 3, change “Representative traces are” to “Representative traces of three independent experiments”
p 81, Figure 3 legend, add to end ‘Traces are shown until a timepoint where they reach plateau, up to 150 hrs.’
p 81, Figure 4 legend, add to end, ‘Error bars represent SEM.’
p 82, Table 2, row 1, change ‘(°C)’ to ‘(°C ± SEM)’ and ‘(hrs)’ to ‘(hrs ± SEM)’
p 93, Figure 4.1 legend line 5: change “carbonyl” to “aliphatic”
p 101, 2nd paragraph, after sentence 4, insert ‘This suggests that Josephin is folded in 50 % acetonitrile, pH 2.0. This is unexpected considering that for several ataxin-3 variants, low pH was previously found to
unfold the protein (69), and highlights the different properties of the Josephin domain when independent of the full length protein.’

p 111, 2nd paragraph, add to end ‘Previously, Masino et al. have determined that Josephin unfolding is reversible, thus allowing ΔG to be calculated (208).’

p 112, Figure 4.12 legend, add to end ‘Error bars for both A and B represent SEM from three independent experiments.’

p 132, line 6, ‘aggregate’ for ‘fibril’

p 142, Figure 3 legend, add to end ‘Error bars represent SEM.’

p 143, Figure 6 legend, part (A), line 3, add ‘The numbers beside the membranes correspond to different lipids as follows: 1-Lysophosphatidic acid, 2-Lysophosphatidlycholine, 3-Phosphatidylinositol (PtdIns), 4-PtdIns(3)P, 5-PtdIns(4)P, 6-PtdIns(5)P, 7-Phosphatidylethanolamine, 8-Phosphatidylcholine, 9-Sphingosine 1-phosphate, 10-PtdIns(3,4)P, 11-PtdIns(3,5)P, 12-PtdIns(4,5)P, 13-PtdIns(3,4,5)P, 14-Phosphatidic acid, 15-Phosphatidylylserine, 16-blank.’

p 157, 1st paragraph, sentence 5, add ‘and that SDS may be pushing aggregation down an alternative pathway.’

p 167, reference 33, ‘(2010) for (2009)’ and add to end ‘395, 717-27’

CHAPTER 1
INTRODUCTION ................................................................. 1

1.1 INTRODUCTION TO AMYLOID DISEASE ................................................................. 1
  1.1.1 Conformational Disorders ................................................................. 1
  1.1.2 Amyloid Disease ............................................................................ 2
  1.1.3 Amyloid Fibril Structure ................................................................. 4
  1.1.4 Species Formed Along the Aggregation Pathway ....................... 8
  1.1.5 Kinetics of Aggregation ................................................................. 10
  1.1.6 Initial Misfolding Events ................................................................. 11

1.2 POLYGLUTAMINE DISEASE ........................................................................ 15
  1.2.1 General Characteristics of PolyQ disease ................................... 15
  1.2.2 Localization of Inclusions in PolyQ disease .............................. 17
  1.2.3 Role of Proteolysis in PolyQ Disease ........................................ 18
  1.2.4 Identification of Toxic Species in PolyQ Disease ....................... 19

1.3 SPINOCEREBELLAR ATAXIA TYPE 3 (SCA3) ........................................ 21
  1.3.1 General Characteristics of SCA3 ............................................... 21
  1.3.2 Ataxin-3 Structure and Function ............................................ 22
  1.3.3 Molecular Characteristics of SCA3 ........................ .................. 25

1.4 POLYGLUTAMINE AGGREGATION ................................................................. 27
  1.4.1 Introduction to the PolyQ Aggregation Pathway ....................... 27
  1.4.2 PolyQ Aggregation and Protein Context ................................ 28
  1.4.3 Interplay Between the PolyQ Domain and the Stability of Adjacent Regions .... 30
  1.4.4 Multi-domain Misfolding and Aggregation ............................ 31

1.5 AIMS OF THIS STUDY ................................................................................... 37
CHAPTER 2
MATERIALS AND METHODS ........................................................................................................................................... 39
2.1 DNA CLONING AND MICROBIAL TECHNIQUES ................................................................................................................................. 39
  2.1.1 Reagents and Enzymes .............................................................................................................................................................................. 39
  2.1.2 2xTY Growth Medium ................................................................................................................................................................................ 39
  2.1.3 Oligonucleotide Design ................................................................................................................................................................................ 40
  2.1.4 QuikChange Site-Directed Mutagenesis .................................................................................................................................................... 40
  2.1.5 Agarose Gel Electrophoresis ................................................................................................................................................................. 41
  2.1.6 Plasmid Purification .................................................................................................................................................................................. 42
  2.1.7 Preparation of Competent Cells ............................................................................................................................................................... 42
  2.1.8 Bacterial Transformations for Expression .................................................................................................................................................. 43
2.2 PROTEIN TECHNIQUES ................................................................................................................................................................................. 43
  2.2.1 Reagents .................................................................................................................................................................................................. 43
  2.2.2 Expression of Ataxin-3 Variants .................................................................................................................................................................. 44
  2.2.3 Protein Purification ...................................................................................................................................................................................... 45
    2.2.3.1 Concentration Determination .................................................................................................................................................................. 46
  2.2.4 Protein Pre-Analysis Techniques ................................................................................................................................................................. 47
    2.2.4.1 SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE) .................................................................................................................................. 47
    2.2.4.2 Western Blot .............................................................................................................................................................................................. 48
    2.2.4.3 Size Exclusion Chromatography .................................................................................................................................................. 49
    2.2.4.4 Ubiquitin Protease Activity Assay ...................................................................................................................................................... 49
  2.2.5 Protein Analysis Techniques ..................................................................................................................................................................... 50
    2.2.5.1 Circular Dichroism (CD) ........................................................................................................................................................................ 50
    2.2.5.2 Equilibrium Denaturation .................................................................................................................................................................... 51
    2.2.5.3 Mass spectrometry (MS) ....................................................................................................................................................................... 52
  2.2.6 Aggregation Assays .................................................................................................................................................................................... 53
    2.2.6.1 ThioflavinT Fluorescence ....................................................................................................................................................................... 53
    2.2.6.2 Filter Trap Assay .................................................................................................................................................................................... 54
    2.2.6.3 Transmission Electron Microscopy .................................................................................................................................................. 55
    2.2.6.4 Phospholipid Overlay Assay .............................................................................................................................................................. 55
CHAPTER 3
THE IMPACT OF FLANKING DOMAIN STABILITY ON AGGREGATION KINETICS OF A POLYQ PROTEIN ................................................................................................................................. 57
CHAPTER 4
CONFORMATIONAL DYNAMICS OF THE JOSEPHIN DOMAIN OF ATAXIN-3 ................................................................................................................. 89
4.1 INTRODUCTION ............................................................................................................................................................................................... 89
  4.1.1 Protein Aggregation and Dynamics ............................................................................................................................................................... 89
  4.1.2 Ataxin-3 Aggregation and Dynamics .................................................................................................................................................... 90
  4.1.3 Theory of Hydrogen Exchange Mass Spectrometry ................................................................................................................................ 92
4.2 Results ..................................................................................................................96
  4.2.1 Josephin MS under Native Conditions .............................................................96
  4.2.2 Josephin Exists in Multiple Conformations in Solution .................................98
    4.2.2.1 Multiple Charge State Envelopes are Not Linked to Parameters on the MS98
    4.2.2.2 Denaturing Conditions Shift the Charge State Envelope .........................101
    4.2.2.3 The Partially Unfolded Conformations are Not Linked to Aggregation ....102
  4.2.3 Josephin HX MS .............................................................................................104
    4.2.3.1 Deuteration of Josephin ...........................................................................104
    4.2.3.2 D→H Exchange ......................................................................................106
    4.2.3.3 H→D Exchange using NanoESI-MS Demonstrates the Multiple
      Conformations of Josephin Exist in Equilibrium ............................................108
    4.2.3.4 Analysis of Josephin HX Kinetics ............................................................110

4.3 Discussion ..........................................................................................................113
  4.3.1 Implications of the Conformational Heterogeneity of Josephin .......................113
  4.3.2 The Dynamic Nature of the Josephin domain ..................................................115

Chapter 5
Aggregation of Ataxin-3 in SDS: Evidence for Alternative Aggregation Pathways ..........117

Chapter 6
General Discussion & Conclusions .............................................................................155
  6.1 Role of Flanking Domains in PolyQ Aggregation ..............................................155
  6.2 Relationship between Folding and Aggregation Pathways of Ataxin-3 ..........159
  6.3 Multi-domain Misfolding and PolyQ Disease .......................................................162
  6.4 General Conclusions .........................................................................................163

Chapter 7
Reference List ............................................................................................................165

Appendices ..............................................................................................................187
  Appendix 1: Small Heat-shock Proteins Interact with a Flanking Domain to
              Suppress Polyglutamine Aggregation ..........................................................188
  Appendix 2: Multi-domain Misfolding: Understanding the Aggregation
              Pathways of Polyglutamine Proteins .............................................................194
ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor Steve for his support and encouragement over the last four years. In particular, I would like to thank him for giving me the freedom to work on the projects which really interested me whilst keeping me realistic as to how many projects I could fit into four years! In addition, I thank my co-supervisor Roberto for making the trip to Monash numerous times in order to stay up to date with my project and provide advice.

The Bottomley lab has been a great place to work, and I would like to thank everyone who went out of their way to help me in some way or share their knowledge. In particular I would like to thank the polyQ people – Andy, Michelle, Bronwen, Amy, SuLing, Glyn and Victoria - for the many useful discussions. I would like to thank all my collaborators at Monash – Phil and Cathy Bird, Wei Wen Dai, Andy and Amy – for their assistance with various projects.

I spent a fantastic 5 months at Leeds University, and I would like to thank Sheena for making me so welcome in her lab, Johnny and David for sharing their mass spec knowledge, and Carol, Wei-Feng and Jenny for some great weekend trips away.

My family and friends have all been a great support to me over the years and I thank them for their understanding all those times I had to run away to take a sample at the lab.

Lastly, I have to thank my partner Jonathon for his support during the last four years - I appreciated your company on the many late night and weekend trips to the lab and couldn’t have done it without you.
GENERAL DECLARATION

In accordance with Monash University Doctorate Regulation 17/Doctor of Philosophy (PhD) and Master of Philosophy (MPhil) regulations the following declarations are made:

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where appropriate reference is made in the text of the thesis.

This thesis contains two unpublished publications. The core theme of the thesis is the characterization of the aggregation pathway of ataxin-3. The ideas, development and writing up of all the papers in the thesis was the principal responsibility of myself, the candidate, working within the Department of Biochemistry and Molecular Biology under the supervision of Professor Stephen Bottomley.

The inclusion of co-authors reflects the fact that the work came from active collaborations between researchers and acknowledges input into team-based research.

In the case of Chapters 3 and 5, my contribution to the work involved the following:

<table>
<thead>
<tr>
<th>Thesis Chapter</th>
<th>Publication Title</th>
<th>Publication Status</th>
<th>Nature and Extent of My Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>The impact of flanking domain stability on aggregation kinetics of a polyQ protein.</td>
<td>Submitted to Protein Science</td>
<td>Mutagenesis, protein purification, experimentation, data analysis and manuscript preparation</td>
</tr>
<tr>
<td>5</td>
<td>Aggregation of ataxin-3 in SDS: evidence for alternative aggregation pathways</td>
<td>Submitted to Journal of Molecular Biology</td>
<td>Protein purification, experimentation, data analysis and manuscript preparation</td>
</tr>
</tbody>
</table>

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Helen Saunders
21 May 2010
DECLARATION BY CO-AUTHORS FOR CHAPTER 3

In the case of Chapter 3, the following co-authors contributed to the work:

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of Contribution</th>
<th>Extent of Contribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helen M Saunders</td>
<td>Mutagenesis, protein purification, experimentation, data analysis and manuscript preparation</td>
<td>80%</td>
</tr>
<tr>
<td>Amy L Robertson</td>
<td>General advice and refinements to the manuscript</td>
<td>5%</td>
</tr>
<tr>
<td>Dimitri Giles</td>
<td>PoPMuSiC analysis and refinements to the manuscript</td>
<td></td>
</tr>
<tr>
<td>Marianne Rooman</td>
<td>PoPMuSiC analysis</td>
<td></td>
</tr>
<tr>
<td>Yves Dehouck</td>
<td>PoPMuSiC analysis</td>
<td></td>
</tr>
<tr>
<td>Stephen P Bottomley</td>
<td>General advice, supervision and refinements to the manuscript</td>
<td></td>
</tr>
</tbody>
</table>

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate’s contribution to this work, and the nature of the contribution of each of the co-authors

(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;

(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

(4) there are no other authors of the publication according to these criteria;

(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and

(6) the original data are stored at the following location and will be held for at least five years from the date indicated below:

**Location:** Monash University, Department of Biochemistry and Molecular Biology, Clayton Campus

**Date:** 21/5/2010
DECLARATION BY CO-AUTHORS FOR CHAPTER 5

In the case of Chapter 5, the following co-authors contributed to the work:

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of Contribution</th>
<th>Extent of Contribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helen M Saunders</td>
<td>Mutagenesis, protein purification, experimentation, data analysis and manuscript preparation</td>
<td>80%</td>
</tr>
<tr>
<td>Andrew M Ellisdon</td>
<td>General advice and refinements to the manuscript</td>
<td>5%</td>
</tr>
<tr>
<td>Roberto Cappai</td>
<td>General advice and refinements to the manuscript</td>
<td></td>
</tr>
<tr>
<td>Stephen P Bottomley</td>
<td>General advice, supervision and refinements to the manuscript</td>
<td></td>
</tr>
</tbody>
</table>

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate’s contribution to this work, and the nature of the contribution of each of the co-authors
(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
(4) there are no other authors of the publication according to these criteria;
(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
(6) the original data are stored at the following location and will be held for at least five years from the date indicated below:

Location: Monash University, Department of Biochemistry and Molecular Biology, Clayton Campus
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFM</td>
<td>Atomic Force Microscopy</td>
</tr>
<tr>
<td>ANS</td>
<td>Anilinonaphthalene Sulfonate</td>
</tr>
<tr>
<td>APS</td>
<td>Ammonium Persulfate</td>
</tr>
<tr>
<td>AR</td>
<td>Androgen Receptor</td>
</tr>
<tr>
<td>A-β</td>
<td>Amyloid-β</td>
</tr>
<tr>
<td>β-me</td>
<td>β-mercaptoethanol</td>
</tr>
<tr>
<td>BCA</td>
<td>Bicinchoninic Acid</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>CACNA1A</td>
<td>Alpha 1A-voltage-dependent Calcium Channel</td>
</tr>
<tr>
<td>CBP</td>
<td>Creb-Binding Protein</td>
</tr>
<tr>
<td>CD</td>
<td>Circular Dichroism</td>
</tr>
<tr>
<td>CMC</td>
<td>Critical Micelle Concentration</td>
</tr>
<tr>
<td>CRABP I</td>
<td>Cellular Retinoic Acid-binding Protein I</td>
</tr>
<tr>
<td>CSE</td>
<td>Charge State Envelope</td>
</tr>
<tr>
<td>CV</td>
<td>Column Volumes</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>DNaseI</td>
<td>Deoxyribonuclease I</td>
</tr>
<tr>
<td>DRPLA</td>
<td>Dentatorubral-Pallidoluysian Atrophy</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>DUB</td>
<td>De-ubiquitinating Enzyme</td>
</tr>
<tr>
<td>ε</td>
<td>Extinction Co-efficient</td>
</tr>
<tr>
<td>ECL</td>
<td>Enymatic Chemiluminescence</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic Acid</td>
</tr>
<tr>
<td>EGCG</td>
<td>(-)-Epi-Gallocatechine Gallate</td>
</tr>
<tr>
<td>EPR</td>
<td>Electron Paramagnetic Resonance</td>
</tr>
<tr>
<td>ESI-MS</td>
<td>Electrospray Mass Spectrometry</td>
</tr>
<tr>
<td>FPLC</td>
<td>Fast Protein Liquid Chromatography</td>
</tr>
<tr>
<td>ΔG</td>
<td>Free Energy of Unfolding</td>
</tr>
<tr>
<td>ΔG_{ex}^{app}</td>
<td>Apparent Free Energy of Exchange</td>
</tr>
<tr>
<td>GdnHCl</td>
<td>Guanidine Hydrochloride</td>
</tr>
<tr>
<td>HD</td>
<td>Huntington's Disease</td>
</tr>
<tr>
<td>His6</td>
<td>Hexa-histidine</td>
</tr>
<tr>
<td>HX</td>
<td>Hydrogen Exchange</td>
</tr>
<tr>
<td>IAPP</td>
<td>Islet Amyloid Polypeptide</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-β-D-Thiogalactopyranoside</td>
</tr>
<tr>
<td>k_{acid}</td>
<td>Acid Catalyzed Second Order Rate Constant</td>
</tr>
<tr>
<td>k_{base}</td>
<td>Base Catalyzed Second Order Rate Constant</td>
</tr>
<tr>
<td>k_{int}</td>
<td>Intrinsic Exchange Rate</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>$k_{obs}$</td>
<td>Observed Rate of Exchange</td>
</tr>
<tr>
<td>$k_w$</td>
<td>Water Catalyzed Second Order Rate Constant</td>
</tr>
<tr>
<td>m/z</td>
<td>Mass to Charge Ratio</td>
</tr>
<tr>
<td>MS</td>
<td>Mass Spectrometry</td>
</tr>
<tr>
<td>NII</td>
<td>Neuronal Intranuclear Inclusions</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>OD$_{600 \text{ nm}}$</td>
<td>Optical Density at 600 nm</td>
</tr>
<tr>
<td>PABPN1</td>
<td>PolyA Binding Protein Nuclear 1</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffered Saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PD</td>
<td>Parkinson's Disease</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethyl Sulphonyl Fluoride</td>
</tr>
<tr>
<td>polyA</td>
<td>Polyalanine</td>
</tr>
<tr>
<td>polyQ</td>
<td>Polyglutamine</td>
</tr>
<tr>
<td>PoPMuSiC</td>
<td>Prediction of Proteins Mutations Stability Changes</td>
</tr>
<tr>
<td>PPII</td>
<td>Proline Type II Helix</td>
</tr>
<tr>
<td>PtdIns</td>
<td>Phosphotidylinositol</td>
</tr>
<tr>
<td>PUF</td>
<td>Partially Unfolded Forms</td>
</tr>
<tr>
<td>QBP1</td>
<td>PolyQ Binding Peptide 1</td>
</tr>
<tr>
<td>R</td>
<td>Gas Constant</td>
</tr>
<tr>
<td>SBMA</td>
<td>Spinal and Bulbar Muscular Atrophy</td>
</tr>
<tr>
<td>SCA</td>
<td>Spinocerebellar Ataxia</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium Dodecyl Sulfate</td>
</tr>
<tr>
<td>SEC</td>
<td>Size Exclusion Chromatography</td>
</tr>
<tr>
<td>SpA</td>
<td>Staphylococcus Protein A</td>
</tr>
<tr>
<td>T</td>
<td>Absolute Temperature</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris acetate - EDTA buffer</td>
</tr>
<tr>
<td>TBP</td>
<td>TATA-Box-Binding Protein</td>
</tr>
<tr>
<td>TBSG</td>
<td>Tris Buffered Saline with Glycerol</td>
</tr>
<tr>
<td>TEM</td>
<td>Transmission Electron Microscopy</td>
</tr>
<tr>
<td>TEMED</td>
<td>N,N,N',N'-tetramethylethylenediamine</td>
</tr>
<tr>
<td>TFA</td>
<td>Tri-fluoro Acetate</td>
</tr>
<tr>
<td>thioT</td>
<td>Thioflavin T</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>UIM</td>
<td>Ubiquitin Interaction Motif</td>
</tr>
<tr>
<td>UPS</td>
<td>Ubiquitin Proteosomal System</td>
</tr>
<tr>
<td>VCP</td>
<td>Valosin-Containing protein</td>
</tr>
</tbody>
</table>
Spinocerebellar Ataxia Type 3 (SCA3) is one of nine polyglutamine (polyQ) diseases which are all characterized by progressive neuronal dysfunction. A hallmark of these diseases is the presence of neuronal inclusions which contain aggregated polyQ protein, implicating protein misfolding as a key part of this disease. The only sequence homology which exists between the polyQ proteins is the polyQ tract, thus suggesting that protein context is responsible for dictating the specific characteristics of each disease. Increasingly, flanking domains to the polyQ tract have been demonstrated to modulate the aggregation of polyQ proteins. Fibrillogenesis of ataxin-3, the causative protein of SCA3, occurs via a multi-stage pathway when studied in vitro. The isolated Josephin domain of ataxin-3 has the intrinsic ability to aggregate under native conditions and thus the multi-stage model of ataxin-3 aggregation has proposed the first stage of aggregation is mediated by interactions of the Josephin domain. This thesis has further characterized the multi-stage aggregation pathway of ataxin-3, with a focus on the role of the Josephin domain during misfolding and aggregation.

This thesis initially determined that changing the stability of the flanking Josephin domain dramatically impacts upon ataxin-3 fibrillogenesis in vitro. Mutations within the Josephin domain that increased stability led to decreased aggregation rates and vice versa. This provided support to the hypothesis that Josephin domain interactions are involved in the first stage of ataxin-3 aggregation. The stability mutants also distinguished between two previously proposed models of misfolding to demonstrate that ataxin-3 aggregates from an intermediate which is on the folding pathway. Defining the conformational landscape of ataxin-3 is one mechanism which may provide insights into the initial
misfolding events involved in aggregation. As Josephin domain interactions were demonstrated to mediate the first stage of aggregation, hydrogen exchange mass spectrometry was used to investigate the dynamics of this folded domain. The Josephin domain was found to be highly dynamic, and under native conditions, exists in rapid equilibrium with several partially unfolded conformations. From these data, it is proposed the unfolded conformations are not involved in the fibrillogenic pathway, as mutants with either increased or decreased aggregation rates do not change the proportions of the partially unfolded conformations. In addition, the proportion of the partially unfolded conformations was unaffected by several protein interaction partners of ataxin-3. Therefore, native ataxin-3 exists within a more complex conformational landscape than previously thought, which is potentially linked to the normal function of the protein. In a further exploration of the conformational landscape of ataxin-3, addition of the membrane mimetic sodium dodecyl sulfate resulted in a number of alternative aggregation pathways. Ataxin-3 was found to bind several phospholipids only when in a fibrillar conformation, which may be linked to the mechanism of toxicity, as has been reported for other amyloidogenic proteins.

Together, these studies have demonstrated that native ataxin-3 exists within a complex conformational landscape, and provided evidence for a multi-domain aggregation mechanism for ataxin-3. This thesis has added to the growing body of evidence that suggests flanking domains influence the misfolding and aggregation of polyQ disease proteins.
1.1 INTRODUCTION TO AMYLOID DISEASE

1.1.1 Conformational Disorders

The correct folding of proteins within the cell is a complex process requiring a significant number of accessory proteins such as chaperones and folding catalysts (1). However even after a protein has obtained its native structural state, subsequent inappropriate conformational changes can occur which may result in disease. In these diseases (termed the conformational disorders), functional proteins undergo a change which results in their self association and irreversible tissue deposition (2). The mechanism of toxicity for these diseases can be due to both a gain-of-function and loss-of-function. Deposition of protein can cause toxicity via the uptake of cellular components into the deposits, resulting in cellular dysfunction and death, whilst there can also be a loss-of-
function of the aggregating protein due to its adoption of inactive conformations. There are a wide range of diseases which are grouped as conformational diseases, and all typically involve a late and gradual onset of pathology.

1.1.2 Amyloid Disease

The amyloidoses are a subset of the conformational disorders, and include sporadic, hereditary as and transmissible diseases (Table 1.1). Amyloid diseases can be systemic, such as light chain and dialysis-related amyloidoses, or localized as for type II diabetes, and includes a number of neurodegenerative diseases such as Alzheimer’s, Parkinson’s (PD) and prion diseases. The pathological hallmark of these neurodegenerative diseases is the presence of insoluble, ordered proteinaceous deposits or ‘amyloid fibrils’ within the brain (Fig 1.1). The presence of these deposits has implicated a role for protein misfolding in disease, which has led to much interest in characterizing both the deposited amyloid fibrils and the mechanism via which fibrils form. Although these deposits predominantly consist of amyloid fibrils formed by a specific protein, a number of other proteins and carbohydrates also form part of the deposits (3).

Figure 1.1 Morphology of Proteinaceous Deposits and Amyloid Fibrils. (A) Proteinaceous deposits within the brain, as highlighted by the arrow, are a characteristic of a number of neurodegenerative amyloid diseases. The arrow highlights a nuclear inclusion formed in Spinocerebellar Type 3 (SCA3) (4). Within the deposits, the aggregated protein is present in ordered, fibrillar species as shown in the TEM images of fibrils formed from (B) islet amyloid polypeptide (IAPP) (5), (C) amyloid-β (6), (D) α-synuclein (7) and (E) huntingtin exon-1 (8).
Table 1.1 Amyloid Diseases

<table>
<thead>
<tr>
<th>Disease</th>
<th>Aggregating Protein</th>
<th>Nature of Deposits</th>
<th>Main Location of Deposits</th>
<th>Nature of Inheritance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer’s Disease</td>
<td>Amyloid-β</td>
<td>Localized</td>
<td>Brain - Hippocampus, cerebral cortex</td>
<td>Sporadic; Acquired</td>
</tr>
<tr>
<td>Parkinson Disease</td>
<td>α-synuclein</td>
<td>Localized</td>
<td>Brain - Substantia nigra, hypothalamus</td>
<td>Sporadic; Hereditary</td>
</tr>
<tr>
<td>Huntington Disease</td>
<td>Huntingtin</td>
<td>Localized</td>
<td>Brain - Striatum, cerebral cortex</td>
<td>Hereditary</td>
</tr>
<tr>
<td>Creutzfeld-Jacob Disease and other prionopathies</td>
<td>Prion protein</td>
<td>Localized</td>
<td>Central nervous system, brain</td>
<td>Sporadic; Transmissible; Hereditary</td>
</tr>
<tr>
<td>Dialysis-related amyloidoses</td>
<td>β-2-microglobulin</td>
<td>Systemic</td>
<td>Musculoskeletal tissues</td>
<td>Acquired</td>
</tr>
<tr>
<td>Dialysis-related amyloidoses</td>
<td>Lysozyme</td>
<td>Systemic</td>
<td>Kidney, liver, spleen</td>
<td>Hereditary</td>
</tr>
<tr>
<td>Familial amyloid polyneuropathy</td>
<td>Transthyretin</td>
<td>Systemic</td>
<td>Most organs and tissues - heart, vessels</td>
<td>Hereditary; Acquired</td>
</tr>
</tbody>
</table>

Despite the formation of fibrillar aggregates with similar properties in these diseases, protein deposition and neuronal loss occurs in different brain regions for each particular disease (Table 1.1). The reasons for this are not fully understood, however it may be due to factors such as local changes in protein expression, different concentrations or activity of chaperones or differences in clearance pathways such as autophagy, which is known to ameliorate toxicity associated with protein aggregation (9). Amyloid was traditionally defined as the extracellular deposition of protein fibrils (3), however this definition has now been expanded to include protein fibrils which are deposited intracellularly (10).
Regardless of whether formed by a peptide, unstructured polypeptide or natively globular protein, the end point of the aggregation pathway, the amyloid fibril, displays a number of key defining characteristics. As determined by transmission electron microscopy (TEM) (Fig 1.1), amyloid fibrils have a long, straight morphology, with a typical diameter of 7-13 nm (11). In addition to this, amyloid fibrils show a number of tinctoral properties. Under polarized light fibrils show green birefringence upon binding to congo red, and thioflavin S and thioflavin T both hyperfluoresce upon binding to fibrils (12). These characteristic features of amyloid fibrils are well established and recent advances in high resolution techniques such as solid state NMR (ssNMR) and X-ray crystallography are also providing new insights into the structure of fibrils at a molecular level.

1.1.3 Amyloid Fibril Structure

Two main models of fibril structure have been proposed; the traditional cross β-sheet model and the β-helix model (Fig 1.2). X-ray diffraction data has played a key role in determining these models, as have a multitude of lower resolution techniques, such as TEM, atomic force microscopy (AFM), electron paramagnetic resonance (EPR), scanning TEM, hydrogen deuterium exchange and limited proteolysis. These techniques provide important constraints to the modeling of the fibril at a molecular level from the higher resolution techniques. At a supramolecular level, cryoelectron microscopy has been used to characterize the structure of fibrils. For example, using this technique insulin fibrils have been found to be a heterogeneous mix of fibrils containing either 2, 4, or 6 protofilaments twisted together (13).
Evidence for the traditional cross-β sheet model comes from X-ray diffraction studies which have shown that fibrils have a common structure whereby the β-strands are positioned 4.8 Å apart and perpendicular to the long axis of the fibril. β-sheets also exist parallel to the fibril long axis and are 8-11 Å apart depending on the side chains involved (14,15). There is no sequence or structural homology between amyloid forming proteins, and thus typically it is main-chain interactions that stabilize the core of the fibril. This supports the hypothesis that under suitable conditions, all polypeptides have an inherent ability to form fibrils (16,17).

Figure 1.2 Cross-β and β-helix Fibril Models. The two main models of fibril structure are (A) the cross-β and (B) the β-helix model. Both models have β-strands which are 4.8 Å apart and perpendicular to the long axis of the fibril, however only the cross β-sheet model has an intersheet distance of 8 – 11 Å. Images are adapted from (13,18) respectively.

Whilst the cross-β structure is generic to amyloid fibrils, the ultrastructure of the fibril differs between different disease proteins. The recent advances in ssNMR have led to the characterization of protofilament structure at a molecular level, and have provided support for the cross-β structure proposed through X-ray diffraction studies. Although not identical, several proteins including amyloid-β (A-β), HET-s prion(218-289), and
transthyretin(105-115) have been shown to have a similar protofilament structure consisting of two layers of parallel β-sheets (Fig 1.3) (6,19,20). However, demonstrating the differences which exist between disease fibrils, the fibril structure of the α-synuclein(30-110) fragment has five layers of parallel in-register β-sheets (21). In many cases only fragments of disease proteins have been used for ssNMR due to technical difficulties, however for A-β, studies have shown that smaller fragments of this peptide results in anti-parallel β-sheets, whilst full length A-β results in parallel β-sheets (6,22-24).

![Figure 1.3 A-β(1-40) Fibril Model. Structure of an A-β(1-40) fibril modeled using ssNMR data. Looking down the long axis of the fibril shows each molecule forms a β-strand – turn – β-strand motif which forms a parallel β-sheet, (adapted from (6)) and two such parallel β-sheets form a protofilament.](image)

Mature fibrils are not amenable to structure determination using X-ray crystallography. Microcrystals have however been formed from small fibrillogenic peptides which have similar diffraction patterns to, and which are able to seed, fibrils and thus provide insight into the side-chain interactions and packing of strands in the fibril at a molecular level (25-27). For example, using a specifically designed peptide, Serpell and coworkers found that beyond the inherent H-bonded network of the cross-β motif, aromatic residue sidechain stacking and charge interactions were important for stabilizing fibrils (25).
The second main fibril model is the β-helix model (Fig 1.2B) where fibrils have parallel β-strands 4.8 Å apart and perpendicular to the fibril axis, which form a helix with a hollow centre (18,28). A β-helix structure has been proposed for several disease fibrils, including polyglutamine (polyQ), prion and β₂-microglobulin fibrils (29-31). Perutz and coworkers have proposed that the structure of polyQ fibrils are a water-filled nanotubule, based on X-ray diffraction data whereby the equatorial diffraction signal was smaller than is typically observed for amyloid fibrils and thus excluded from analysis (31). However, these data have been reanalyzed by others who propose that it is consistent with a cross-β structure whereby the polar polyQ sidechains hydrogen bond with the main-chain and thus allow dense packing (32). The second analysis of the data is also supported by molecular modeling of X-ray diffraction data, and it is thought that the different diffraction pattern may have resulted from an atypical data collection method, whereby the sample had a film texture, as opposed to the typical fiber structures generally used (33).

Amyloid fibrils exhibit a common core cross-β structure, however there are a number of differences between different fibrils, such as the inter sheet differences seen in the X-ray diffraction equatorial reflection which ranges from 8-11 Å (15), due to factors such as the steric bulk of side chains (34). In addition to this, the proportion of the polypeptide chain which is incorporated into the fibril core varies greatly, ranging from 100 % of fibrillogenic 7mer peptides to approximately 50 % of A-β and 13 % of HET-s prion protein (6,20,27). The question of whether the β-sheets are arranged parallel or anti-parallel within the fibril differs for each protein, with both having been observed. Margittai and Langen suggest that it depends upon the number of amino acids within the fibril core, with less than 20 amino acids leading to anti-parallel β-sheets and more than 20 resulting in
parallel β-sheets, due to the stereochemistry involved (35). Therefore it appears that although fibril formation may be a generic property of all proteins, the type of fibril formed will differ depending on the protein sequence and size.

**1.1.4 Species Formed Along the Aggregation Pathway**

A poor correlation exists between the mass of fibrillar material deposited in the tissue and the presentation of clinical features (36,37). This implicates a role for other factors in pathogenesis, and one aspect into which there has been considerable interest is the species formed along the aggregation pathway. To fully understand the aggregation pathway requires determination of the thermodynamic and kinetic properties of the species populated during aggregation, a number of which are common to amyloidogenic proteins (38). The pre-fibrillar species common to most amyloid pathways include the oligomer, protofibril and protofilament.

![Figure 1.4 Structural Species Formed During Aggregation](image)

**Figure 1.4 Structural Species Formed During Aggregation.** The conformational species formed during the process of fibril formation include the oligomer, protofibril and protofilament. All these species typically contain β-sheet rich structures. Pre-fibrillar oligomers are also formed by some amyloidogenic proteins.

The initial species formed in the aggregation pathway for most proteins is the oligomer (Fig 1.4). Some proteins such as A-β, Ure2p and insulin form soluble pre-fibrillar
oligomers which retain native structure, before forming an oligomer bearing cross-β-like structure (39-41). The pre-fibrillar oligomer has not, however, been observed for all proteins, and other proteins progress directly to a cross-β oligomer, which typically has a diameter of 2-5 nm (38). The transient nature of the cross-β oligomer limits structural analysis, however, oligomers from numerous amyloid proteins – both disease and non-disease related – form a common structure which is distinct to the mature fibril. This is evidenced by the ability of oligomers to bind the A11 antibody (42). Although the epitope of this antibody remains unknown, it has been demonstrated to bind to oligomers but not fibrils (42). Depending on the protein, oligomers may be a heterogeneous mix of structures, or can be more defined such as the distinct ‘donut’ shaped species formed by several proteins, although not all oligomers progress to form fibrils (43).

Elongation of the on-pathway oligomers results in short, thin curvilinear protofibrils, as visualized using TEM and AFM, which typically have a similar diameter of 3-5 nm but have increased lengths of 20-100 nm (44,45). These protofibrils then form larger protofilaments which are the units of the mature fibril, with 2-6 protofilaments per fibril which generally form a long-range twist (13,46). Aside from the pre-fibrillar oligomer, the pre-fibrillar species are all characterized by a cross-β structure, and the ability to bind the dye thioflavin T. In the process of forming these species, internal structural rearrangements can occur (47), thus explaining the ability of oligomers but not protofibrils to bind to the A11 antibody (42).

In accordance with the lack of correlation between the fibril load and toxicity, increasing evidence suggests that the pre-fibrillar species are responsible for toxicity in a number of cell and animal models of disease. Both A-β transgenic mice and drosophila
models show symptoms before the accumulation of amyloid plaques (48,49), whilst non-native transthyretin oligomers have been identified as the toxic species in a cellular model (50). Genetic support for this mechanism of toxicity comes from juvenile PD, where early neurodegeneration occurs without the formation of the intracellular deposits or Lewy bodies typically seen in late onset PD (51,52). The key pre-fibrillar species involved in toxicity has been demonstrated in vitro to be the oligomer, and this appears to be a generic mechanism of toxicity as both disease and non-disease proteins form oligomers which are toxic to cultured neurons (53,54). Kayed et al. suggest that the oligomers may be cytotoxic due to their recruitment of cellular components into aggregates, their disruption of membranes and resultant depolarization events, and their exacerbation of oxidative stress (42).

1.1.5 Kinetics of Aggregation

The aggregation pathway can be modeled kinetically as a nucleated polymerization mechanism (Fig 1.5) (55). This is characterized by an initial lag phase, in which formation of a critical nucleus occurs. Formation of this nucleus is a rate-limiting step in the pathway, and the number of molecules within the critical nucleus varies depending on the protein involved (56,57). Following the lag phase, the elongation phase then involves rapid extension of the initial critical nucleus, through the pre-fibrillar species, to form the mature fibril. Kinetic modeling suggests that the mechanism of association can occur through monomer addition or association of oligomers (58,59).

In support of the nucleated growth mechanism, the addition of pre-formed species as ‘seed’ eliminates the lag phase as it removes the need for the formation of the thermodynamically unfavourable critical nucleus before extension can occur (60).
Interestingly, the efficiency of seeding varies greatly for different proteins (61-63). Although the ability to form amyloid fibrils is considered an inherent ability of all proteins, their propensity to do so differs greatly, as evidenced by differences in their aggregation kinetics. Kinetic modeling provides an important quantitative tool for determining the impact of other factors, from solution conditions to binding partners to potential drugs, on the modulation of the aggregation pathway.

**Figure 1.5 Nucleated Polymerization Aggregation Kinetics.** The aggregation curve shows nucleated polymerization kinetics with a lag phase in which the rate-limiting nucleus is formed, followed by a rapid extension phase. The dotted line shows the elimination of the lag phase upon the addition of seed.

### 1.1.6 Initial misfolding events

For fibril formation to occur, proteins must undergo a conformational change from their native state. All potential conformations formed by a protein, whether during folding or misfolding, can be described in terms of their free energy. The heterogeneity of
unfolded and partially folded conformations formed by a protein, due to inherent fluctuations of the polypeptide chain, has led to protein conformations being described within an energy landscape (64). Although all conformations are represented on the energy landscape, for most proteins the transition from conformations populated in folding to those of aggregation has not been clearly defined. A number of factors may play a role, such as the cellular environment, factors that affect the global thermodynamic or kinetic stability, or local fluctuations of the protein.

As described by Anfinsen, the typically rapid timeframe for folding indicates that proteins spontaneously fold to their native structures without searching the entirety of conformational space (65). When a protein approaches its native state, the favourable enthalpy earned from forming native-like interactions offsets the decrease in entropy paid for adopting an ordered structure. This results in protein molecules folding correctly within a short timeframe due to native-like interactions being more stable than non-native interactions (66). Therefore, during folding the free energy available to the protein decreases and thus the energy landscape of folding is represented as a funnel.

The initial events of aggregation vary depending on whether or not the protein is natively folded. Analysis of primary sequence can be very informative when predicting amyloidogenic regions in unstructured proteins, where the initial misfolding event involves formation of a partially folded amyloidogenic intermediate (67). For natively folded proteins, the conformational change hypothesis of aggregation proposes they must undergo either partial or global unfolding in order to initiate aggregation (Fig 1.6) (68). Alternatively, some proteins have been proposed to aggregate via an intermediate which is off the folding pathway (69,70). Evolution selects for burial of hydrophobic amyloidogenic
regions, and so predicted amyloidogenic regions may be buried in a stable fold and shielded from forming the interactions required for aggregation (71). For folded proteins, two mechanisms exist via which change can result in increased formation of partially unfolded intermediates and thus aggregation; these are protein destabilization and increased protein dynamics (72).

Many disease proteins have destabilizing mutations when compared to wild type non-disease causing proteins. The destabilizing mutations result in a decreased free energy of unfolding and thus a higher population of amyloidogenic intermediates. For example, mutations in α-synuclein which destabilize the protein result in increased aggregation propensity and therefore early onset of disease (73). Therefore determining the thermodynamic stability can provide insight into the mechanism of aggregation of a protein.

Local or global fluctuations are another mechanism via which some natively folded proteins sample amyloidogenic conformations. This can occur when a low energy barrier exists to forming amyloidogenic intermediates which results in them being frequently sampled but transiently populated (68). Alternatively, some proteins populate transient
amyloidogenic native-like states without crossing the free energy barrier of unfolding (74). It has been demonstrated that the residues which are key for folding can be distinct to those important for aggregation, and therefore protein conformational fluctuations may allow sampling of conformations where these residues are more exposed (71). When factors such as mutation increase the conformational dynamics this can then result in a greater proportion of the transient intermediate and thus faster aggregation kinetics. This is demonstrated by both lysozyme and transthyretin, which both can have amyloidogenic mutations which increase protein dynamics, as determined using hydrogen deuterium exchange mass spectroscopy and NMR (75-77).

In many cases, amyloidogenic intermediates necessary for aggregation form from partially unfolded intermediates which are on the folding pathway. Therefore full characterization of both folding and the transition to aggregation is important in understanding the aggregation mechanism.
1.2 POLYGLUTAMINE DISEASE

1.2.1 General Characteristics of PolyQ Disease

A subset of the amyloid disorders, the polyQ diseases are an inherited group of nine neurodegenerative disorders (Table 1.2) whereby mutation, in the form of a CAG repeat expansion, results in disease. The pathologic hallmark of the polyQ diseases is the presence of neuronal intranuclear inclusions (NII) which contain fibrillar protein (78). As for amyloid disorders in general, formation of NII is now considered to be a protective mechanism which acts by sequestering misfolded polyQ proteins from the cell (79). The NII also sequester a number of other cellular factors, including components of the ubiquitin-proteasome pathway, chaperones and transcription factors which may disturb the homeostasis of the cell (80-82).

The only homology between the nine proteins associated with polyQ disease is the expanded polyQ tract, however the characteristics of the diseases themselves show substantial similarity. With the exception of the X-linked SBMA, all the polyQ proteins show autosomal dominant inheritance (83). An inverse exponential correlation exists between repeat length and both age of disease onset and disease severity (84). Due to repeat instability, anticipation occurs in several polyQ diseases, whereby subsequent generations have an increased repeat length and thus earlier onset and increased disease severity (85-87). Each protein has a threshold glutamine repeat length, with polyQ repeat lengths of over 40 generally leading to disease. The same polyQ tract length does not result in the same age of onset between different disease proteins, suggesting disease characteristics are not solely due to the expanded polyQ tract, but are modulated by protein context. In addition to this, the polyQ diseases all show loss of different subsets of
neurons, despite most of the polyQ proteins showing widespread expression throughout the brain (88).

### Table 1.2 Characteristics of PolyQ Disease Proteins

<table>
<thead>
<tr>
<th>Disease</th>
<th>Protein</th>
<th>PolyQ Tract Location (Black) &amp; Relative Size</th>
<th>Pathogenic Q Length</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>HD</td>
<td>Huntingtin</td>
<td><img src="image" alt="Huntingtin PolyQ Tract" /></td>
<td>36 - 121</td>
<td>(89,90)</td>
</tr>
<tr>
<td>SCA1</td>
<td>Ataxin-1</td>
<td><img src="image" alt="Ataxin-1 PolyQ Tract" /></td>
<td>40 - 83</td>
<td>(91,92)</td>
</tr>
<tr>
<td>SCA2</td>
<td>Ataxin-2</td>
<td><img src="image" alt="Ataxin-2 PolyQ Tract" /></td>
<td>35 - 59</td>
<td>(93)</td>
</tr>
<tr>
<td>SCA3</td>
<td>Ataxin-3</td>
<td><img src="image" alt="Ataxin-3 PolyQ Tract" /></td>
<td>45-84</td>
<td>(94,95)</td>
</tr>
<tr>
<td>SCA6</td>
<td>CACNA1A</td>
<td><img src="image" alt="CACNA1A PolyQ Tract" /></td>
<td>21 - 29</td>
<td>(96)</td>
</tr>
<tr>
<td>SCA7</td>
<td>Ataxin-7</td>
<td><img src="image" alt="Ataxin-7 PolyQ Tract" /></td>
<td>37 - 306</td>
<td>(97,98)</td>
</tr>
<tr>
<td>SCA17</td>
<td>TBP</td>
<td><img src="image" alt="TBP PolyQ Tract" /></td>
<td>47 - 63</td>
<td>(99,100)</td>
</tr>
<tr>
<td>DRPLA</td>
<td>Atrophin 1</td>
<td><img src="image" alt="Atrophin 1 PolyQ Tract" /></td>
<td>49 - 88</td>
<td>(101)</td>
</tr>
<tr>
<td>SBMA</td>
<td>Androgen Receptor</td>
<td><img src="image" alt="Androgen Receptor PolyQ Tract" /></td>
<td>40 - 62</td>
<td>(102)</td>
</tr>
</tbody>
</table>

**Abbreviations:** DRPLA-Dentatorubral and Pallidoluysian Atrophy, HD-Huntington’s Disease, SCA-Spinocerebellar Ataxia, SBMA-Spinal and Bulbar Muscular Atrophy, CACNA1A-alpha 1A-voltage-dependent calcium channel, TBP-TATA-box-binding protein

The functions of the polyQ proteins are unrelated, ranging from a calcium channel protein to a deubiquitinating enzyme, and a focus of recent research has been characterizing the functions of these proteins. Targeting the function of the native protein, as opposed to trying to target a conformational change involved in aggregation, has had some success as a
therapeutic approach. In the polyQ disease SBMA, targeting the substrate of the causative protein was demonstrated to inhibit aggregation within a mouse model of disease (103).

1.2.2 Localization of Inclusions in PolyQ Disease

In disease, a redistribution of protein occurs, and for most polyQ diseases the location of aggregates has been demonstrated to be important for toxicity (104-106). Aside from SCA2 and SCA6, nuclear localization and formation of nuclear inclusions is a common characteristic of polyQ disease. Within a recent SCA3 mouse model, the presence of a nuclear export signal alleviated the disease phenotype whilst addition of a nuclear localization signal accelerated aggregation of pathogenic length ataxin-3 (104). This has also been demonstrated in a number of other polyQ proteins, with ataxin-1, huntingtin and DRPLA all showing increased toxicity when targeted to the nucleus (105-108).

Aggregates are predominantly found within the nucleus as NII, but smaller soluble cytoplasmic inclusions are also observed in cell and animal models (109,110). A drosophila model for HD has suggested that soluble cytoplasmic aggregates may also play a role in disease progression, through the trapping of polyQ proteins and blocking of axonal transport (111). The increased toxicity when aggregates are localized to the nucleus as opposed to the cytoplasm may be due to the ameliorating effects of autophagy. Autophagy, or lysosome-mediated protein degradation is a cytoplasmic process which clears polyQ aggregates from the cytoplasm but not the nucleus (112,113). Upregulation of autophagy has been demonstrated to decrease aggregation and toxicity of ataxin-3, huntingtin and the androgen receptor (AR) in a number of cell and animal models (113-115).

During the redistribution of protein which occurs in disease, polyQ proteins may also interact with cellular membranes. The impact of an expanded polyQ tract has been
predominantly studied for huntingtin, and it has been shown that expanded huntingtin inserts into synthetic lipid bilayers more than non-expanded, and also alters the specific phospholipids with which huntingtin interacts (116-118). Similarly to a number of other amyloidogenic proteins, a channel hypothesis has been proposed for HD, whereby expanded polyQ tracts are able to form ion channels in membranes that contribute to cellular disruption and death (119,120). This theory is supported by a study of ataxin-3 which demonstrated that expanded ataxin-3 altered the electrophysiological properties of the cell (121).

1.2.3 Role of Proteolysis in PolyQ Disease

There has been much debate over the role of proteolysis as an initiator of aggregation within polyQ disease. The polyQ tract is the only similarity between the disease proteins, and expansion of the polyQ tract is linked to disease severity both in patients and model systems (104,122). The key involvement of the polyQ tract has led to the ‘toxic fragment hypothesis’ which proposes disease pathogenesis is caused by release of a polyQ containing fragment through cleavage (123-125). In support of this, simple polyQ peptides have been demonstrated to form fibrils \textit{in vitro} and these fibrils are toxic when introduced into mammalian cells (126,127).

The only polyQ disease where there is a general consensus for the involvement of proteolysis is huntingtin, whereby huntingtin exon-1 is released and subsequently aggregates (128-131). However a number of other polyQ proteins have been shown to be highly susceptible to proteolysis \textit{in vitro}. The mechanism of cleavage has been investigated by several groups, and ataxin-3, huntingtin, AR and atrophin-1 are all targets
for caspases (123,132). In addition to this, huntingtin and ataxin-3 are cleaved by calpains *in vitro* (124,133).

### 1.2.4 Identification of Toxic Species in PolyQ Disease

To date, it has been proposed that soluble polyQ oligomers are the species within the aggregation pathway responsible for toxicity, however some uncertainty remains as to the mechanism of toxicity (42,53,79). Within the polyQ diseases, substantial evidence suggests that NII are not linked to toxicity. In fact in a number of both cell and animal models, as well as disease brain slices, the formation of NII has been demonstrated to be inversely correlated with disease (134-138). In a convincing experiment, Arrasate et al. used 4D fluorescent microscopy to follow expression of huntingtin exon-1 in single neurons over time, and found that the bigger the inclusions, the fewer cells died in the primary neuronal cultures (139). In a polyQ model system, a toxic monomeric conformation was shown to be responsible for toxicity, however this has not been observed for polyQ disease causing proteins (140). Most evidence suggests toxicity of the proteins is linked to early aggregates within the pathway, and thus defining the initial steps involved in aggregation may be key to understanding toxicity in the polyQ diseases.

There is evidence that polyQ containing fragments are more toxic than full length proteins with a number of studies demonstrating both faster aggregation rates *in vitro*, and increased toxicity in cell studies overexpressing truncated ataxin-3, the AR and huntingtin (124,141-143). There is general agreement that huntingtin exon-1 is cleaved from full length huntingtin before disease progression, and therefore the toxicity of polyQ fragments must be considered (130,131). Although expression of huntingtin exon-1 in mice leads to aggregation, NII formation and disease, the mice do not show all the characteristics of HD.
Introduction

(131). Aggregation and disease progression have been shown to be slower when full length huntingtin is expressed in mice, however a closer model of HD results, which suggests that full-length huntingtin modulates disease (144). Therefore, there is currently evidence for the role of both full length and truncated polyQ proteins in disease and it appears that the involvement of cleavage may differ between the polyQ proteins.
1.3 Spinocerebellar Ataxia Type 3 (SCA3)

1.3.1 General Characteristics of SCA3

SCA3, also known as Machado-Joseph Disease, is an autosomal dominant disease and is the most common dominantly inherited ataxia (145-147). SCA3 was originally described in families native to the Portuguese Azores islands (148,149). It has since been reported in other populations of non-Portuguese descent, including in America, Europe, Asia, South America, Africa and Australia (150,151).

Although a relatively rare disease, SCA3 is prevalent in specific populations such as in the Portuguese islands of Sao Miguel and Flores, and in an Australian aboriginal community from Groote Eylandt (150,152). There is still debate as to whether SCA3 has evolved from several independent mutations worldwide, or there has been a founder mutation originally occurring in Portugal, which has been spread worldwide via Portuguese explorers (151,153,154).

SCA3 is characterized clinically by progressive cerebellar ataxia, and can be distinguished from other dominant cerebellar ataxias by the presence of ophtalmoplegia, a weakness of the muscles that control eye movement (152). SCA3 shows significant phenotypic variation, with ataxia of gait and stance, cerebellar dysarthria, severe peripheral neuropathy and defective temperature discrimination all being seen (145,155).

The gene responsible for SCA3, \( ATXN3 \) maps to chromosome 14q32.1, with the CAG repeat expansion responsible for disease located within the tenth exon (156,157). Individuals homozygous for expanded repeat alleles are not common, however in a few cases homozygotes have presented with more severe phenotypes than heterozygous
individuals with the same length CAG expansion (158). The smallest observed polyQ repeat length that has resulted in a disease phenotype is 45 glutamines (95).

In SCA3, although the encoded protein ataxin-3 is ubiquitously expressed throughout the brain, only specific subsets of neurons undergo degeneration. Degeneration is observed within the cerebellum and thoracic cord in most patients, with varying degeneration of the striatum, substantia nigra, basis pons, oculomotor nuclei and peripheral nerves (159). Recently, new imaging techniques have shown more widespread degeneration to the cerebellum, thalamus, midbrain, pons, medulla oblongata and spinal cord (160).

1.3.3 Ataxin-3 Structure and Function

Ataxin-3 is one of the smaller polyQ proteins. It consists of a folded N-terminal Josephin domain, and a predominantly unfolded C-terminal tail which contains two ubiquitin interaction motifs (UIM) in addition to the polyglutamine tract (Fig 1.7). There are two major isoforms of ataxin-3 found in disease, the first containing just two UIMs between the Josephin domain and the polyQ tract, and the second containing a third UIM C-terminal to the polyQ tract (161). Both ataxin-3 isoforms show similar polyubiquitin binding properties and aggregation (162-164). The NMR structure of Josephin was solved in 2005 and consists of two subdomains separated by a cleft, and a mixed α/β fold (Fig 1.8) (165,166). NMR experiments suggest that it is a predominantly rigid structure, however it also possesses a flexible helical hairpin consisting of the α2 and α3 helices (Fig 1.8) (166). Ataxin-3 belongs to the family of papain-like cysteine proteases, and has the catalytic triad Cys-14, His-119 and Asp-134 within the Josephin domain (Fig 1.8) (165,166).
Ataxin-3 consists of an N-terminal Josephin domain, two ubiquitin interacting motifs (UIM) and the polyQ tract. A non-expanded tract of 20 glutamine results in an ataxin-3 molecule of 354 amino acids.

Ataxin-3 functions as a de-ubiquitinating enzyme (DUB) as part of the ubiquitin proteosomal system (UPS). The expansion of the polyglutamine tract does not affect either the ability of monomeric ataxin-3 to bind Lys\textsuperscript{48}-linked tetra-ubiquitin, nor change the rate of degradation (162,167). This suggests that it is not a change in function of the monomeric expanded ataxin-3 which is responsible for toxicity.

The UIMs also play an important role in ataxin-3 function, through the binding of ubiquitin chains and also conferring specificity of cleavage (168). If the UIMs are mutated such that they do not bind ubiquitin, this results in the ubiquitination of ataxin-3 being greatly inhibited (162). In addition to binding ubiquitin via the UIMs, ataxin-3 also has two ubiquitin binding sites within the Josephin domain (169). Ataxin-3 causes lower toxicity when the full length protein is present in a native and active conformation. This has been demonstrated within a *Drosophila melanogaster* model system whereby the DUB function of Josephin, in combination with the UIMs, abrogates the level of toxicity in comparison to cleaved C-terminal fragments (170).
Unlike most DUBs which only cleave polyubiquitin chains with a specific linkage, ataxin-3 is currently thought to be unique in its ability to cleave mixed-linkage chains. Winborn et al. showed that ataxin-3 cleaves both Lys$^{48}$ and Lys$^{63}$ linkages even when both are present in mixed linkage chains (168). This suggests that ataxin-3 has more conformational flexibility than the other proteins in this family, which may be linked to its high propensity to aggregate under native conditions.

As part of its normal cellular function, ataxin-3 interacts with a number of proteins including HHR23A and B, and valosin-containing protein (VCP). The proteins HHR23A and B, which are homologs of yeast DNA repair proteins, have been demonstrated to interact with ataxin-3 via a ubiquitin-like domain (171). In addition to this, via modulation of the molecular chaperone VCP (also known as p97), ataxin-3 mediates endoplasmic reticulum associated degradation (172,173). VCP interacts with ataxin-3 via an arginine/lysine rich motif located between the second UIM and polyQ tract of ataxin-3. Neither the interactions with HHR23A or VCP are altered by pathogenic length ataxin-3, however both proteins co-localize with ataxin-3 in NII (171,173).
1.3.2 Molecular Characteristics of SCA3

A number of studies have investigated the role of proteolytic cleavage of ataxin-3 in SCA3. Goti et al. have identified a fragment in some regions of disease brain from a SCA3 patient (141). However in contrast to this, a number of other studies have not observed any fragments in either disease brain or a mouse model (174-176). It should be noted, however, that aggregation of fragment and full length ataxin-3 are not mutually exclusive. Haake et al. have shown that fragments accelerate the rate of full length ataxin-3 aggregation in vitro (124), thus providing a possible mechanism for the involvement of both full length protein and fragments in disease.

Ataxin-3 is highly susceptible to cleavage by both caspases and calpains in vitro, and a number of cleavage sites have been identified in ataxin-3 which release the polyQ containing C-terminal region from the folded Josephin domain (132,174,177). Although studies show cleaved fragments to have increased aggregation kinetics (178), full length ataxin-3 also forms aggregates and NII in cell models (179). The question remains as to whether it is full length or cleaved ataxin-3 which is the initiator of aggregation.

Pathogenic length ataxin-3 forms NII which also contain non-expanded ataxin-3 (4). In addition to this, a number of other cellular proteins are sequestered, including ubiquitin, components of the proteosome, molecular chaperones and transcription factors such as TATA-binding protein (TBP) (81,178,180). Interestingly NII are found in both affected and unaffected brain regions (178,181), and as for a number of other polyQ diseases their formation does not correlate strongly with toxicity (180). However, even if NII are not directly toxic, studies suggest that sequestration of the cellular proteins may contribute to pathogenesis (182).
NII are heavily ubiquitinated, and contain major components of the UPS such as the 20S and 19S proteosomal complexes (81). Within the cell, the UPS functions as a quality control system via the degradation of misfolded or damaged proteins (183). Whilst the UPS has been demonstrated to degrade flanking domains of polyQ proteins, the ability of the UPS to degrade polyQ tracts remains under debate. It has been suggested that polyQ peptides may block the UPS system and thus result in increased ataxin-3 aggregation or alternatively that polyQ peptides may be degraded in the cytoplasm by amino peptidases (81,184-186). Chaperones are another component of SCA3 NII which indicate the involvement of misfolded proteins. The chaperones HSP40 and HSP70 are upregulated upon expression of pathogenic length ataxin-3, thus suggesting that expanded ataxin-3 induces a stress response within the cell (80).

Transcription factors are another element commonly sequestered by SCA3 NII. This corresponds with the widespread transcriptional dysregulation often seen in cellular models of polyQ disease (187). An example of this is creb binding protein (CBP) which is a co-activator for many different transcription factors, and found in NII of ataxin-3 (179). Therefore although formation of NII is considered a protective mechanism for the sequestration of misfolded and fibrillar polyQ proteins, they may play a role in the later stages of toxicity through the sequestration of multiple cellular factors.
1.4 POLYGLUTAMINE AGGREGATION

1.4.1 Introduction to the PolyQ Aggregation Pathway

The aggregation pathways of several polyQ proteins have been well characterized in vitro both structurally and kinetically. Within these proteins, the regions flanking the polyQ tract have been shown to play a role in determining the molecular species formed during aggregation (Fig 1.9). Pathogenic length pure polyQ peptides form two significantly populated species during aggregation, monomer and fibril (Fig 1.9) (188). In the context of the disease protein additional intermediate species such as spherical oligomers and curvilinear protofibrils are also formed (164,189). As discussed previously, polyQ oligomers have also been shown to be toxic (53,79,190), highlighting the importance of flanking domains within the context of disease.

Figure 1.9 Structural Species in Pathological PolyQ Aggregation. The common structural species which are formed during aggregation of pathological length polyQ proteins, with the dotted region indicating species mediated by the flanking domains and the solid region indicating species specifically formed as a result of the pathogenic length polyQ tract.

Use of oligomer-specific antibodies has shown that polyQ proteins and peptides form oligomers with a fold common to well characterized amyloid proteins such as A-β and
α-synuclein (42,164). However the endpoint polyQ disease fibrils, generally characterized by resistance to breakdown by detergents, do not form the typical long straight amyloid fibrils, but instead are more heterogeneous and irregular in appearance (191). This is not the case for polyQ peptides, which form long straight fibrils and show the similar cross-β diffraction pattern typical of amyloid fibrils (31,127). No X-ray diffraction data of polyQ disease proteins exists, which in combination with the morphological data, suggests that the flanking domains prevent formation of typical well-ordered endpoint fibrils.

The polyQ fibrillogenic pathway has been characterized kinetically as nucleation-dependent polymerization (127,192). A kinetic analysis has been completed for polyQ peptides and ataxin-3 and in both cases the size of the critical nucleus is one molecule, thus implying a single rarely populated misfolded monomer is sufficient to initiate aggregation (127,192). For ataxin-3, this partially folded intermediate has been proposed to form off the folding pathway, unlike the general paradigm for misfolding and subsequent aggregation (69).

1.4.2 PolyQ Aggregation and Protein Context

In different contexts the same length polyQ expansion results in varying toxicity and disease pathology, and thus many studies have investigated the role that protein context plays in modulating the extrinsic toxicity of the polyQ sequence. The ability of polyQ flanking sequences to modulate aggregation was clearly demonstrated by Nozaki and colleagues who saw differing aggregation rates when the surrounding 17 amino acids of ataxin-2, ataxin-3, huntingtin and atrophin 1 were added to the same length polyQ tract (193). In addition to this, they found that mutating four hydrophobic residues in the
upstream flanking sequence of ataxin-2 to charged amino acids decreased the aggregation rate, further highlighting the impact of sequence specificity upon aggregation.

Subsequently, a number of studies have investigated the flanking regions to the polyQ tract within huntingtin exon-1. Through the use of peptide studies, the poly proline (polyP) region immediately C-terminal to the polyQ tract in huntingtin (Fig 1.10), has been shown to slow aggregation and decrease the stability of aggregates compared to a polyQ peptide alone (194). This study also found that the protective effects of the polyP region are directional, having no impact on aggregation when N-terminal to the polyQ tract. The mechanism of inhibition was investigated by Darnell et al. (195) who show that the polyP region exists as a PPII helix (proline type II helix), inducing a PPII-like structure in the polyQ tract that opposed the aggregation prone β-sheet conformation formed by this region. In addition to the in vitro study using polyQ peptides, it was shown that the polyP sequence decreases toxicity in a yeast huntingtin exon-1 model system (196). Duennwald and colleagues also investigated the impact of two commonly used tags and show the soluble GST domain prevents toxicity whilst the negatively charged FLAG tag increases toxicity. Therefore the polyQ surrounding sequence within huntingtin exon-1 has a significant modulating effect on huntingtin exon-1 aggregation kinetics.

The sequence specific effects of the flanking domain are unsurprising, given that primary protein sequences vary in their amyloidogenicity (197). Aggregation is known to depend on factors determined by the primary sequence such as hydrophobicity, charge and propensity to form secondary structure (197), with certain amino acids having specific effects upon aggregation propensity. In keeping with its inhibitory effect on huntingtin exon-1 aggregation, proline is considered to be a ‘β-breaker’ as it is structurally
incompatible with β-sheet formation (198). There is also a group of ‘gatekeeper’ residues which are commonly found flanking aggregation prone regions. These include proline, lysine, arginine, glutamate and aspartate, and all have low hydrophobicity, are charged, and show a low propensity to form β-sheet structures (199) as demonstrated in the mutagenesis study on ataxin-2 (193).

In addition to having a direct impact upon the polyQ tract, the flanking sequences allow for varying protein-protein interactions which also modulate aggregation. Ubiquitin is one protein present in the NII of most polyQ diseases, however typically the subset of proteins sequestered within the NII are specific to each disease. The importance of the flanking domain in determining interactions is seen in ataxin-3, where differences between components of NII have been shown between full length versus truncated ataxin-3 (179). In addition to this, huntingtin has over 100 interaction partners within the cell, all of which could potentially affect aggregation (200). Therefore flanking sequences and domains, although not the key cause of toxicity in polyQ proteins, can have a modifying effect on polyQ aggregation through their interaction partners.

1.4.3 The Interplay Between the PolyQ Domain and the Stability of Adjacent Regions

The interplay between the polyQ tract and flanking domains clearly depends upon the characteristics of the domains. The stability of the flanking domain is known to play an important modulatory role in aggregation. In accordance with the general amyloid aggregation theory, an early hypothesis proposed that the expanded polyQ tract destabilized the protein, leading to an increased propensity to aggregate (201). Ataxin-3 is the only disease protein for which kinetic and thermodynamic stability has been determined. Contrary to the initial hypothesis Chow et al. have shown that the addition of a non-
pathogenic length polyQ tract leads to a destabilization of the protein in comparison to the Josephin domain alone, yet expanding the polyQ tract into the pathogenic length has no further effect on stability (69,163).

The complexity of the effect of the polyQ tract on protein stability is demonstrated by a number of polyQ model systems. On the one hand the addition and subsequent expansion of a polyQ tract did not result in a change in the thermodynamic stability of a Staphylococcus aureus protein A domain (202). However this domain is one of the most stable of the model proteins used, so it is unsurprising that the addition of a pathogenic length polyQ tract was not sufficient to disrupt the fold. The CRABP I domain has a similar stability to the Josephin domain of ataxin-3, however the addition of a non-expanded polyQ tract does not impact upon stability, and it is only when the polyQ tract is expanded that there is a corresponding decrease in the stability of the protein. Furthermore, although in a GST-polyQ model the stability was not determined, aggregation could only be induced above 50 °C, suggesting the stability and/or solubility of the GST domain overrides any effects of the polyQ tract at physiological temperatures (203). Therefore these model systems demonstrate the complexity of the relationship between the polyQ tract and the flanking domains and show that whilst in some situations a polyQ tract can decrease the global stability of a protein, it is not necessary in order to induce aggregation.

1.4.4 Multi-domain Misfolding and Aggregation

It is well established that protein domains can fold independently and recently it has been shown that within a protein multiple domains can misfold in separate steps (191,204). Due to the technical difficulties involved in producing these proteins in vitro, the aggregation kinetics and pathways of only three of the nine polyQ proteins have been
characterized. Ataxin-1, ataxin-3 and huntingtin exon-1 all contain a flanking region which is able to aggregate independently of the polyQ tract, albeit with slower kinetics (Fig 1.10). A key tool in the identification of these polyQ-independent aggregation regions has been the polyQ binding peptide QBP1, which inhibits aggregation of the polyQ tract (205). The ability of flanking regions to aggregate independently accounts for the aggregation seen by non-expanded ataxin-1 and ataxin-3 in vitro (164,206-208).

The aggregation pathway of ataxin-3 has been well defined in vitro and it has been shown that the 24 kDa Josephin domain of ataxin-3 aggregates when isolated (206,208). The Bottomley group has shown that full length ataxin-3 aggregates with a two-stage
mechanism (191). The first stage of aggregation involves interactions of the Josephin domain and results in formation of soluble protofibrils. Both pathogenic and non-pathogenic length ataxin-3, in addition to the Josephin domain, form morphologically similar stage 1 protofibrils (191,206,208). The addition of QBPl demonstrates that the polyQ tract is not directly involved in the first step, however the expanded ataxin-3 has the fastest aggregation kinetics, suggesting it still impacts the Josephin domain in some way. The second stage in the polyQ aggregation mechanism is only undertaken by pathogenic length ataxin-3 and mediated by interactions of the polyQ tract, it involves a subsequent remodeling to form an SDS-insoluble fibril (191). Thus the two domains in ataxin-3, both with an intrinsic ability to misfold, aggregate sequentially (Fig 1.11).

Figure 1.11 PolyQ Multi-domain Aggregation Mechanism. A schematic representation of the multi-stage polyQ aggregation mechanism. In this pathway an initial misfolding event occurs in the polyQ flanking folded domain (square), followed by the formation of soluble species mediated by flanking domain interactions. There is then a conformational rearrangement of the polyQ tract which results in formation of end stage SDS-insoluble fibrils.

Although a complete kinetic analysis has not been completed for ataxin-1 due to the limitations caused by its size, the AXH domain of ataxin-1 has similarly been found to aggregate when isolated (207). Within the context of ataxin-1, the AXH domain has a pro-aggregation effect, as replacement with a homologous non-aggregation-prone sequence decreases the formation of nuclear inclusions within a cell model (207). The formation of
detergent insoluble aggregates, commonly attributed to polyQ tract aggregation has not been investigated, nor has the use of QBP1 to ascertain whether the polyQ tract misfolds and aggregates subsequently to the AXH domain.

A recent study by the Wetzel group has shown that huntingtin exon-1 also undergoes multi-domain misfolding during aggregation (189). This study found the 17 amino acids N-terminal of the polyQ tract accelerated aggregation, whether N- or C-terminal to the polyQ tract. This differs from ataxin-1 and ataxin-3 in that the 17 amino acids are not a globular structured domain, but instead exist in a collapsed state, which can be induced into a more aggregation competent extended conformation by the presence of an expanded polyQ tract.

Intriguingly, a kinetic analysis of the huntingtin peptide with the N-terminal sequence resulted in a critical nucleus of minus one. This suggests there is no kinetic barrier to aggregation, and spontaneous aggregation of oligomers occurs in a ‘downhill’ fashion (189). They also found that amorphous aggregates formed in the initial step of aggregation had an exposed polyQ tract that could still bind antibodies. Later aggregates masked the polyQ epitope thus suggesting an additional misfolding event and conformational rearrangement of the polyQ tract to become involved in the fibrils (189). Computer simulations of this region have suggested that the 17 amino acids mirror the conformation of the polyQ tract as opposed to acting independently (209). Thus early studies suggest that huntingtin exon-1 may involve two different misfolding events which each result in formation of different species along the pathway.

With the presence of an amyloidogenic flanking domain being common to polyQ disease proteins, model polyQ protein containing flanking domains with an intrinsic ability
to aggregate provide the most representative insight into the aggregation mechanism of polyQ proteins. One model polyQ system involving the cellular retinoic acid-binding protein I (CRABP I) domain fused to huntingtin exon-1 has proved to have an aggregation pathway which is analogous to ataxin-3. The CRABP I domain aggregates independently and when fused to huntingtin exon-1 the aggregation pathway is characterized by an initial step involving CRABP I, followed by a second step mediated by polyQ tract interactions (204). Using proteolysis, this study also observed the two-stage aggregation pathway in an E. coli cell model, which is the first evidence of the multi-domain mechanism within a cell.

Multi-domain misfolding is not exclusive to polyQ proteins, having also been observed in several other aggregating proteins such as the poly alanine (polyA) protein Poly A Binding Protein Nuclear 1 (PABPN1) and the yeast prion protein Ure2p (210,211). Expansion of the polyA tract occurs at the N-terminal of PABPN1, with an increase from 10 to 17 alanines leading to disease. Similarly to ataxin-1 and ataxin-3, aggregation still occurs in the absence of the homo-sequence domain, resulting in aggregates which are smaller and more numerous than full length PABPN1 (210). A second example, Ure2p, consists of a highly amyloidogenic unstructured N-terminal domain which is responsible for the prion phenotype, and a compactly folded C-terminal catalytic domain (211). Dimerization of the protein occurs through the C-terminal domain, which is also able to aggregate independently, forming high molecular weight insoluble oligomers, that do not progress to fibrillar aggregates (212). Interestingly, the presence of the amyloidogenic N-terminal domain has no impact upon the stability of the full length protein (212).

The second stage of the multi-domain misfolding pathway involves a conformational change within the polyQ tract. A number of studies, both in vitro and in
silico have suggested to date that polyQ peptides have a collapsed structure which is conformationally dynamic (213,214). Using simulations, Leitgeb et al. found that a non-pathogenic length polyQ peptide forms unordered structures in addition to a heterogeneous ensemble of conformations containing various secondary structures (215). Molecular dynamics simulations also suggest that above 30 glutamines, there is a stabilization of structure which may allow time for amyloidogenic interactions to form (216).

The three characterized polyQ proteins all show domain-independent misfolding which translates into multi-step aggregation pathways, whereby the different domains aggregate sequentially. Common to all three cases is the role of a non-polyQ containing domain in forming the initial interactions, thus suggesting a mechanism where these domains initially self-associate, increasing the local concentration of the polyQ tract before the polyQ tract plays a significant role in fibril formation.

The aggregation pathways of polyQ proteins characterized to date have all involved multiple domains. With the variation seen between polyQ diseases, it is unsurprising that their specific flanking domains have the ability to independently misfold and modulate the fibrillogenic mechanism of the full length polyQ protein in vitro. In these studies, the importance of using a number of complementary techniques to follow the multiple aggregation steps has been highlighted. The mechanism via which an expanded polyQ tract results in accelerated aggregation remains unknown, with no clear relationship between the effects of the expanded polyQ tract and global stability of the protein. Further investigation into polyQ protein dynamics and the effects of the expanded polyQ tract upon local stability will therefore be important for determining the early aggregation events.
1.5 AIMS OF THIS STUDY

The flanking domains of polyQ proteins have increasingly been shown to modulate aggregation (193,207). In particular, ataxin-3 has been proposed to form aggregates initially mediated by interactions of the flanking Josephin domain (164,191). Therefore the overall aim of this study is to characterize the aggregation pathway of ataxin-3, with a focus on the role of the Josephin domain.

This study will begin by testing the two-stage aggregation pathway hypothesis that the Josephin domain mediates the initial stages of aggregation. Following on from this the initial steps of aggregation will be investigated through the characterization of the conformational dynamics of the Josephin domain. Lastly, having further defined the aggregation pathway of ataxin-3, the effects of the membrane mimetic sodium dodecyl sulfate (SDS) on aggregation are determined. With very few of the polyglutamine disease proteins able to be produced recombinantly, the further characterization of the aggregation pathway of ataxin-3 may provide insight into the aggregation of the nine disease proteins. Therefore in summary, the aims of this study are to -

1. **Probe the effects of changing the stability of the flanking Josephin domain and thus test the two-stage aggregation hypothesis of ataxin-3**

2. **Determine the conformational dynamics of the Josephin domain of ataxin-3**

3. **Characterize the effects of the membrane mimetic SDS on the ataxin-3 aggregation pathway**
CHAPTER 2

MATERIALS AND METHODS

2.1 DNA CLONING AND MICROBIAL TECHNIQUES

2.1.1 Reagents and Enzymes

The restriction enzyme $DpnI$ and pfu DNA polymerase was purchased from Promega. Agarose I, biotech grade, was purchased from Amresco and SYBR Green was purchased from Applied Biosystems.

2.1.2 2xTY Growth Medium

For 1L medium: To 900 mL of double deionised water (ddH2O) were added:

- Bacto-tryptone 16 g
- Bacto-yeast 10 g
- NaCl 5 g
The medium was stirred until all solutes had dissolved and the volume then adjusted to 1 L. The medium and all other materials for molecular biology and microbial purposes were sterilized by autoclaving at 120 °C for 20 min.

Agar plates were prepared by the addition of 1.5 % (w/v) agar to 2xTY medium prior to autoclaving. 10 cm diameter plates were poured to a depth of approximately 1 cm and set at room temperature. Where appropriate, antibiotics were added once the media had cooled to below 50 °C following sterilization. Antibiotics were added at final concentrations of 0.1 mg/mL for ampicillin and 0.025 mg/mL for kanamycin.

### 2.1.3 Oligonucleotide Design

All oligonucleotides were purchased from Geneworks (Australia). 40 nmol of lyophilized PCR/Sequencing grade oligonucleotides were obtained, reconstituted in sterile ddH2O to a final concentration of 100 μM and stored at -20 °C.

**Oligonucleotides:**

<table>
<thead>
<tr>
<th>Name</th>
<th>Direction</th>
<th>Sequence</th>
<th>Melting Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>R103Gfwd</td>
<td>(5’ – 3’)</td>
<td>TATCAGAGGCTCGGAATCGATCCTATA</td>
<td>58 °C</td>
</tr>
<tr>
<td>R103Grev</td>
<td>(5’ – 3’)</td>
<td>TATAGGATCGATTCCGAGCCTCTGATA</td>
<td>58 °C</td>
</tr>
<tr>
<td>S81Afwd</td>
<td>(5’ – 3’)</td>
<td>ATTCAGGTTATAGCAAATGCCTTGAAA</td>
<td>54 °C</td>
</tr>
<tr>
<td>S81Arev</td>
<td>(5’ – 3’)</td>
<td>TTTC AAGGCATTTGCTATAACCTGAAT</td>
<td>54 °C</td>
</tr>
<tr>
<td>L169Hfwd</td>
<td>(5’ – 3’)</td>
<td>GTTAAGGGTGATCACCAGTTGCGAA</td>
<td>60 °C</td>
</tr>
<tr>
<td>L169Hrev</td>
<td>(5’ – 3’)</td>
<td>TTCGCAATCTGGGTGATCACCCTTAAC</td>
<td>60 °C</td>
</tr>
</tbody>
</table>

### 2.1.4 QuikChange Site-Directed Mutagenesis

A number of variants were created in pExp_lic_his_ataxin3(Q64), ataxin3(Q15) and Josephin using the QuikChange site-directed mutagenesis method. Using the relevant gene as a template, PCR was performed to produce double–stranded circular DNA. The primers
used for PCR each contained a single base substitution as described in 2.1.4. Each PCR mixture contained 100 μM dNTPs, 125 ng of each primer, 25 ng purified vector, 1.5 units pfu DNA polymerase, enzyme buffer and sterile water to a final volume of 25 μL. Five samples were set up to use the gradient function of the Corbett Research Palm Cycler. The following program was run:

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>95 °C</td>
<td>1 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95 °C</td>
<td>1 min</td>
</tr>
<tr>
<td>Annealing</td>
<td>X °C</td>
<td>1 min</td>
</tr>
<tr>
<td>Extension</td>
<td>72 °C</td>
<td>13 min</td>
</tr>
<tr>
<td>Final Extension</td>
<td>72 °C</td>
<td>13 min</td>
</tr>
<tr>
<td>Hold</td>
<td>25 °C</td>
<td>-</td>
</tr>
</tbody>
</table>

X = A gradient ranging from 10 °C to 2 °C below the melting temperature of the primers.

Following amplification, the non-mutated, methylated parental strain DNA was digested by incubation with DpnI for 3 hr at 37 °C. 5 μL of the remaining single-stranded vector was then transformed into 50 μL of NovaBlue™ (Novagen) competent cells and incubated overnight. Resulting colonies were sequenced by PCR amplification using the BigDye Terminator reagent according to the manufacturer’s instructions, and then sequenced at the Micromon sequencing facility at Monash University.

### 2.1.5 Agarose Gel Electrophoresis

For preparation of 1 % agarose gels, 0.25 g of agarose I was added to 25 mL Tris acetate buffer (TAE, 0.04 M Tris acetate, 0.001 M ethylene diaminetetra-acetic acid (EDTA), pH 8.0). The agarose was dissolved by heating, and when it had cooled slightly but was still in liquid form, the solution was poured into a previously prepared gel tray taped firmly at both ends. The tape was removed once the gel had set. A Pharmacia Biotech mini-gel system
was used to run gels, with TAE buffer used as tank buffer. Loading buffer and SYBR green was added to DNA samples and gels were run at a constant voltage of 110 V for approximately 30 min. The DNA sample was visualized using fluorescent light.

### 2.1.6 Plasmid Purification

All expression plasmids used carried a gene conferring ampicillin resistance to the host cells. 5 mL of 2xTY medium containing 0.1 mg/mL ampicillin was inoculated with *E. coli* JM107 cells containing the relevant plasmid and grown overnight at 37 ºC with shaking. Following overnight amplification, the plasmid was extracted and purified from the cells using a Wizard Plus SV Minipreps DNA Purification System (Promega) according to the manufacturer’s instructions.

### 2.1.7 Preparation of Competent Cells

Competent cells were prepared using the rubidium chloride method (217).

**Transformation buffer 1:**

100 mM Rubidium Chloride, 50 mM Manganese Chloride, 10 mM Calcium Chloride, 30 mM Potassium Acetate, 15 % (w/v) Glycerol, pH 5.8

**Transformation buffer 2:**

10 mM Rubidium Chloride, 75 mM Calcium Chloride, 10 mM MOPS, 15 % (w/v) Glycerol, pH 6.8

Solutions were filter sterilized and stored at 4 ºC.

A scraping of JM107 cells from a glycerol stock was added to 5 mL 2xTY media and grown overnight at 37 ºC. 500 μL of the starter culture was used to inoculate 50 mL of
2xYT media in a 250 mL flask, and incubated at 37 °C under shaking conditions until an 
OD₆₀₀ nm of 0.5 was reached. The culture was chilled on ice for 10 min, and then 
centrifuged for 10 min at 3000 rpm. The supernatant was discarded, and the pelleted cells 
were resuspended in 15 mL of transformation buffer 1. The cells were incubated on ice for 
15 – 30 min then centrifuged again. The cell pellet was then resuspended in 4 mL of 
transformation buffer 2 and incubated on ice for 15 min. Cells were aliquoted into 100 μL 
volumes, snap frozen in a dry ice/ethanol bath and immediately transferred to -80 °C for 
storage.

2.1.8 Bacterial Transformations for Expression

2 μL of the desired plasmid was added to 100 μL of previously prepared competent E. coli 
BL21(DE3) cells and left on ice for 30 min. Cells were heat-shocked at 42 °C for 90 sec 
and placed on ice for 2 min to recover, before being plated on 2xTY agar (1.5 % agar) 
containing 0.1 mg/mL of antibiotics as dictated by the transformed plasmid.

2.2 PROTEIN TECHNIQUES

2.2.1 Reagents

Thioflavin T (ThioT), guanidine hydrochloride (GdnHCl), lysozyme, β-
mercaptoethanol (β-me), ethylenediaminetetraacetic acid (EDTA), phenylmethyl sulphonyl 
fluoride (PMSF), ubiquitin, Proteinase K, mass spectrometry grade cesium iodide, ultrapure 
grade ammonium acetate and ultrapure grade ammonium bicarbonate were purchased from 
Sigma Aldrich Co. ThioflavinT (ThioT) and ultrapure grade guanidine (GdnHCl) were
Materials and Methods

prepared analytically by weight and passed through a 0.22 μm filter. The concentration of guanidine was determined from the refractive index of the prepared sample (218).

Isopropyl-β-D-thiogalactopyranoside (IPTG), which was purchased from Astral Scientific, was prepared analytically by weight and passed through a 0.22 μm filter. N,N,N’,N’-Tetramethylethylenediamine (TEMED) and ammonium persulfate (APS) were obtained from ICN and dithiothreitol (DTT) and mass spectrometry grade Trypsin Gold were purchased from Promega. Benzenesulfonyl fluoride hydrochloride (AEBSF), under the trade name Pefabloc, in addition to deoxyribonuclease I (DNaseI) were obtained from Roche. Enzymatic chemiluminesence reagents (ECL) were obtained from Amersham. All quartz cuvettes were purchased from Starna.

2.2.2 Expression of Ataxin-3 Variants

BL21(DE3) E.coli cells containing the appropriate expression vector (pExp_Lic_His) were used to inoculate a starter culture of 2xTY medium containing 0.1 mg/mL ampicillen, and this was grown overnight in a shaking incubator at 37 ºC. The following morning, this culture was added to 2xTY medium at 1 % (v/v) and incubated at 37 ºC and 190 rpm until the OD600 nm was between 0.6-0.8. Subsequently, the incubation temperature was reduced to 28 ºC and protein expression was induced by the addition of 0.5 mM IPTG. The cells were incubated for a further 4 hr and harvested by centrifugation at 5000 rpm for 10 min at 4 ºC. The harvested cells were resuspended in Lysis buffer (2.2.3), snap frozen and stored at –80 ºC.
2.2.3 Protein Purification

**Ataxin-3 and variants**

Solutions:

- **Lysis Buffer:** 50 mM NaH2PO4, 300 mM NaCl, 1 mM β-me, 0.1 % (v/v) Triton X-100, 10 % (v/v) glycerol, pH 8.0
- **Wash Buffer:** Lysis buffer + 20 mM Imidazole
- **Elution Buffer:** Lysis buffer + 250 mM Imidazole

Tris(hydroxymethyl)aminomethane (Tris) buffered saline (TBSG): 100 mM TrisHCl, 80 mM NaCl, 10 % Glycerol, pH 7.4 (at the required temperature)

The cell lysate was incubated with lysozyme (1.4 mM), DNaseI (0.1 mg/mL) and pefabloc (0.5 mg/mL) for 20 min at 4 ºC. The cells were lysed by sonication and after the addition of pefabloc (0.5 mg/mL) the cell lysate was centrifuged for 30 min at 20,000 rpm at 4 ºC. The pellet was subsequently discarded, and the supernatant was filtered through a 0.45 μm syringe filter then added to 10 % (v/v) p-aminobenzamidine agarose and incubated on a rotary orbitor for 1 hr at 4 ºC.

Pefabloc and imidazole were added to the collected flow-through at concentrations of 0.5 mg/mL and 10 mM respectively. The solution was then loaded at a flow rate of 4 mL/min onto a 5 mL HisTrap high performance (HP) column (Amersham), which had been pre-equilibrated with wash buffer. The column was then washed with 20 column volumes (CV) of wash buffer using an Amersham Äkta fast protein liquid chromatography (FPLC) system. The protein was eluted from the HisTrap HP column with a 10 CV gradient from wash buffer to elution buffer. 1 mL fractions were collected and run on SDS-PAGE (2.2.4.1) to determine those containing ataxin-3. The appropriate fractions were pooled and pefabloc added to a final concentration of 0.5 mg/mL.
The chosen fractions were loaded onto a Hiload Superdex™ 200 (16/60) prep grade column which had been pre-equilibrated with TBSG. 1 mL fractions from this column were collected and again analyzed by SDS-PAGE. The purist fractions were pooled and concentrated using an Amicon Ultra (Millipore) centrifuge device. The protein was filtered with a 0.45 µm syringe filter, divided into 50 – 100 µL aliquots and stored at – 80 ºC.

Proteins other than Ataxin-3

The proteins HHR23A, HHR23B and valosin-containing protein(VCP)/p97 were a gift from Wei Wen Dai.

2.2.3.1 Concentration Determination

The concentration of ataxin-3 following purification was determined by absorbance at 280 nm and also using a bicinchoninic acid (BCA) assay (Pierce). Extinction coefficients (ε) for all ataxin-3 variants were calculated using equation 2.1 (Gill and von Hippel, 1989, Pace et al., 1995):

\[ \varepsilon = \text{Number(Tyrosine)} \times \text{Ext(Tyrosine)} + \text{Number(Tryptophan)} \times \text{Ext(Tryptophan)} + \text{Number(Cysteine)} \times \text{Ext(Cysteine)} \]

Equation 2.1

where Ext(Tyrosine) = 1490, Ext(Tryptophan) = 5500, Ext(Cysteine) = 125. BCA assays were completed according to the manufacturer’s guidelines, using bovine serum albumin (BSA) as a protein standard. The values derived from both methods correlated well for all ataxin-3 variants.
2.2.4 Protein Pre-Analysis Techniques

2.2.4.1 SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Stock solutions:

4 x Running gel buffer: 1.5 M TrisHCl, pH 8.8
4 x Stacking gel buffer: 0.5 M TrisHCl, pH 6.8
2 x Loading buffer: 0.125 M TrisHCl, 4 % SDS, 20 % glycerol, 0.2 M DTT, 0.02 % bromophenol blue, pH 6.8
10 % (w/v) sodium dodecyl sulphate (SDS) solution
25 % ammonium persulfate (APS) solution
Tank buffer: 0.025 M TrisHCl, 0.192 M glycine, 0.1 % SDS, pH8.3

Table 2.2: Preparation of two SDS polyacrylamide gels, with concentrations (% v/v) of acrylamide as indicated

<table>
<thead>
<tr>
<th></th>
<th>Running Gel</th>
<th>Stacking Gel</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 % Bis-acrylamide (monomer) (mL)</td>
<td>3.30 4.12 0.44</td>
<td></td>
</tr>
<tr>
<td>4 x Running buffer (mL)</td>
<td>2.50 2.50 0.92</td>
<td></td>
</tr>
<tr>
<td>4 x Stacking buffer (mL)</td>
<td>- - 0.83</td>
<td></td>
</tr>
<tr>
<td>ddH2O (mL)</td>
<td>4.10 3.18 2.36</td>
<td></td>
</tr>
<tr>
<td>25 % w/v APS (μL)</td>
<td>60.0 60.0 20.0</td>
<td></td>
</tr>
<tr>
<td>TEMED (μL)</td>
<td>8.0 8.0 3.0</td>
<td></td>
</tr>
</tbody>
</table>

SDS gels were prepared according to Table 2.2 and run using a Biorad minigel system. Samples were diluted in 2 x loading buffer and boiled at over 90 °C for 5 min before being applied to the stacking gel. A constant current of 45 mA was run for 1 hr. The gels were then stained in Coomassie blue staining solution (0.025 % w/v Coomassie brilliant blue
R250, 40 % v/v methanol, 7 % v/v acetic acid) for 1 hr and destained in 40 % v/v methanol, 7 % v/v acetic acid for 12 hr.

2.2.4.2 Western Blot

Transfer Buffer: 50 mM TrisHCl, 20 mM glycine, 20 % (v/v) methanol, 0.01 % (v/v) SDS.

Tris-Buffered Saline (TBS): 20 mM TrisHCl, 0.2 M NaCl, pH 7.4

Tween Tris-Buffered Saline (TTBS): 0.5% Tween in TBS

Blocking Buffer: 5 % (w/v) low fat milk powder in TBS

Primary antibody:
Mouse anti-histidine tag (Serotec), 1:5000 dilution

Secondary antibody:
Horseradish peroxidase (HRP) - conjugated anti-mouse (Amersham), 1:6000 dilution

Western blots were performed using a Mini Trans-Blot Cell (Biorad). SDS-PAGE gels were run as described in Section 2.2.2.1, using pre-stained molecular weight markers (Biorad). Following the completion of the gel run, the gel was transferred to a cassette against a nitrocellulose membrane pre-soaked in transfer buffer. The gel and membrane, sandwiched between two pieces of filter paper and two sponges, were placed in the cassette which was placed in the Mini Trans-Blot Cell kit. The protein was transferred from the gel to the membrane by running a constant voltage of 250 V for 1 hr. Following the transfer, the nitrocellulose was removed from the cassette and soaked in blocking buffer for one hr.
After blocking, the nitrocellulose was washed in TTBS for 3 x 10 min washes. The membrane was then soaked in a solution of the chosen primary antibody in TBS with 0.05 % (w/v) bovine serum albumin (BSA) for two hr. Following another round of washing in TTBS (3 x 10 min), the secondary antibody, again in TBS with 0.05 % (w/v) BSA, was applied to the membrane for 1 hr. The membrane was then washed again with TTBS (3 x 10 min). The blot was treated with ECL reagents (GE Healthcare) for 1 min, and then exposed onto film in a Hypercassette™ (Amersham) which was then developed in a Fuji Processor FPM-100A.

2.2.4.3 Size Exclusion Chromatography

Prior to experimentation, purified proteins were analyzed by size exclusion chromatography (SEC) to confirm they were in their native monomeric state. A pre-calibrated Amersham Biosciences Superose 12 HR 10/300 column was used in conjunction with an Amersham Äkta FPLC system and the associated Unicorn program. The equilibration and elution buffer was TBSG, and 20 – 200 µg of purified protein was injected onto the column. The flow rate was maintained at the recommended rate and the absorbance monitored at a wavelength of either 214 nm or 280 nm.

2.2.4.4 Ubiquitin Protease Activity Assay

To confirm the protein was active, the ubiquitin protease activity of protein was determined. Ataxin-3 variants and the linear substrate GST-Ub52 were incubated at a ratio of 1:4 (enzyme:substrate) at 37 ºC for 1 hr in a reaction buffer containing 100 mM Tris, 1 mM DTT, pH 8.0. The activity was assessed by coomassie-stained SDS-PAGE (2.2.4.1).
Materials and Methods

The pGEX-Ub52 vector was kindly donated by Roger Everett at the MRC Virology Unit, Glasgow, Scotland. The fusion protein was expressed and purified as described previously (219).

2.2.5 Protein Analysis Techniques

2.2.5.1 Circular Dichroism (CD)

Secondary Structure Analysis

CD spectra were obtained either on a Jasco 810 or a Jasco J-815 spectropolarimeter, both thermostatted by a Jasco peltier. Far-UV CD measurements were obtained using a 0.1 mm path length quartz cuvette with a protein concentration of 20 – 30 μM. Spectra were obtained from 190 – 260 nm using a scanning speed of 50 nm/min, and three scans were accumulated for each sample. The CD measurements were converted to molar ellipticity (θ) using equation 2.2:

\[
\theta = \frac{\theta_{obs}}{10lxC}
\]

\(\theta_{obs}\) is the CD signal in millidegrees, \(l\) is the path length in cm and \(C\) is the protein concentration (M). The units for molar ellipticity are mdeg.cm\(^2\).dmol\(^{-1}\). Secondary structure content of proteins was estimated by deconvolution of the far-UV CD spectra using the CONTINLL algorithm (220,221) available at the online DICHRORWEB facility (222,223).
**Tertiary Structure Analysis**

Near-UV CD spectra were obtained using a 1 cm path length quartz cuvette. Spectra were obtained from 260 – 350 nm at a scanning speed of 50 nm/min, and three accumulations were averaged for each scan.

**Thermal Denaturation**

Thermal denaturation was performed in a 1 cm path length stoppered quartz cuvette. The temperature was increased from 20 ºC to 80 ºC at a rate of 1 ºC/min. Each protein was at a concentration of 1 mg/mL in PBS (137 mM NaCl, 10.1 mM NaH$_2$PO$_4$, 2.68 mM KCl, 1.76 KH$_2$PO$_4$, 10 % Glycerol, pH 7.4), and the change in CD signal at 280 nm was monitored.

2.2.5.2 Equilibrium Denaturation

Equilibrium unfolding was completed using the chemical denaturant GdnHCl. Stock solutions of GdnHCl were prepared in TBSG or 50 mM ammonium acetate, 50 mM ammonium bicarbonate, pH 7.4 buffer. Equilibrium unfolding samples were set up with GdnHCl concentrations ranging from 0 M to 6 M, with a constant protein concentration of 1 µM and the presence of 5 mM EDTA and 15 mM β-me. Samples were left to equilibrate for 3 hr at 4 ºC before CD analysis at 222 nm was completed. Unfolding data was fit to a two-state unfolding model as previously described (224):

\[
I = \frac{(I_f + m_f[U]) + (I_u + m_u[U]) \exp(-\Delta G^o - m[U])/(RT)}{1 + \exp(-\Delta G^o - m[U])/(RT)}
\]

**Equation 2.3**

where $I$ is the measured signal, $[U]$ is the GdnHCl concentration, T is the temperature in degrees Kelvin and R is the gas constant (1.98 cal/mol/K). $I_f$ and $I_u$ are the values of the native and unfolded states respectively, and $m_f$ and $m_u$ are the respective gradients of the
pre- and post- unfolding base lines. $\Delta G^o$ is the change in free energy between the native and unfolded states at 0 M GdnHCl and $m$ is the dependence of the free energy of unfolding on denaturant concentration.

2.2.5.3 Mass spectrometry (MS)

Native Conditions
Spectra under native conditions were obtained using a LCT Premier (Waters) connected to a NanoMate (Advion Biosciences, USA) nanoESI interface and sample delivery system. A capillary voltage of 1.8 kV and a nitrogen nebulizing gas pressure of 0.7 psi were used, together with the following instrument parameters for the LCT premier: sample cone 70 V, ion guide 1 70 V, aperture 1 10 V and ion energy 100 V. Mass accuracy was ensured by calibration with a separate introduction of cesium iodide. Mass spectra were analyzed using MassLynx v4.1 software (Waters UK Ltd.). Spectra were baseline corrected, mildly smoothed using a Sovitzy-Golay algorithm, and then peak centered by peak area.

Denaturing Conditions
Spectra of Josephin in acetonitrile (pH 2.0) were obtained using a Micro Q-Tof (Bruker). Samples were infused using a syringe pump (KD Scientific, USA) at 10 $\mu$L min$^{-1}$ and the following instrument parameters were set: gas temperature 325 K, drying gas 5 L min$^{-1}$, nebulizer gas pressure 20 psi, capillary voltage 5000 V, fragmenter voltage 350 V and skimmer voltage 65 V.

Hydrogen Exchange MS
Josephin in protonated buffer A (50 mM ammonium acetate, 50 mM ammonium bicarbonate, pH 7.4) was diluted 25 fold into deuterated buffer A to initiate exchange. The
sample was incubated in a 20 °C water bath and 10 μL samples were removed over 6 hrs and analyzed using positive nanoESI-MS under native conditions as described above.

In order to eliminate back exchange and determine that Josephin reached the theoretical exchangeable mass during HX, exchange was followed using a closed electrospray source attached to the LCT Premier. In these experiments the sample was infused at 10 μL min\(^{-1}\) using a syringe pump (Harvard Apparatus, MA).

2.2.6 Aggregation Assays
Ataxin-3 variants were incubated at a concentration of 30 μM in TBSG (pH 7.4) buffer containing 5 mM EDTA, 15 mM β-me and 2 mM PMSF. Samples were incubated at 37 °C under non-shaking conditions in air-tight containers to eliminate evaporation. At various time points, samples were removed and assayed as described in the relevant methods section.

2.2.6.1 ThioflavinT Fluorescence

**Discontinuous Analysis**
A 10 μL sample was taken from the aggregation assay (2.2.6) and diluted into 500 μL buffer containing 30 μM thioT. Samples were placed in a quartz fluorescence cuvette (Starna), and using an excitation wavelength of 440 nm, the emission intensity at 480 nm was monitored on a FluoroMax-4 fluorimeter (Horiba Jobin Yvon).

**Continuous Analysis**
ThioT fluorescence was monitored using a Fluostar Optima fluorescence platereader (BMG laboratories). Continuous assays were set up as described in section 2.2.6, but with the
addition of 30 μM thioT. Excitation and emission filters of 450 nm and 490 nm were used respectively. The reactions were performed in black 384 well clear bottom plates (Corning) with the top of the plate sealed to prevent evaporation. All reactions were incubated at 37 °C without shaking, and excitation and emission read from the bottom of the plate. The aggregation kinetics obtained from continuous plate reader assays and thus performed in the presence of thioT, were comparable to kinetic data from discontinuous thioT measurements.

2.2.6.2 Filter Trap Assay

Solutions:

- Stability Buffer: 4 % (w/v) SDS, 100 mM DTT
- TBS: 20 mM TrisHCl, 0.2 M NaCl, pH 7.4
- TTBS: 20 mM TrisHCl, 0.2 M NaCl, 0.5 % (w/v) Tween, pH 7.4
- Blocking Buffer: 3 % (w/v) skim milk powder in TBS
- Primary Antibody: Mouse anti-histidine tag (Serotec), 1:5000 dilution in TBS containing 0.5 % (w/v) BSA
- Secondary Antibody: HRP-conjugated anti-mouse (Amersham), 1:6000 dilution in TBS containing 0.5 % (w/v) BSA

Protein aliquots were diluted into stability buffer at a 1:1 ratio. The samples were heated for 5 min at 100 °C and then diluted with 200 μL of 2 % (w/v) SDS. 70 μL of each sample was filtered through a pre-equilibrated 0.2 μm cellulose acetate membrane (Schleicher and Schuell) using a Bio-Dot SF microfiltration unit (Biorad). The membrane was washed twice with 0.1 % (w/v) SDS. The membrane was soaked in blocking buffer for 1.5 hr, and rinsed briefly with TTBS. The membrane was then soaked for 1 hr in the primary antibody solution. The membrane was washed in TTBS (3 x 10 min) and then incubated for 2 hr in the secondary antibody solution. The membrane was again washed with TTBS (3 x 10
Materials and Methods

min). The blot was analyzed using the ECL reagents as described previously in section 2.2.4.2.

2.2.6.3 Transmission Electron Microscopy

Samples were adsorbed onto a carbon grid and negatively stained with 1 % (w/v) aqueous uranyl acetate (pH 8.0). Samples were analyzed by transmission electron microscopy on a Hitachi H7500 electron microscope using an acceleration voltage of 80 kV.

2.2.7 Phospholipid Overlay Assay

Samples taken from an aggregation assay set up as described in section 2.2.6 were incubated with the phospholipids on the PIP strip membrane (Molecular Probes) overnight at 4 ºC. Each lipid dot contained 100 pmol of phospholipids. The membrane was then blotted using the antibodies described in section 2.2.6.2 using the method described previously (225).
Ataxin-3 aggregation was proposed in 2006 to occur via a multi-stage mechanism in a key study which first detailed investigation the aggregation of a pathogenic length polyQ protein in vitro (191). This study used ataxin-3 variants with polyQ tract lengths of 1, 15 and 64 to demonstrate that all the variants form SDS-soluble fibrils with a curvilinear morphology, termed stage 1 fibrils. However, only pathogenic length ataxin-3(Q64) fibrils formed SDS-insoluble or stage 2 fibrils with a larger and more filamentous morphology (Fig 3.1).

The study proposed that the fibrils formed in stage 1 were mediated by interactions of the Josephin domain based upon two main sets of evidence (191). Firstly, cross-seeding experiments demonstrated that variants could seed the stage 1 aggregation pathways of each other, suggesting the stage 1 aggregates were structurally similar, whilst ataxin-
3(Q64) stage 2 fibrils could only seed the formation of stage 2 fibrils in ataxin-3(Q64). Secondly, the polyQ binding peptide QBP1, which inhibits polyQ mediated aggregation (205), did not affect the formation of stage 1 aggregates yet inhibited the formation of stage 2 aggregates. Therefore based upon these results, in combination with the similar morphology of stage 1 fibrils formed by all variants, Ellisdon et al. concluded that the stage 1 fibrils were mediated by interactions of the Josephin domain (191).

![Ataxin-3 Multi-stage Aggregation Pathway](image)

**Figure 3.1 Ataxin-3 Multi-stage Aggregation Pathway** The two-stage mechanism of ataxin-3 aggregation involves formation of stage 1 fibrils by all variants, and stage 2 fibrils by only pathogenic length ataxin-3. Adapted from (191).

The Josephin domain is approximately half the size of ataxin-3, and has been previously isolated, characterized and demonstrated to be intrinsically fibrillogenic (206,208). Although Ellisdon et al. clearly demonstrated that the polyQ tract was not involved in the formation of stage 1 fibrils, they did not consider whether as an alternative to Josephin interactions mediating stage 1 aggregation, interactions in the remainder of the C-terminal region were involved in this pathway. Therefore this study was designed to test the proposed mechanism whereby Josephin interactions mediate the first stage of ataxin-3 aggregation.
The Impact of Flanking Domain Stability on Aggregation Kinetics of a PolyQ Protein

Helen M. Saunders*, Dimitri Gilis#, Marianne Rooman#, Yves Dehouck#, Amy L. Robertson* & Stephen P. Bottomley*

*Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800, Australia.
#Genomic and Structural Bioinformatics, CP 165/61, Université Libre de Bruxelles, 1050 Brussels, Belgium

Correspondence should be addressed to: Stephen P. Bottomley, Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800, Australia.
Tel: 61-3-9902; Fax. 61-3-9902; Email: steve.bottomley@med.monash.edu.au

Running title: Josephin domain stability modulates ataxin-3 aggregation
Abstract

Spinocerebellar Ataxia Type 3 (SCA3) is one of nine polyglutamine (polyQ) diseases which are all characterized by progressive neuronal dysfunction. A hallmark of these diseases is the presence of neuronal inclusions which contain aggregated polyQ protein, thus suggesting that protein misfolding is a key part of this disease. Ataxin-3, the causative protein of SCA3, contains a globular, highly structured N-terminal domain (the Josephin domain) and a flexible polyQ containing C-terminal tail whose length modulates pathogenicity. The fibrillogenesis pathway of ataxin-3 has been characterized *in vitro* and proposed to have a non-polyQ dependent step mediated by Josephin domain interactions, followed by a polyQ dependent step.

In order to test the involvement of the Josephin domain in ataxin-3 fibrillogenesis, we have created both pathogenic and non-pathogenic length ataxin-3 variants with a stabilized Josephin domain, and have both stabilized and destabilized the isolated Josephin domain. We show that changing the stability of the Josephin domain dramatically impacts upon ataxin-3 fibrillogenesis *in vitro*. These data support the hypothesis that the first stage of ataxin-3 fibrillogenesis observed *in vitro* is caused by interactions involving the non-polyQ containing Josephin domain and that the stability of this domain is linked to the ataxin-3 aggregation propensity.
Introduction

The polyglutamine (polyQ) diseases are a group of nine inherited neurodegenerative diseases, whereby the expansion of a polyglutamine tract above a specific threshold leads to disease. These diseases are characterized by the presence of neuronal nuclear inclusions which contain fibrillar protein aggregates (Diaz-Hernandez et al. 2004). Thus characterization of the pathway from the natively folded protein to the misfolded fibrillar aggregate is important for understanding the role of these misfolded structures in disease initiation, progression and toxicity.

Of the causative polyQ proteins, aside from their expanded polyQ tract, there is no homology between the host proteins. The importance of protein context is increasingly being recognized in the modulation of aggregation for both polyQ proteins and peptides. Studies have reported that flanking domains affect both aggregation rates and the species formed by polyQ proteins (Nozaki et al. 2001; Chen et al. 2002; Duennwald et al. 2006; Thakur et al. 2009).

Recent studies have demonstrated that a number of the polyQ proteins contain domains which have an intrinsic ability to aggregate (Chow et al. 2004b; Masino et al. 2004; de Chiara et al. 2005b; Thakur et al. 2009). Whilst one theory suggests that under the right conditions, the ability to fibrillise is a generic property of all proteins (Dobson 1999; Fändrich et al. 2001), it is significant that a number of polyQ flanking domains are highly aggregation prone. The Josephin domain of ataxin-3, AXH domain of ataxin-1 and first 17 amino acids of huntingtin all have a propensity to aggregate under physiological conditions (Chow et al. 2004b; Masino et al. 2004; de Chiara et al. 2005b; Thakur et al. 2009). In addition to their intrinsic ability to aggregate when isolated, these domains also
play a key role during aggregation when in the context of the full length protein. In a cellular environment, the load of aggregates formed by ataxin-1 is decreased when the AXH domain is substituted with a non-fibrillogenic homologue (de Chiara et al. 2005a; de Chiara et al. 2005b). Similarly, when the Josephin domain of ataxin-3 is substituted with a similarly sized domain, the formation of large aggregates is inhibited (Menon and Pastore 2006). These studies highlight the importance of the flanking domains in the aggregation of polyQ proteins.

An early study suggested that the expansion of the polyQ tract led to protein destabilization and subsequent aggregation (Bevivino and Loll 2001), however a number of recent studies using both disease and model proteins have investigated the effects of expanding the polyQ tract on overall protein stability, with mixed results. The model system CRABP I-huntingtin exon-1 only showed a decrease in global stability with an expanded polyQ tract, whereas no change in stability was observed in a SpA-polyQ model system with various length polyQ tracts (Ignatova et al. 2007; Robertson et al. 2008). In the disease protein ataxin-3, the addition of a polyQ tract destabilized the protein, yet expansion into the pathogenic range had no further effect (Chow et al. 2004a; Chow et al. 2004c). A comprehensive study into the relationship between flanking domain stability and polyQ aggregation has not however been undertaken.

One of the most extensively characterized of the polyQ proteins, ataxin-3 shows a two-stage aggregation mechanism. It has been proposed that the first stage is characterized by interactions involving the flanking Josephin domain, and the second step involves the polyQ tract (Ellisdon et al. 2006). This two-stage mechanism appears to be similar for ataxin-1 and huntingtin, which are the only other disease proteins with well characterized
aggregation mechanisms (de Chiara et al. 2005b; Thakur et al. 2009). Thus, we hypothesize that intramolecular interactions between Josephin domains are the key first step responsible for initiating the ataxin-3 aggregation process, and that the stability of this domain modulates polyQ aggregation kinetics. Therefore if this hypothesis is correct, changing the Josephin domain stability will correspondingly affect ataxin-3 aggregation kinetics. We have rationally designed both stabilizing and destabilizing mutations within the Josephin domain of pathogenic and non-pathogenic length ataxin-3 using the PoPMuSiC algorithm, and show that in support of our hypothesis, the corresponding aggregation kinetics of these mutants are modulated \textit{in vitro}.
Results

Prediction of mutations with changed stability using PoPMuSiC

The PoPMuSiC algorithm was used in order to assist with the identification of mutations that would result in a changed stability of the Josephin domain. We introduced successively, with the PoPMuSiC program, all the single-site mutations in the Josephin domain (PDB code 1YZB (Nicastro et al. 2005)) and computed the resulting change in folding free energy ($\Delta\Delta G_{\text{computed}}$). To reduce the error on the computed $\Delta\Delta G$ values due to dataset limitations, we used two different datasets, i.e. $\mathcal{DB}_1$ and $\mathcal{DB}_2$ (see Experimental Procedures), to evaluate two $\Delta\Delta G$ values: $\Delta\Delta G_{\mathcal{DB}_1}$ and $\Delta\Delta G_{\mathcal{DB}_2}$. The $\Delta\Delta G_{\text{computed}}$'s presented in this paper refer to the average of these two values.

We excluded from our analysis mutations involving prolines, as they are likely to provoke backbone structural rearrangements which are not modeled by the PoPMuSiC program. We also excluded mutations close to the catalytic triad (Cys14, His119, Asn134) to avoid alteration of the biological activity of the protein. We selected the stabilizing mutations that present a $\Delta\Delta G_{\text{computed}}$ value lower than $-1.0$ kcal/mol and the destabilizing mutations with a $\Delta\Delta G_{\text{computed}}$ value larger than $+2.0$ kcal/mol. This resulted in a short list of five potentially stabilizing mutations (Q24G, Q24Y, S81A, R103G, Q129I) and 232 potentially destabilizing mutations.

We chose two of the most stabilizing mutations (Table 1), and four destabilizing mutations which were predicted to have varying stabilities. Of the destabilizing mutations, only the L169H mutation resulted in soluble expression and was used for this study. The lack of expression of the destabilized recombinant proteins suggests protein stability is strongly linked to solubility and/or aggregation.
In the Josephin domain, ataxin-3(Q15) and ataxin-3(Q64) variants, three potentially stabilizing mutations were constructed, R103G, S81A and the double mutant R103G S81A, in addition to the destabilizing mutation L169H (Fig 1). Soluble expression was not achieved for L169H in either ataxin-3(Q15) or ataxin-3(Q64). Purification of Josephin L169H resulted in a yield ten fold lower than that typically achieved for wild type Josephin, however this was sufficient for use throughout this study. For all mutants, ubiquitin-protease activity was confirmed as previously described (Chow et al. 2004b), and the secondary structure and intrinsic fluorescence was determined to be unchanged from wild type (data not shown).

*Stability mutations in Josephin correspond as predicted in silico*

This study then went on to characterize the stability of these mutants in vitro. To determine whether the mutations resulted in a changed thermodynamic stability of Josephin, we used thermal denaturation (Fig 2). Thermal stability was determined by heating the sample 1 ºC per minute, and following the near-UV CD signal at 280 nm. Although a significant loss of tertiary structure occurred upon thermal denaturation, the negative signal observed at high temperatures (data not shown) suggests that some tertiary structure is retained as previously described (Masino et al. 2004). For wild type Josephin the midpoint of the transition was calculated to be 51.3 ± 0.6 ºC (Fig 2A). The two stabilizing mutants Josephin R103G and Josephin S81A both showed an increase in stability with a shift in the transition to midpoints of 55.9 ± 0.2 ºC and 54.1 ± 0.4 ºC respectively, whilst the destabilizing mutant Josephin L169H showed a decrease in stability to a midpoint of 47.5 ºC (Table 2), in agreement with the PoPMuSiC predictions.
Interestingly, the double mutant Josephin R103G S81A showed the same thermal stability as the single mutation Josephin S81A, with a midpoint of 55.4 ± 0.2 ºC. This trend was also observed for ataxin-3(Q15) (Fig 2B) and ataxin-3(Q64) (Fig 2C, Table 2). We have produced a set of stabilizing mutations in all the ataxin-3 variants, in addition to a destabilizing mutation in Josephin, with which to test the hypothesis that Josephin interactions are involved in the two-stage aggregation pathway of ataxin-3.

**Stage 1 aggregation kinetics are modulated by the stability mutations in Josephin**

In the first stage of ataxin-3 aggregation there is the formation of SDS-soluble curvilinear protofibrils, which are detectable using thioT fluorescence. As demonstrated previously, the aggregation kinetics of the ataxin-3 variants differed, with ataxin-3(Q64) showing the fastest kinetics, followed by ataxin-3(Q15) and lastly the Josephin domain (Chow et al. 2004b; Ellisdon et al. 2006). All of the ataxin-3 variants including the stabilized mutants showed a sigmoidal aggregation curve, consisting of a lag phase followed by rapid elongation, which is indicative of nucleated polymerization kinetics (Fig 3).

In Josephin, ataxin-3(Q15) and ataxin-3(Q64), both of the single stabilizing mutations result in a decreased rate of aggregation in comparison to wild type (Fig 3, Table 2). The aggregation rates of the R103G variants are slower than those of the S81A variants and this fits with the slightly greater degree of stabilization of the R103G mutants (Fig 2). Of the stabilizing mutants, the double mutant R103G S81A slowed the aggregation rates of all three proteins to the greatest extent. This was unexpected considering that it had the
same thermal stability as the R103G mutant, and highlights the complex nature of the aggregation mechanism in polyQ proteins.

The destabilized mutant Josephin L169H showed a dramatic change in aggregation kinetics, with the near elimination of the lag phase and a midpoint of aggregation of \(4.0 \pm 0.9\) hours (Fig 3C). Considering that the change in the midpoint of thermal denaturation is similar in magnitude for both the stabilized Josephin R103G and destabilized Josephin L169H, for this set of mutants it appears that destabilizing the Josephin domain has a greater impact on aggregation.

**Stage 2 SDS-insoluble aggregation is correspondingly slowed by the stabilizing mutations**

The SDS-insoluble aggregates formed by pathogenic length ataxin-3 during the second stage of the aggregation process were monitored by a filter trap assay. Through the use of QBP1, the stage 1 protofibrils have been demonstrated to form independently of the polyQ tract, whereas the stage 2 fibrils form through interactions of the polyQ tract (Ellisdon et al. 2006). In this study, the stabilized ataxin-3(Q64) mutants showed slower aggregation kinetics of SDS-insoluble fibrils.

During stage 1 aggregation, there was a 2-4 fold decrease in the aggregation rates of the stabilizing mutants compared to wild type ataxin-3(Q64) (Fig 3C). The differences in aggregation kinetics were much less for SDS-insoluble aggregation. In the stage 2 aggregation pathway, the aggregation midpoints of the stabilized ataxin-3(Q64) mutants only increased by 20-40 % compared to wildtype ataxin-3(Q64). The midpoints of both stage 1 and stage 2 aggregation were very similar for ataxin-3(Q64) R103G S81A. This
implies that stage 2 aggregation is not dependent on the rate of formation of the stage 1 curvilinear aggregates.

*Morphology of end-point fibrils of the ataxin-3(Q64) stabilizing mutants is unchanged*

The morphology of end point ataxin-3(Q64) fibrils was observed using TEM, in order to determine that the mutants formed typical stage 2 fibrils. The aggregates were analyzed after 90 hours of incubation. The diameters of the fibrils ranged from 45 - 90 nm, with average lengths of 2 to 4 μm (Fig 5). Thus all of the stabilized ataxin-3(Q64) mutants showed the formation of the large, filamentous stage 2 aggregates.
Discussion

Confirmation of a multi-stage aggregation pathway for ataxin-3

There is growing evidence that the flanking domain affects the aggregation of a number of polyQ disease proteins, and this study provides further support for the involvement of the Josephin domain in the ataxin-3 aggregation pathway. The role of the flanking domain in aggregation may contribute to the significant variation in the polyQ tract length threshold which leads to disease, from 21 glutamines in SCA6 to 49 glutamines in Dentatorubral-pallidoluysian atrophy (DRPLA), and the differences in disease severity between polyQ proteins with the same length polyQ tract (Komure et al. 1995; Zhuchenko et al. 1997).

The results of this study support our initial hypothesis that Josephin interactions are involved in the first step of the ataxin-3 aggregation pathway, as suggested by the multi-stage aggregation model (Gales et al. 2005; Ellisdon et al. 2006). Specifically, increasing the thermodynamic stability of the Josephin domain slowed down the rate of stage 1 ataxin-3 fibrillogenesis and vice versa (Fig 3). The destabilized Josephin mutant showed the greatest change in aggregation kinetics, with the elimination of the lag phase and an aggregation midpoint of 4 hours (Fig 3C). Thus in the aggregation pathway of the Josephin domain, changing the stability of the protein significantly affects the kinetics of aggregation. Although only stabilizing mutations were studied for ataxin-3(Q15) and ataxin-3(Q64), in both cases the stabilizing mutations slowed the aggregation rate. This suggests that in the context of full length ataxin-3, the Josephin domain is involved in the initial interactions of stage 1 aggregation.
The mechanism via which stage 1 aggregates transition to stage 2 aggregates has not been determined. If this pathway is linear in nature, and stage 1 fibrils must be formed prior to stage 2, then it assumes there is a rearrangement of the stage 1 fibril to form a stage 2 fibril which is mediated by interactions of the polyQ tract. Whether this transition occurs once a critical concentration of stage 1 fibrils is formed, or from a specific point in the formation of the stage 1 fibril is unknown. Alternatively this pathway could be non-linear in nature, and the two types of fibrils formed independently, albeit with different kinetics. Whilst this may be the case, at the endpoint of aggregation there are only stage 2 type fibrils, suggesting that stage 1 fibrils ultimately form or extend into stage 2 aggregates in ataxin-3(Q64).

This study demonstrated that stabilizing mutations located within the Josephin domain of ataxin-3(Q64) had a greater effect on stage 1 aggregation than on stage 2. These data suggest that the two stages of aggregation are not independent and that it is not a linear pathway whereby there must be completion of stage 1 aggregation before stage 2 occurs. Mutations within the Josephin domain had the same impact on aggregation of both isolated Josephin domain and full length ataxin-3, and therefore this study also supports the hypothesis that both proteins form SDS-soluble fibrils via the same pathway.

Although mutants with changed thermodynamic stability were identified in this study, there were also some unexpected results which highlight the complex nature of ataxin-3. Ataxin-3(Q64) had a lower thermal stability than ataxin-3(Q15) (Table 2), which was unexpected considering previous studies have shown that different length polyQ tracts do not change the stability of ataxin-3 (Chow et al. 2004a). This difference may result from the different measures of stability used, as Chow et al. (2004a) followed changes in
intrinsic fluorescence and far-UV CD using acid-induced denaturation. Thermal denaturation is irreversible and ultimately results in aggregation of the sample, however it has been previously demonstrated that aggregation occurs subsequent to the loss of tertiary structure, and thus thermal denaturation can be used as a measure of the stability of the protein (Masino et al. 2004). In addition to this, even though both of the single stability mutations resulted in increased stability in all ataxin-3 variants, the double mutant R103G S81A only showed an increase in stability that was equal to the most stabilized single mutation, R103G. This non-additive effect on stability is somewhat surprising since the residues S81 and R103 are located far away from each other in the native structure of the Josephin domain (Fig 1).

There are a number of factors which impact upon aggregation even in an in vitro system where the interactions which occur within the cell are lacking. The lack of increase in stability when both of the stabilizing mutations were present in ataxin-3, compared to the single-site mutations, demonstrates the complex nature of the aggregation process. This may be because in addition to changing the thermodynamic stability of the proteins, the mutations may also affect the dynamics, solubility and amyloidogenicity of the protein which is impacting upon aggregation. Further study is needed to investigate the role of local dynamics and aggregation, as hydrogen exchange NMR of the mutants would indicate if there was a change in local dynamics resulting from the mutations.

**Aggregation occurs from an intermediate state within the folding pathway**

Two distinct mechanisms have been proposed for the initiation of ataxin-3 aggregation. Firstly, that the aggregation pathway is distinct to that of folding, and
secondly that aggregation is initiated from an intermediate within the folding pathway (Bevivino and Loll 2001; Chow et al. 2004a). The data presented in this study allow us to distinguish between these two mechanisms. If folding and aggregation occurred through different pathways which both initiate from the native conformation of ataxin-3, then the stability mutations would not be expected to affect the aggregation rate. Therefore the change in aggregation kinetics of the ataxin-3 stability mutants suggests these pathways are not independent, and supports the model whereby aggregation occurs off an intermediate in the folding pathway. This is in keeping with the conformational change hypothesis of aggregation (Kelly 1998). There are examples of proteins which undergo both these mechanisms of aggregation, however the conformational change mechanism is the most common for amyloid proteins. As seen in this study, a mutation which decreases the thermodynamic stability of the protein often results in the increased population of an aggregation prone intermediate conformation which exists on the folding pathway, and thus increased aggregation rates (Horwich 2002).

**Determinants of ataxin-3 aggregation in the cell**

The presence of numerous competing interactions and more complex environment of the cell make the aggregation rates difficult to compare with *in vitro* studies. Warrick *et al.* suggest that within a *Drosophila* model of ataxin-3, the full length ataxin-3 is less toxic than a truncated C-terminal region which does not contain the Josephin domain (ataxin-3(tr)) (Warrick et al. 2005). The two-stage mechanism identified *in vitro* suggests that the initial stages of aggregation are characterized by interactions of the Josephin domain, which conflicts with the amelioration of toxicity when the Josephin domain is present in
Drosophila. This difference may be attributed to the Josephin interaction partners present within the cell, some of which have been demonstrated to also modulate aggregation in vitro (Boeddrich et al. 2006; Robertson et al. 2010).

In a screen for interactors in Drosophila, ataxin-3 bound to chaperones such as the small heat shock protein αB-crystallin (Bilen and Bonini 2007). This chaperone has subsequently been demonstrated in vitro to transiently interact with the Josephin domain to inhibit Josephin-mediated aggregation, but has no effect upon aggregation caused through polyQ interactions (Robertson et al. 2010). This highlights how the impact of external interactions in vivo can dramatically affect the aggregation rates, and also the complementary nature of in vitro characterization to determine the effect of extrinsic modifiers on ataxin-3 aggregation.

The involvement of cleavage within the polyQ diseases raises the relevance of studying the impact of the Josephin domain in ataxin-3 aggregation. As opposed to various other polyQ diseases, in SCA3, there is conflicting evidence as to the presence of ataxin-3(tr) fragments, with only one of several studies reporting the presence of fragments (Paulson et al. 1997; Cemal et al. 2002; Berke et al. 2004; Goti et al. 2004). Even within the study which identified ataxin-3(tr) within a patient brain, there was an equal proportion of full length ataxin-3 (Goti et al. 2004). Thus the question still remains as to whether full length or cleaved ataxin-3 is the initiator of aggregation.

Although interactions in the Josephin domain mediate ataxin-3 aggregation in vitro, the binding of interaction partners to Josephin in vivo may cause the slowed aggregation of full length ataxin-3 in comparison to truncated ataxin-3 which lacks the Josephin domain. Therefore, the characterization of the in vitro aggregation pathway provides a means to
investigate the effects of cellular factors or therapeutics within a controlled environment. With the flanking domains directly involved in aggregation, targeting these domains as a means to slow the aggregation pathway provides a potential therapeutic strategy.
Experimental Procedures

Materials – Phenylmethylsulfonyl fluoride, β-mercaptoethanol and thioflavin T (ThioT) were all obtained from Sigma.

PoPMuSiC Algorithm – The PoPMuSiC program aims to predict the folding free energy changes (ΔΔG=ΔGmutant−ΔGwild-type) that results from single-site mutations in a protein (Gilis and Rooman 2000; Kwasigroch et al. 2002). All the possible point mutations are successively introduced in silico in the protein and their ΔΔG’s are computed by means of statistical potentials derived from datasets of known protein structures. Two datasets are used here, referred to as \(DB_1\) and \(DB_2\): the first contains 141 well resolved X-ray structures (resolution ≤ 2.5 Å) of protein chains with less than 25 % pairwise sequence identity (Wintjens et al. 1996), and the other 735 high-resolution X-ray structures (resolution ≤ 2.0 Å) of protein chains with no more than 20 % pairwise sequence identity (Hobohm et al. 1992). Two types of potentials are considered. \(ΔG_{\text{torsion}}\) is a torsion potential that describes local interactions along the sequence, and is obtained from the frequencies of association between residues or residue pairs and backbone torsion angle domains. \(ΔG_{\text{distance}}\) is a distance potential computed from the propensities of residue pairs to be separated by a spatial distance between average side chain centroids \(C_c\) (Kocher et al. 1994). These potentials take implicitly the solvent into account as they are derived from solution structures. The folding free energy changes ΔΔG’s are evaluated by linear combinations of torsion and distance potentials:

\[
ΔΔG_{\text{computed}} = \alpha \ ΔΔG_{\text{torsion}} + \beta \ ΔΔG_{\text{distance}} + \gamma
\]

where \(\alpha\), \(\beta\) and \(\gamma\) are weighting factors depending on the solvent accessibility of the mutated residue (Gilis and Rooman 2000).
The PoPMuSiC program requires as input the structure of the target protein in PDB format (Berman et al. 2000) and uses a simplified representation of protein structures consisting of the backbone atoms and the pseudo-atom C. The backbone structure is assumed to be unchanged upon mutation.

*Mutagenesis* – Using the QuickChange mutagenesis method, the following mutations were introduced into Josephin, ataxin-3(Q15) and ataxin-3(Q64): L169H, S81A, R103G.

*Expression and Purification of Ataxin-3 variants* – All ataxin-3 variants were expressed and purified as previously described (Chow et al. 2004c), and the proteins were stored at -80 °C. Following purification the deubiquitinating activity of the proteins was measured (Chow et al. 2004b) and before use all were analyzed using gel filtration to ensure that no multimeric species were present.

*Thermal Denaturation* – Thermal unfolding was monitored on a thermostatted Jasco-810 spectropolarimeter and performed by increasing the temperature by 1 °C per min from 20 °C to 80 °C. Changes in the near-UV CD signal were monitored at 280 nm, using a stoppered quartz cuvette with a path length of 1 cm. Protein samples were measured at a concentration of 1 mg/ml.

*Fibrillogenesis Time Course Assays* – Ataxin-3 variants (30 µM) were incubated in TBSG (80 mM Tris, 100 mM NaCl, 10 % (v/v) glycerol, pH 7.4) containing 5 mM EDTA, 15 mM β-mercaptoethanol and 2 mM phenylmethylsulfonyl fluoride. Samples were incubated without shaking at 37 °C in air-tight containers to prevent evaporation.

*ThioT Fluorescence* – ThioT fluorescence measurements were recorded on a BMG Fluorostar plate reader, using the buffering conditions described previously with the addition of 30 µM thioT. Excitation and emission wavelengths of 430 nm and 480 nm with
a cut off filter of 455 nm were used, and both excitation and emission were read from the underneath of a black clear bottom plate. All reactions were completed at 37 °C with no shaking.

Membrane Filter Trap Assay – Aliquots containing 7.4 µg of protein were taken from the fibrillogenesis reaction, diluted 1:1 with a 4 % (w/v) SDS / 100 mM DTT solution and then boiled for 5 minutes at 100 ºC. 200 µl of 2 % (w/v) SDS was then added to each of the samples, and 2.5 µg of protein was filtered through a 0.2 µm cellulose acetate membrane (Schleicher and Schuell) using a Bio-Rad Bio-Dot SF microfiltration unit. The membrane was then washed twice by filtering 200 µl of 0.1 % (w/v) SDS, and blotted with a hexahistidine (His6) antibody (Serotec). Densitometry was completed using the program ImageQuant.

Transmission Electron Microscopy (TEM) – TEM images were obtained using a Hitachi TEM with an accelerating voltage of 80 kV. Samples were adsorbed onto a carbon-coated grid, and stained with 1 % (w/v) uranyl acetate.

Acknowledgements

SPB is a Senior Research Fellow of the NHMRC and the work was supported by both the ARC and NHMRC of Australia.
References


Figure Legends

**Figure 1** Predicted stability mutations in the Josephin domain. The two stabilizing mutations R103G and S81A and the destabilizing mutation L169H are highlighted on the 1YZB Josephin structure.

**Figure 2** Thermal denaturation of Josephin stability mutants. Thermal denaturation was followed at 280 nm using near-UV CD. The temperature was increased by 1 °C min⁻¹ in a stoppered cuvette with a sample concentration of 1 mg ml⁻¹. Representative traces are shown for (A) Josephin (B) ataxin-3(Q15) and (C) ataxin-3(Q64) with wild type (black), R103G (blue), S81A (green), R103G S81A (purple) and L169H (red).

**Figure 3** Ataxin-3 stage 1 aggregation kinetics are changed by the Josephin mutants. Stage 1 aggregation was monitored using thioT fluorescence at 30 μM and pH 7.4 with a fluorescence plate reader. Wild type (black), R103G (blue), S81A (green), R103G S81A (purple) and L169H (red) are shown in (A) ataxin-3(Q64), (B) ataxin-3(Q15) and (C) Josephin.

**Figure 4** SDS-insoluble aggregation is slowed by a stabilized Josephin domain to a lesser extent. Formation of SDS-insoluble aggregates was analyzed over time, taking aliquots from an aggregation assay at 30 μM and pH 7.4. (A) A representative filter trap membrane shows the formation of SDS-insoluble aggregates formed from ataxin-3(Q64) and the stabilizing ataxin-3(Q64) mutants. (B) Densitometry is shown for ataxin-3(Q64) (black), ataxin-3(Q64) R103G (blue), ataxin-3(Q64) S81A (green) and ataxin-3(Q64) R103G S81A (purple). Analysis of the membrane by densitometry was completed on 3 independent experiments and fit to a sigmoidal curve.

**Figure 5** Morphology of endpoint aggregates is unchanged. Transmission Electron Microscopy of (A) ataxin-3(Q64), (B) ataxin-3(Q64) R103G, (C) ataxin-3(Q64) S81A and (D) ataxin-3(Q64) R103G S81A. After 90 hr of aggregation, samples were negatively stained using 1 % uranyl acetate. Scale bar represents 200 nm.
Table 1  PoPMuSiC Predicted Changes in Josephin Stability

<table>
<thead>
<tr>
<th>Mutation Type</th>
<th>Mutation</th>
<th>(\Delta\Delta G_{\text{computed}}) (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stabilizing</td>
<td>R103G</td>
<td>-1.37</td>
</tr>
<tr>
<td>Stabilizing</td>
<td>S81A</td>
<td>-1.10</td>
</tr>
<tr>
<td>Destabilizing</td>
<td>L169H</td>
<td>3.00</td>
</tr>
</tbody>
</table>

Table 2  Thermal Denaturation and Aggregation Midpoints

<table>
<thead>
<tr>
<th></th>
<th>Thermal Denaturation Midpoint (^\circ\text{C})</th>
<th>ThioT Midpoint (hrs)</th>
<th>SDS-Stability Midpoint (hrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Josephin</td>
<td>51.3 ± 0.6</td>
<td>79.4 ± 6.7</td>
<td></td>
</tr>
<tr>
<td>Josephin R103G</td>
<td>55.9 ± 0.2</td>
<td>147.5 ± 7.3</td>
<td></td>
</tr>
<tr>
<td>Josephin S81A</td>
<td>54.1 ± 0.4</td>
<td>115.2 ± 6.9</td>
<td></td>
</tr>
<tr>
<td>Josephin R103G S81A</td>
<td>55.4 ± 0.2</td>
<td>&gt;150 hr</td>
<td></td>
</tr>
<tr>
<td>Josephin L169H</td>
<td>47.5 ± 0.1</td>
<td>4.0 ± 0.9</td>
<td></td>
</tr>
<tr>
<td>Ataxin-3(Q15)</td>
<td>49.2 ± 0.1</td>
<td>24.0 ± 1.3</td>
<td></td>
</tr>
<tr>
<td>Ataxin-3(Q15) R103G</td>
<td>52.9 ± 0.8</td>
<td>48.9 ± 1.8</td>
<td></td>
</tr>
<tr>
<td>Ataxin-3(Q15) S81A</td>
<td>50.6 ± 0.9</td>
<td>32.2 ± 1.9</td>
<td></td>
</tr>
<tr>
<td>Ataxin-3(Q15) R103G S81A</td>
<td>53.4 ± 0.4</td>
<td>50.0 ± 2.8</td>
<td></td>
</tr>
<tr>
<td>Ataxin-3(Q64)</td>
<td>51.0 ± 0.3</td>
<td>11.7 ± 1.2</td>
<td>35.0 ± 2.7</td>
</tr>
<tr>
<td>Ataxin-3(Q64) R103G</td>
<td>52.1 ± 0.6</td>
<td>25.3 ± 2.1</td>
<td>42.9 ± 2.6</td>
</tr>
<tr>
<td>Ataxin-3(Q64) S81A</td>
<td>50.8 ± 0.2</td>
<td>26.3 ± 4.9</td>
<td>41.0 ± 7.1</td>
</tr>
<tr>
<td>Ataxin-3(Q64) R103G S81A</td>
<td>52.4 ± 0.3</td>
<td>40.6 ± 2.2</td>
<td>48.7 ± 1.3</td>
</tr>
</tbody>
</table>
Figure 1
Figure 2

A

Josephin Mutants

B

Ataxin-3(Q15) Mutants

C

Ataxin-3(Q64) Mutants
Figure 3

A

Ataxin-3(Q64) Mutants

B

Ataxin-3(Q15) Mutants

C

Josephin Mutants
Figure 4

A

B
Figure 5
CONFORMATIONAL DYNAMICS OF THE JOSEPHIN DOMAIN OF ATAXIN-3

4.1 INTRODUCTION

4.1.1 Protein Aggregation and Dynamics

The conformational change hypothesis of aggregation proposes that folded proteins undergo an initial conformational change to adopt a partially unfolded amyloidogenic intermediate (68,226). The propensity to form such an intermediate is linked to the propensity to aggregate (68), and this has been demonstrated for a number of amyloidogenic proteins, including β2-microglobulin, α-synuclein and lysozyme (77,227,228). For these proteins, mutations or solution conditions that promote aggregation result in an increased population of partially unfolded intermediates (77,227-229).

The proportion of the amyloidogenic intermediate in solution can be affected by both its thermodynamic equilibrium with the native conformation as well as the rate of
conversion and conformational dynamics (74). In several amyloid diseases, a destabilizing mutation increases the population of a partially unfolded amyloidogenic intermediate (75,230,231). However in other diseases, mutation leads to an unchanged thermodynamic stability and it is the protein dynamics which are affected. This is observed for both lysozyme and transthyretin, whereby some aggregation prone mutants have increased rates of conformational fluctuations (75-77).

Both global and local dynamic fluctuations can enable a folded protein to momentarily visit partially unfolded conformations, however their influence on aggregation varies for different proteins. Calloni et al. suggest that for proteins which form an inherently fluctuating intermediate with non-persistent structure, the region with the highest intrinsic aggregation propensity, as opposed to the most dynamic region, will drive aggregation (229). Conversely, in another study local differences in dynamics between human and mouse α-synuclein were used to explain their altered aggregation propensity (232). Therefore to fully characterize the aggregation pathway of a folded protein, it is important to consider the conformational dynamics.

4.1.2 Ataxin-3 Aggregation and Dynamics

Flanking domains within the polyQ disease proteins have increasingly been shown to modulate polyQ aggregation (189,191,196,207). For example, several studies have shown that both the immediate polyQ flanking sequence in addition to flanking domains change aggregation both in vitro and in cell models (189,193,194,196). The dynamics of the flanking domains of polyQ proteins such as the Josephin and AXH domains have not been well characterized. When isolated, the folded Josephin domain is intrinsically
fibrillogenic and forms SDS-soluble curvilinear fibrillar species (233,234). In the context of ataxin-3, it is involved in the multi-stage pathway of ataxin-3 aggregation, with the first stage of aggregation being mediated by Josephin domain interactions (Ch 3)(191). Similarly, the AXH domain has also been demonstrated to be intrinsically fibrillogenic when isolated, and shown to increase aggregation in cells in the context of ataxin-1 (207,235). Therefore determining the conformational changes and dynamics of fibrillogenic flanking domains, such as the AXH domain and Josephin, may provide insights into the initial events of aggregation of their respective proteins.

In SCA3, as for many other polyQ diseases, only an expanded polyQ tract with greater than 45 glutamines leads to disease (236). As many amyloid forming proteins contain a mutation which decreases thermodynamic stability, the effect of an expanded polyQ tract on protein stability has been investigated in vitro. Mixed results have been obtained with model polyQ systems, with different studies showing either no change in stability or destabilization of the model protein upon polyQ expansion (237,238). Chow et al. showed there was no corresponding decrease in kinetic or thermodynamic stability upon expansion of the polyQ tract into the pathological range (69). In contrast to this, there is a difference in aggregation kinetics of ataxin-3 in vitro, which suggests that other properties of the protein are affected by the presence of the expanded polyQ tract.

In ataxin-3 the Josephin domain plays a key role in both aggregation and cellular function (165,168,233,234). Ataxin-3 is a de-ubiquitinating protease (DUB), and the Josephin domain contains a papain-like fold (Fig 4.2) (165,166). It contains two ubiquitin binding regions, one of which is within the helical hairpin region, with the second adjacent to it, and the catalytic triad Cys-14, His-119 and Asp-134 (165,166,169). Ataxin-3 is an
unusual DUB in that it has the unique ability to cleave multi-linked ubiquitin chains, as opposed to most DUBs which cleave ubiquitin chains of a single linkage (168). This suggests that Josephin may have considerable conformational flexibility and undergo structural changes to function in this manner. Therefore determining the conformation dynamics of the Josephin domain may provide insight into both its activity and aggregation propensity.

4.1.3 Theory of Hydrogen Exchange Mass Spectrometry

Hydrogen Exchange (HX) is a technique used to characterize the structure and dynamics of proteins, as within a folded protein the rate of exchange depends upon factors such as whether a proton is involved in H-bonded secondary structure, or is buried within the core of the protein (239). HX is also used to study the conformational dynamics of proteins as conformational fluctuations allow exchange to occur during the breaking and reforming of H-bonds. Proteins contain a number of protons which exchange on different timescales, and it is the backbone amide protons which exchange over a timeframe amenable to analysis (Fig 4.1).

Although commonly detected by NMR, mass spectrometry (MS) can also be used to analyze HX due to the increase in mass which occurs as amide protons exchange to deuterons. Although MS does not give residue specific information, the advantages of using MS over NMR are that a larger mass range can be analyzed, co-existing populations can be detected and different mechanisms of exchange distinguished (240). The kinetic analysis discussed below has been described previously by a number of groups (239-242).
The intrinsic exchange rate ($k_{\text{int}}$) of an amide hydrogen is the rate at which it exchanges if freely exposed to solvent. This rate is sensitive to pH (242), as it is catalyzed by both acids and bases as described:

$$k_{\text{int}} = k_w + k_{\text{acid}} [H^+] + k_{\text{base}} [OH^-]$$

Equation 4.1

where $k_w$ is the water catalyzed, $k_{\text{acid}}$ the acid catalyzed and $k_{\text{base}}$ the base catalyzed second-order rate constants. In addition to being pH-dependent, $k_{\text{int}}$ is also dependent upon temperature, with rates changing by a factor of 10 per 22 °C (243).

The intrinsic exchange rate of an amide proton is sequence specific as bulky amino acid side chain can sterically block exchange, or conversely neighbouring polar amino acids can increase exchange (242). Bai and colleagues have quantified the effects of neighbouring amino acids on exchange for all amino acids and thus using these values, $k_{\text{int}}$ can be calculated both individually for each amino acid in a sequence, and also for the protein as a whole (244).

In folded proteins, exchange involves a two-state transition from a closed state (C) not accessible to exchange, to an open accessible state (O) which is described as:

---

**Figure 4.1 Timescale of exchange of protons.** Amide protons (pink) exchange within a timescale which is amenable to analysis. The side chain protons (green) and the carbonyl protons (blue) exchange at rates which are too fast and too slow respectively for analysis.
Conformational Dynamics of Josephin

\[ C^H \xrightarrow{k_{op}} C^D \xrightarrow{k_{cl}} O^D \xrightarrow{k_{op}} O^H \]

Equation 4.2

where \( C^H \) and \( O^H \) represent amide protons which are in closed or open conformations either locally or globally, \( C^D \) and \( O^D \) are the exchanged or deuterated amides in a closed or open conformation and \( k_{op} \) and \( k_{cl} \) are the rates of opening and closing respectively. In a folded protein, the mechanism of exchange is dependent upon the rate of unfolding in addition to the intrinsic rate of exchange. Therefore according to Eq. 4.2, the exchange rate of a proton \((k_{ex})\) becomes:

\[
k_{ex} = \frac{k_{op}k_{int}}{k_{op} + k_{cl} + k_{int}}
\]

Equation 4.3

Under native conditions, generally \( k_{cl} \gg k_{op} \), and therefore Eq. 4.3 can be simplified to:

\[
k_{ex} = \frac{k_{op}k_{int}}{k_{cl} + k_{int}}
\]

Equation 4.4

There are two extreme limits of Eq. 4.4, which are termed EX1 and EX2. EX2 is most commonly observed under physiological conditions and occurs when \( k_{cl} \gg k_{int} \), such that:

\[
k_{ex} = \frac{k_{op}k_{int}}{k_{cl}} = K_o k_{int} \quad \text{where} \quad K_o = \frac{k_{op}}{k_{cl}}
\]

Equation 4.5

Thus under an EX2 mechanism a protein will undergo random rapid visits to a conformation where exchange is possible, however the probability of exchange each time is low. As \( k_{ex} \) is dependent upon \( K_o \), the equilibrium constant between a closed and open conformation, the apparent free energy of exchange \( (\Delta G_{ex}^{app}) \) can be determined where \( R \) is the gas constant and \( T \) the absolute temperature:
\[
\Delta G_{ex}^{app} = - RT \ln \left[ \frac{k_{obs}}{k_{int}} \right]
\]

**Equation 4.6**

EX1 is the other limit that can be observed, and occurs where \( k_{int} \gg k_{cl} \). Less common under native conditions, this corresponds to a situation where proteins rarely visit a conformation in which exchange can occur, however upon visiting a conformation amenable to exchange then complete exchange will occur. Therefore under this regime Eq. 4.4 is simplified to:

\[ k_{ex} = k_{op} \]

**Equation 4.7**

The EX1 and EX2 mechanisms of exchange are easily distinguished when HX is analyzed using MS (240). EX1 results in a significant increase in mass to the fully deuterated mass and then an increase in intensity of this mass over time as more molecules visit an open conformation. In comparison, EX2 exchange results in a gradual increase in mass over time, as the protein frequently visits an open conformation yet only a small amount of exchange occurs each time.

In this study a MS approach is used to investigate the conformation of Josephin under both native and denaturing conditions, and show that multiple conformations of Josephin are populated in solution under native conditions. Using HX these conformations are demonstrated to be in equilibrium, thus highlighting the highly dynamic nature of Josephin.
4.2 Results

4.2.1 Josephin MS under Native Conditions

Electrospray mass spectroscopy (ESI-MS) is a technique which under native conditions can be used to determine the number of protein conformations which exist in solution, as the more compact the structure, the fewer charges that can be accommodated on its surface (245). Initially, the characteristic mass spectrum displayed by the Josephin domain of ataxin-3 (Fig 4.2) under native conditions was investigated.

The three volatile buffers predominantly used in combination for native state MS are ammonium acetate, ammonium bicarbonate and ammonium formate. When buffered in ammonium formate, Josephin immediately precipitated, and thus this buffer was not used. The use of an ammonium acetate buffer resulted in a stronger signal than ammonium bicarbonate, however this buffer did not buffer as well at pH 7.4 over time. Thus a buffer containing both ammonium acetate and ammonium bicarbonate was chosen for all native MS experiments as it resulted in clean spectra with sufficient intensity (Fig 4.3).
Conformational Dynamics of Josephin

Figure 4.3 Josephin mass spectra under native conditions. (A) The mass spectra of Josephin under native conditions obtained using nanoESI-MS shows multiple charge state envelopes (CSE) centered about n = +8, +13 and +18. The spectra was obtained using a low ionic strength buffer. (B) The mass spectra of Josephin in a high ionic strength buffer shows a loss of signal due to noise.

Both a low and high ionic strength buffer were tested in order to determine which resulted in the cleanest spectra with the highest intensity (Fig 4.3). A 50 mM ammonium acetate, 50 mM ammonium bicarbonate pH 7.4 buffer resulted in a significantly greater signal intensity than a 90 mM ammonium acetate, 90 mM ammonium bicarbonate pH 7.4 buffer which had a great deal of noise, and thus the former buffer was used for all native condition MS experiments.

Unlike other spectroscopic techniques such as CD and NMR where the signal represents an average of the conformations which exist in solution, if native and non-native conformations co-exist in solution, then a bimodal distribution is observed using ESI-MS (246). Therefore under native conditions, the number of charge state envelopes in ESI-MS corresponds to the number of protein conformations in solution.

Despite the well defined NMR structure of the Josephin domain (247), the mass spectra shows a larger charge distribution (n = +22 to +7) than would be expected for a folded protein, in addition to multiple charge state envelopes (Fig 4.3). The first
distribution centres on the charge state \( n = +8 \), and corresponds to approximately 60 % of the molecules, as calculated from summing the area under the peaks in each distribution. With the ability to bind the least amount of charge, this population is the most folded in solution (246). The second and third populations centre around \( n = +13 \) and \( n = +18 \), together corresponding to approximately 40 % of molecules in solution. These two charge state envelopes correspond to molecules which exist in solution in a partially unfolded conformation.

4.2.2 Josephin Exists in Multiple Conformations in Solution

4.2.2.1 Multiple charge state envelopes are not linked to parameters on the MS

A series of control experiments were performed to ensure these conformations were not merely an artifact of the mass spectrometer. The cone voltage and ionization source temperature were systematically altered, as was the sample temperature. Increasing the cone voltage can shift unfolded conformations to a higher m/z and thus lower charge state, due to a ‘charge stripping’ effect (248,249). The cone voltage was increased from 50 V to 110 V, however changing it had no effect on the ratio of molecules in each population (Fig 4.4A). Increasing the ionization source temperature can unfold proteins at high temperatures during the ionization process (248). The effect of having a temperature of 50 °C to 100 °C was investigated, and similarly did not change the proportion of each charge state envelope (Fig 4.4B). Therefore a standard cone voltage of 70 V and ionization source temperature of 50 °C was used for all experiments.

The effect of changing the sample temperature on the multiple charge state envelopes was also determined. A sample temperature of between 4 °C and 42 °C did not
affect the Josephin mass spectra (Fig 4.4C). In addition, using two different ESI-MS instruments (Waters and Bruker), and two expression constructs of Josephin all resulted in Josephin mass spectra with multiple charge state envelopes (Tomé, Radford and Macedo-Ribeiro, personal communication). This suggests that the presence of multiple charge state envelopes is not an artifact of the ionization process of ESI-MS.

![Figure 4.4 (A) Effects of changing the cone voltage on the Josephin mass spectra.](image)

**Figure 4.4 (A) Effects of changing the cone voltage on the Josephin mass spectra.** Changing the cone voltage from 50 – 110 V does not significantly affect the proportions of the multiple charge state envelopes.
Figure 4.4 (B) Effects of changing the ionization source temperature on the Josephin mass spectra. Changing the ionization source temperature from 50 – 100 °C does not significantly affect the proportions of the multiple charge state envelopes.

Figure 4.4 (C) Effects of changing the sample temperature on the Josephin mass spectra. Changing the sample temperature from 4 – 42 °C does not significantly affect the proportions of the multiple charge state envelopes.
4.2.2.2 Denaturing Conditions Shift the Charge State Envelope

The mass spectrum of Josephin under ‘denaturing’ conditions (50 % (v/v) acetonitrile, 1 % (v/v) TFA, pH 2.0) showed a single charge state envelope from $n = +22$ to $+8$, which was centered around $n = +18$ (Fig 4.5A). This is similar to the partially unfolded conformation seen under native conditions which centers on $n = +18$. The overlap suggests that either the partially unfolded conformation is similar in conformation to the denatured state, or it has partially unfolded in such a way that a similar amount of charge is able to bind during the ionization process.

The far-UV CD spectrum of Josephin has previously been well characterized (206,208). Far-UV CD was used to confirm that the conformations seen in the mass spectra were not due to effects of the atypical buffers affecting Josephin structure (Fig 4.5B). The volatile ammonium acetate/ammonium bicarbonate pH 7.4 buffer used for native conditions, and the low pH acetonitrile buffer are quite different to the strong tris or phosphate buffered salines typically used to buffer Josephin. However, in all of these conditions, the secondary structure of Josephin is clearly unchanged. Spectral analysis using CONTIN indicated that Josephin has approximately 39 % $\alpha$-helix, 26 % random coil and 34 % $\beta$-sheet and turn structure which is identical to that previously reported. Although widely used to measure protein structure in solution, CD is a measure of the average structure of a protein, and thus does not detect multiple conformations which co-exist in solution.
Figure 4.5 **Josephin displays one conformation under denaturing conditions.** (A) The ESI-MS of Josephin in 50 % (v/v) acetonitrile, 1 % (v/v) TFA, pH 2.0 shows a single charge state envelope centered on \( n = +18 \). (B) Far-UV CD spectra of Josephin in 50 % (v/v) acetonitrile, 1 % TFA (v/v), pH 2.0 (red), 100 mM Tris, 80 mM NaCl, 10 % (v/v) Glycerol, pH 7.4 (blue) and 50 mM ammonium acetate, 50 mM ammonium bicarbonate, pH 7.4 (green). The spectra were obtained at 20 ºC at a protein concentration of 30 μM.

4.2.2.3 The Partially Unfolded Conformations are Not Linked to Aggregation

The intrinsic ability of Josephin to aggregate (233,234) led to the question: are the partially unfolded conformations linked to the aggregation pathway? In order to investigate this, two mutants of Josephin which have altered thermodynamic stability and aggregation kinetics were used. Josephin L169H and Josephin R103G S81A respectively have midpoints of aggregation of 19 fold faster and 2.5 fold slower than Josephin (Ch 3). Analysis of the peak area in each charge state envelope show similar proportions of the folded and unfolded conformations for both the mutants and wildtype Josephin (Fig 4.6), and thus the partially unfolded conformations cannot be directly linked to the aggregation pathway.
Fig 4.6 The multiple conformations are unaffected by amyloidogenic mutation or Josephin binding partners. (A) The proportion of the folded conformation (grey bar) to the partially unfolded conformations (white bar) was calculated from the area under each charge state envelope from Josephin, the stabilizing mutant Josephin R103G S81A and the destabilizing mutant Josephin L169H. (B) The mass spectra of Josephin L169H and Josephin R103G S81A under native conditions obtained using nanoESI-MS show multiple charge state envelopes similar to wildtype Josephin.

The effect of known protein interactors of Josephin on the proportions of the conformations were also investigated. Josephin was incubated with HHR23A, HHR23B and α-B-crystallin, which all impact the aggregation of Josephin in vitro (Ellisdon &
Bottomley, personal communication)(App 1) and in each case the proportions of the charge state envelopes in the Josephin mass spectra were unchanged (data not shown).

Ubiquitin was incubated with Josephin at a ratio of 2:1 Ubiquitin:Josephin, as it has been demonstrated previously that Ubiquitin binds Josephin in this ratio (169). Although the ubiquitin signal is significantly stronger than that of Josephin due to its higher concentration, multiple charge state envelopes are still observed from Josephin, suggesting that the binding of Ubiquitin does not change the presence of the partially unfolded forms (PUF) (Fig 4.7).

**Figure 4.7 Ubiquitin does not change the Josephin spectra.** (A) The mass spectra of ubiquitin alone. (B) The mass spectra of Josephin with 2:1 Ubiquitin:Josephin. The peaks corresponding to ubiquitin are indicated using arrows.

### 4.2.3 Josephin HX MS

#### 4.2.3.1 Deuteration of Josephin

Josephin has a molecular mass of 24,093 Da, and has 366 exchangeable protons which thus gives a theoretical deuterated mass of 24,459 Da (Table 4.2). For some proteins, such as lysozyme, there is a protected region within the protein which does not
undergo deuteration (75). Therefore we initially investigated whether Josephin reaches the theoretical deuterated mass under experimental exchange conditions.

H→D exchange was initially completed using an enclosed electrospray source to eliminate back exchange from the atmosphere, and showed that Josephin has little protection from exchange. 325 protons were exchanged within the first time point of 2 min and the theoretical end point mass of 24,444 Da in a 96 % deuterated solvent was reached after 3.25 hr (Fig 4.8).

### Table 4.2 Calculation of Theoretical Deuterated Mass of Josephin

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Number of Amino Acids</th>
<th>Increase in Mass</th>
<th>Deuterated Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>10</td>
<td>1</td>
<td>720.80</td>
</tr>
<tr>
<td>Arg</td>
<td>8</td>
<td>4</td>
<td>1281.52</td>
</tr>
<tr>
<td>Asn</td>
<td>10</td>
<td>3</td>
<td>1170.90</td>
</tr>
<tr>
<td>Asp</td>
<td>9</td>
<td>2</td>
<td>1053.81</td>
</tr>
<tr>
<td>Cys</td>
<td>4</td>
<td>2</td>
<td>420.56</td>
</tr>
<tr>
<td>Gln</td>
<td>15</td>
<td>3</td>
<td>1966.95</td>
</tr>
<tr>
<td>Glu</td>
<td>18</td>
<td>2</td>
<td>2359.98</td>
</tr>
<tr>
<td>Gly</td>
<td>13</td>
<td>1</td>
<td>754.65</td>
</tr>
<tr>
<td>His</td>
<td>10</td>
<td>2</td>
<td>1391.40</td>
</tr>
<tr>
<td>Ile</td>
<td>12</td>
<td>1</td>
<td>1369.92</td>
</tr>
<tr>
<td>Leu</td>
<td>26</td>
<td>1</td>
<td>2968.16</td>
</tr>
<tr>
<td>Lys</td>
<td>8</td>
<td>3</td>
<td>1049.36</td>
</tr>
<tr>
<td>Met</td>
<td>6</td>
<td>1</td>
<td>793.20</td>
</tr>
<tr>
<td>Phe</td>
<td>13</td>
<td>1</td>
<td>1926.34</td>
</tr>
<tr>
<td>Pro</td>
<td>6</td>
<td>-</td>
<td>582.72</td>
</tr>
<tr>
<td>Ser</td>
<td>16</td>
<td>2</td>
<td>1425.28</td>
</tr>
<tr>
<td>Thr</td>
<td>6</td>
<td>2</td>
<td>618.60</td>
</tr>
<tr>
<td>Trp</td>
<td>3</td>
<td>2</td>
<td>564.63</td>
</tr>
<tr>
<td>Tyr</td>
<td>8</td>
<td>2</td>
<td>1321.36</td>
</tr>
<tr>
<td>Val</td>
<td>7</td>
<td>1</td>
<td>700.91</td>
</tr>
<tr>
<td>Plus N-H₂ &amp; -OH on ends</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Theoretical Deuterated Mass:</strong></td>
<td><strong>24461.05</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Although hydrogen exchange can be analyzed using ESI-MS, the lower signal due to the electrospray source in addition to the decrease in signal which occurs due to deuteration resulted in only the folded charge state envelope being observed (Fig 4.8). Therefore to try to improve the quality of the spectra, we investigated whether the nanoESI source could be used to follow Josephin HX. As the nanoESI source is open to the atmosphere, analyzing D→H exchange is preferable in order to eliminate the effects of back exchange during ionization.

**Figure 4.8 Complete deuteration of Josephin.** Josephin is fully deuterated after 3.25hr at 37 °C with a +9 charge state of 2716.12. Using an electrospray source, the partially unfolded forms were not observed after deuteration.

### 4.2.3.2 D→H Exchange

To complete D→H exchange, the protein must be fully deuterated prior to initiating exchange. A number of techniques were used to buffer exchange Josephin into a deuterated buffer due to problems with Josephin aggregating during either the deuteration process or soon after the initiation of exchange. Buffer exchange was attempted using a number of chromatography columns (Superdex75, NAP5 and Desalting HiTrap), dilution
using a centrifugal concentrator and dialysis (Fig 4.9). Except for dialysis, which was incubated at 37 °C during exchange, Josephin was then incubated at 37 °C for 3.5 hrs to allow full deuteration to occur.

Figure 4.9 Techniques used to achieve deuteration of Josephin for D→H exchange.
Josephin was incubated at a dilute concentration in order to prevent aggregation. Following incubation, the deuterated Josephin then needed to be concentrated as the HX reaction is initiated via a 1:20 dilution into protonated buffer. Concentration was completed using either lyophilization or a centrifugal concentrator. Regardless of which technique was used to deuterate and concentrate Josephin, aggregation occurred within 1 hr of initiation of HX, resulting in no signal on the MS (data not shown).

4.2.3.3 H→D Exchange Using NanoESI-MS Demonstrates the Multiple Conformations of Josephin Exist in Equilibrium

Initially, there was a concern that nanoESI-MS could not be used for H→D exchange, due to the back exchange which occurs during the ionization process with an open air source. However using β2-microglobulin, a protein for which the HX rate has been previously established (250), it has been determined that although back exchange prevents the theoretical end mass being reached, the kinetics of exchange are unchanged (Hodkinson & Radford, personal communication). Therefore in order to firstly eliminate the extra deuteration step needed for D→H exchange and to secondly observe the exchange rates of the partially unfolded conformations, H→D exchange of Josephin was followed using a nanoESI source.

Josephin HX was initiated by diluting Josephin 25 fold into deuterated buffer, pH 7.4 (corrected) and analyzing time points using nanoESI-MS. HX is dependent on temperature, and as at 37 °C Josephin was fully deuterated in 3.25 hr, 20 °C was used for kinetic analysis as it resulted in slower and thus more reproducible kinetics.
Therefore following H→D exchange using nanoESI-MS under native conditions at 20 °C, there is a gradual shift in the n = +8 charge state to a higher m/z over time (Fig 4.10A). This is indicative of an EX2 mechanism of exchange which is characterized by small rapid fluctuations in the protein that gradually allow the incorporation of deuterons (240). The smaller peak in Figure 4.10A is due to Cu$^{2+}$ adducts.

The incorporation of deuterium was plotted as a change in mass over time for all three conformations (Fig 4.10B), and the observed rate of exchange ($k_{obs}$) of 0.028 ± 0.01 min$^{-1}$ was the same for all three conformations. This confirms that the conformations are present in an equilibrium, and must be rapidly interconverting at a rate faster than that of exchange, as has been seen previously in apomyoglobin (251).
Figure 4.10 Josephin HX confirms multiple conformations are in equilibrium. (A) H→D HX of the Josephin domain at 20 ºC, pH 7.4 monitored using nanoESI-MS. The increase in the n = +8 charge state is shown at 2 min, 20 mins and 290 mins. (B) The increase in molecular weight over time was plotted for each charge state envelope: the folded conformation centered around n = +8 (■), the first PUF centered around n = +13 (▲) and the second PUF centered around n = +18 (▼). The data were fit to a single exponential function.

4.2.3.4 Analysis of Josephin HX Kinetics

The contribution of structure to the HX rate of a particular sequence can be considered by comparing it to the intrinsic exchange rate of the sequence if it were in a random coil structure (244, 252). Using model peptides, Bai et al. have investigated the contribution of sequence to the intrinsic exchange rate of unprotected amino acids to determine the overall exchange rate for each amino acid in an unfolded sequence (244).
Using these data to determine the rate of exchange of Josephin if it were in a random coil structure under the conditions used in this study, the intrinsic rate of exchange or $k_{int}$ is 3700 events min$^{-1}$. The difference between these two rates thus fits with the high level of structure known to be present in Josephin. Based on these rates of exchange occurring under an EX2 mechanism, the apparent free energy of hydrogen exchange ($\Delta G_{ex}^{app}$) for Josephin was calculated to be $0.69 \pm 0.06$ kcal mol$^{-1}$.

For comparison, the equilibrium unfolding curve for Josephin under the conditions used for hydrogen exchange was determined using the denaturant GdnHCl and following the change in secondary structure at 222 nm using far-UV CD (Fig 4.11). The data were fit to a two-state unfolding mechanism, and the midpoint of the transition occurred at 2.47 M GdnHCl. Similar to results previously published for Josephin (234), this suggests that the volatile buffer does not significantly impact upon the thermodynamic stability of the protein. From the equilibrium unfolding data, the free energy of unfolding ($\Delta G$) was calculated to be $1.7 \pm 0.3$ kcal mol$^{-1}$.

![Figure 4.11 Josephin equilibrium unfolding is unaffected by a volatile buffer. Equilibrium unfolding curve for Josephin performed in the conditions used for HX (50 mM ammonium acetate, 50 mM ammonium bicarbonate, pH 7.4 at 20 °C), using the denaturant GdnHCl. The data were fit to a two-state unfolding curve.](image)

The analysis of the HX data relies upon the assumption that the protein remains monomeric throughout the timescale of the experiment. As Josephin is prone to
aggregation, we used size exclusion chromatography (SEC) and thioflavin T (thioT) fluorescence to confirm that the protein remained monomeric during the experiments (Fig 4.12). Josephin was incubated under the same conditions used for hydrogen exchange, and at specific time points aliquots were removed and analyzed using SEC and thioT. Both thioT and SEC showed no change over time, thus suggesting that there are no multimeric species present and Josephin has remained monomeric throughout the timescale of the HX experiments.

Figure 4.12 Josephin remains monomeric throughout the timescale of HX. (A) Size exclusion chromatography of Josephin incubated in the conditions used for HX. The area under the monomer peak at 19 mL was calculated and is shown at timepoints across the length of the HX experiment. No loss of monomer occurred during this timescale. (B) ThioT fluorescence also demonstrates there is no formation of fibrillar species over 5 hrs at 20 °C, as there is no increase in thioT fluorescence in comparison to the thioT only negative control. A positive control using Josephin endpoint fibrils is also shown.
4.3 DISCUSSION

4.3.1 Implications of the Conformational Heterogeneity of Josephin

The ESI-MS data provides new insight into the conformational heterogeneity of the Josephin domain. Techniques such as CD, NMR and intrinsic fluorescence have previously suggested that Josephin is in a stable, highly structured conformation (166,233,234). The mass spectra under native conditions instead shows that Josephin co-populates a number of conformations in solution (Fig 4.3). Although the spectra of most proteins show a single charge state envelope, there have been a number of proteins reported which also co-populate several conformations under native conditions and correspondingly have multiple charge state envelopes (248,251,253). More commonly, multiple conformations can be induced using denaturing conditions and used to identify partially unfolded intermediates which may exist either on or off the folding pathway (254).

The multiple charge state envelopes were however present under native conditions for Josephin. The HX data (Fig 4.10) suggest that the multiple conformations are rapidly interconverting in equilibrium, as if a fixed population in an unfolded conformation existed then different rates of HX within each charge state envelope would be obtained. The most folded population binds the least number of charges (245), and thus in the Josephin mass spectra the folded conformation is represented by the charge state envelope centred around n = +8 (Fig 4.3A). In addition to this, there were two charge state envelopes with lower m/z ratios which thus corresponded to PUF (255). Based on the peak area under each charge state envelope, 40% of the population was calculated to be unfolded. However previous studies show that non-native polypeptides have a higher signal relative to native, with up to two fold difference between the ion distribution and the distribution actually in
solution (256). Thus based on ESI-MS alone, we cannot definitively estimate the proportion of partially unfolded Josephin molecules in solution, and a complementary technique such as single molecule fluorescence is needed to determine this.

The increase in charge which binds to unfolded proteins during positive ESI-MS is generally attributed to an increase in surface area and thus an increased steric accessibility of possible protonation sites (257,258). The extent of unfolding in the PUF of Josephin is not clear. A clue to this is provided by the mass spectra of Josephin in acetonitrile, whereby only the most unfolded charge state envelope remains (Fig 4.5 A). Although there is a significant difference in the mass spectra between acetonitrile and native conditions, the CD spectra under both conditions are identical (Fig 4.5 B), therefore suggesting that the secondary structure is unchanged and thus the PUF are not significantly unfolded. NMR dynamics studies of Josephin have shown that the helical hairpin, which contains the $\alpha_2$ and $\alpha_3$ helices (Fig 4.2), is significantly more dynamic than the rest of the protein (166). Therefore one possibility is that the flexibility of the helical hairpin region allows for the binding of additional charge in some conformations and thus the multiple charge state envelopes seen in the mass spectra. Alternatively, the conformations could be due to the unfolding of another region that has not been identified as dynamic in the NMR spectra due to the timescale of exchange.

The intrinsic aggregation propensity of Josephin means these PUF are of particular interest, however the proportion of each conformation was unchanged in several mutants with differing amyloidogenicity (Fig 4.6). This suggests that the partially unfolded conformations are not linked to the aggregation pathway. For some proteins which populate multiple conformations in solution, one conformation can be stabilized via the
interactions of a binding partner (248, 259-261). This was not however the case for Josephin as known protein interactors ubiquitin, HHR23A or B, and α-B-crystallin all did not affect the proportions of the conformations. Although the role of the PUF could not be identified within this study, the conformational heterogeneity could be linked to the unusual DUB activity displayed by Josephin (168).

The folding pathway of Josephin has been previously characterized (233, 234), and it is important to consider how the PUF may fit into this pathway. A scheme is proposed in Fig. 4.13 whereby the partially unfolded conformations are in equilibrium with the native state. Evidence for this comes from the Josephin stability mutants (Fig 4.6) which have thermodynamic stabilities that are either increased (Josephin R103G S81A) or decreased (Josephin L169H) from wild type. Although the changed stability of these mutants affect the N ↔ I equilibrium, they do not change the relative proportions of N and the PUF, and this suggests that the PUF are not directly linked to the intermediate states on the folding pathway.

4.3.2 The Dynamic Nature of the Josephin Domain

This study found that Josephin exists in an equilibrium between multiple conformations, however although ESI-MS of Josephin has also been completed by another
group, the multiple conformations were not previously observed (234). This may be attributed to nanoESI-MS being used in this study, as it results in cleaner signals and thus allows the characterization of the less populated PUF during HX. The HX rate could not be compared with the previous study as key information such as the temperature used for exchange was not disclosed.

The HX kinetics of Josephin demonstrates that it is a highly dynamic protein in comparison to other amyloidogenic proteins, which may explain its high propensity to form fibrils \textit{in vitro} under native conditions. For example β2-microglobulin, a protein which readily forms fibrils under low pH conditions, has a similar HX rate reported to Josephin, however the experiments were completed at 37 °C which is 17 °C higher than the temperature Josephin HX was completed under. Given the temperature dependence of HX (262), this suggests that Josephin is significantly more dynamic. This dynamic nature potentially links to the unique ability of Josephin to cleave multi-linked ubiquitin chains (168).

In this study a system was set up to determine the conformational dynamics of the Josephin domain in order to provide insights into the normal function and early misfolding events of ataxin-3. Future experiments to determine the impact of full length ataxin-3 both with and without an expanded polyQ tract upon the dynamics of this domain may provide insight into the increased aggregation kinetics observed with increased polyQ tract length (191).
CHAPTER 5

AGGREGATION OF ATAXIN-3 IN SDS: EVIDENCE FOR ALTERNATIVE AGGREGATION PATHWAYS

PREFACE

The focus on characterizing aggregation in vitro in this thesis allows the conformational changes of ataxin-3 which occur during aggregation to be determined. The limitation of this approach is that ataxin-3 containing any length polyQ tract has been demonstrated to aggregate in vitro, which does not mirror disease. This therefore implicates the role of other cellular components in ataxin-3 pathogenesis. Cellular membranes are one such factor which have been demonstrated to affect the aggregation of a number of amyloid proteins (263-266). Within the polyQ proteins, Kegel and colleagues have completed several studies that demonstrate that huntingtin interacts with phospholipids at the plasma membrane within the cell, and an expanded polyQ tract
increases this interaction (116-118). The question remains as to whether the interactions observed between huntingtin and phospholipids are common to all polyQ proteins.

There are a number of studies that suggest ataxin-3 may be involved with lipids as part of its normal cellular function. Firstly, ataxin-3 has been demonstrated to transiently associate with the ER membrane through its interaction with binding partner VCP (267). In addition to this, monomeric ataxin-3 has been isolated from mitochondrial membranes (268). Although it has been demonstrated that ataxin-3 with an expanded polyQ tract results in disrupted electrophysiological properties in a cell model, the impact of lipids on aggregation has not been investigated for ataxin-3. Therefore the last part of this thesis investigates the effects of the membrane mimic SDS on the in vitro ataxin-3 aggregation pathway, in addition to determining the ability of ataxin-3 to bind to various phospholipids.
Conformational Behavior and Aggregation of Ataxin-3 in SDS

Helen Saunders*, Andrew M. Ellisdon*, Roberto Cappai‡ and Stephen P. Bottomley*

*Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia.
‡Department of Pathology and Bio21 Molecular Science & Biotechnology Institute, The University of Melbourne, Parkville, Victoria 3010, Australia.

Correspondence should be addressed to: Stephen P. Bottomley, Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800, Australia. Tel: 61-3-99029362; Fax. 61-3-99029500; Email: steve.bottomley@monash.edu

Running title: Aggregation and conformation of ataxin-3 in SDS
Abstract

Spinocerebellar ataxia type 3 (SCA3) is one of nine polyglutamine (polyQ) diseases all characterized by the presence of intraneuronal inclusions which contain aggregated protein. Aggregation of ataxin-3, the causative protein of SCA3, has been well characterized \textit{in vitro}, with both pathogenic and non-pathogenic length ataxin-3 undergoing fibrillogenesis. However, only ataxin-3 containing an expanded polyQ tract leads to SCA3. Therefore other cellular factors, not present in previous \textit{in vitro} studies, may modulate aggregation during disease. The interactions between fibrillar species and cell membranes have been characterized in a number of amyloid diseases, including the polyQ disease Huntington Disease, and these interactions affect aggregation and toxicity. We have characterized the effects of the membrane mimetic sodium dodecyl sulfate (SDS) on ataxin-3 structure and aggregation, to show that both micellar and non-micellar SDS have differing effects on the two stages of ataxin-3 aggregation. We also demonstrate that fibrillar ataxin-3 binds phospholipids, in particular phosphorylated phosphotidylinositol. These results highlight the mitigating effect of intracellular factors on the ataxin-3 misfolding landscape and implications of these data in SCA3 and polyQ diseases in general are discussed.
Introduction

Ataxin-3 misfolding and its subsequent aggregation underlies the autosomal dominant neurodegenerative disease Spinocerebellar ataxia type 3 (SCA3). This disease is characterized by progressive neuronal dysfunction and the presence of neuronal nuclear inclusions which contain aggregated ataxin-3. The polyglutamine (polyQ) protein ataxin-3 functions as a de-ubiquitinating enzyme, and consists of an N-terminal catalytic Josephin domain which has structural homology to papain-like cysteine proteases, and a comparatively unordered C-terminal region containing two ubiquitin interaction motifs and the polyQ tract. Expansion of this polyQ tract to greater than 45 residues results in protein aggregation and disease, with the age of onset inversely correlated with repeat length.

Knowledge of the kinetic and structural changes involved in ataxin-3 misfolding and aggregation will help us to understand the molecular events and disease progression involved in SCA3. The structural changes and kinetics involved with the in vitro aggregation mechanism of ataxin-3 has been characterized. These data indicate that ataxin-3 aggregation involves a two-stage aggregation pathway with interactions facilitated initially by the Josephin domain and subsequently by the polyQ tract. Despite both the Josephin domain as well as the non-pathogenic length ataxin-3 forming the first stage fibrils in vitro, contradicting data exists regarding the presence of non-pathogenic length ataxin-3 aggregates in cells. However, evidence from various polyQ proteins and model systems increasingly suggests that this multi-stage mechanism is not unique to ataxin-3, and that the flanking regions of the polyQ tract impact upon polyQ aggregation (see review).
The intrinsic fibrillogenic nature of both pathogenic and non-pathogenic length ataxin-3 implicates other cellular factors in disease pathogenesis. As a significant proportion of the cellular environment, membranes of varying compositions influence the aggregation of amyloid proteins such as amyloid β-peptide, α-synuclein and prion protein. Of the polyQ proteins, huntingtin binds various cellular membranes with some evidence that it forms ion channels within bilayer membranes. Aggregates formed from polyQ peptides are internalized by mammalian cells and cross the cell membrane to gain access to the cytoplasmic compartment. Ataxin-3 has been proposed to associate with cellular membranes in several ways. Within the cell ataxin-3 transiently associates with membranes via its binding partner VCP, in addition to directly binding mitochondrial membranes. Interestingly, both huntingtin and ataxin-3 perturb the structure of synthetic lipid bilayers when oligomeric in structure, however the impact of membranes and specific lipids on ataxin-3 structure and aggregation is unknown.

Acidic phospholipids, which are present in a number of intracellular membranes, accelerate the aggregation of numerous fibrillogenic proteins including huntingtin. The concentration of positively charged phospholipids within the membrane creates a local environment with an acidic pH, which in turn accelerates aggregation. The detergent Sodium Dodecyl Sulfate (SDS) is a membrane mimetic at micellar concentrations and at low non-micellar concentrations mimics anionic phospholipids due to its positively charged head group and long tail. SDS has the ability to induce changes in secondary structure and it can be used to probe the conformational change events occurring during aggregation. The impact of SDS on the aggregation kinetics of amyloidogenic proteins such as β2-
microglobulin, amyloid-β and α-synuclein has been determined\textsuperscript{35; 36; 37}; however the effects of SDS on polyQ proteins have not been investigated to date.

Using biophysical techniques, we demonstrate that in the presence of SDS, ataxin-3 is able to form aggregates via a number of alternate pathways. We investigate the effects of both micellar and sub-micellar concentrations of SDS on ataxin-3 and show that there are differential effects of SDS at different points of the multi-stage ataxin-3 aggregation pathway. Finally, we show that oligomeric and fibrillar ataxin-3 binds acidic phospholipids, in particular phosphotidylinositolns, with different specificities.
Results

SDS induces an $\alpha$-helical conformation in ataxin-3

SDS forms micelles at concentrations above the critical micelle concentration (CMC) and in the buffer conditions used within this study the CMC of SDS was determined to be 1.2 mM (data not shown)\textsuperscript{35, 38}. SDS has previously been demonstrated to induce helical secondary structure in a range of proteins at concentrations above the CMC\textsuperscript{34, 35, 36}. In this study, the effects of SDS on pathogenic length ataxin-3(Q64), non-pathogenic length ataxin-3(Q15) and the Josephin domain were investigated. Changes in the secondary structure of the ataxin-3 variants with the addition of up to 10 mM SDS were analyzed using far-UV CD spectroscopy (Fig. 1). Consistent with previous reports, all ataxin-3 variants in the absence of SDS displayed spectra with minima at 208 nm and 222 nm, indicative of predominantly $\alpha$-helical secondary structure\textsuperscript{12, 39}. Only minor changes in secondary structure were observed when SDS was added. In the presence of 1 mM SDS, no significant change in structure occurred for any of the proteins, whereas above 5 mM SDS there was an average increase in $\alpha$-helical structure of 5 % for all proteins (Table 1). The magnitude of the structural changes induced by SDS in ataxin-3(Q64) (Fig. 1A), ataxin-3(Q15) (Fig. 1B) and the isolated Josephin domain (Fig. 1C) were similar, thus suggesting that the changes in secondary structure occur predominantly within the Josephin domain.

SDS modulates SDS-soluble aggregation of ataxin-3

Ataxin-3 aggregation occurs via a two-stage mechanism. The first stage involves the formation of SDS-soluble curvilinear protofibrils and is common to all ataxin-3
variants. In the second stage of aggregation, only pathogenic length ataxin-3 forms SDS-insoluble fibrils which have a straighter morphology. Formation of the first stage SDS-soluble fibrils was monitored by following changes in thioflavin T (thioT) fluorescence as previously described. Without SDS, all ataxin-3 variants show a sigmoidal aggregation curve indicative of a nucleation-dependent process, with a lag phase followed by exponential growth which then plateaus. The overall aggregation kinetics vary such that the isolated Josephin domain has the slowest aggregation rate and ataxin-3(Q64) the fastest (Fig. 2). The presence of 1 mM SDS eliminated the lag phase of all the ataxin-3 variants, resulting in an immediate exponential growth phase with a rate independent of polyQ length.

The midpoint of aggregation decreased to two hours for all proteins with 1 mM SDS present, suggesting that the small conformational change induced by SDS is similar for all proteins and results in a highly amyloidogenic species. In addition, 1 mM SDS also resulted in hyperfluorescence of thioT (Fig. 2) which may be related to the greater number of short fibrils being formed, as visualized using transmission electron microscopy (TEM) (Fig. 5), thereby creating more binding sites. In contrast, at both 5 mM and 10 mM SDS, there is no increase in thioT fluorescence for any of the ataxin-3 variants, thus suggesting that fibril formation is suppressed at these micellar SDS concentrations. These results, in which a specific range of SDS concentrations around the CMC modulate thioT detected fibrillogenesis, are consistent with those previously reported for a range of other non-polyQ amyloid proteins.
**SDS-insoluble aggregation of pathogenic length ataxin-3 is affected by SDS**

Having demonstrated that the effects of SDS on the formation of SDS-soluble stage 1 fibrils are similar to those previously reported for other amyloid proteins, we then investigated the effects of SDS on the second stage of the multi-stage ataxin-3 aggregation pathway. The formation of stage 2 aggregates can be monitored using a membrane filter trap assay in which after boiling in 2 % (w/v) SDS, only SDS-insoluble fibrils are retained upon the filter. It should be noted that the concentration of SDS used to solubilize the fibrils (2 % (w/v) SDS) is six times greater than the concentrations around the CMC which impact upon fibrillogenesis (0.03 % - 0.29 %). It can clearly be seen that in the absence of SDS ataxin-3(Q64) forms SDS-insoluble aggregates following 51 hrs of incubation (Fig. 3A). With 0–10 mM SDS, both non-pathogenic length ataxin-3(Q15) and the Josephin domain did not form SDS-insoluble aggregates when incubated up to 200 hours (data not shown).

Although it was anticipated that due to the sequential nature of the aggregation pathway modulation of stage 1 aggregation would mediate similar kinetic effects on stage 2 aggregation, quite different results were obtained. Despite 1 mM SDS significantly accelerating the rate of SDS-soluble fibril formation as detected by thioT fluorescence (Fig. 2A), the formation of SDS-insoluble fibrils by ataxin-3(Q64) with 1 mM SDS is slower and has a midpoint 24 hours later than in the absence of SDS (Table 2). In contrast, ataxin-3(Q64) with 5 mM SDS shows no increase in thioT fluorescence over 100 hours (Fig. 2A) yet the formation of SDS-insoluble aggregates appear in the filter trap assay from 51 hours (Fig. 3A), suggesting that the aggregates are forming on an alternative aggregation
pathway. Addition of the polyQ-binding peptide QBP1 eliminated the formation of SDS-insoluble aggregates formed by ataxin-3(Q64) with 5 mM SDS.

In a strongly micellar environment (10 mM SDS), ataxin-3(Q64) does not show the formation of either SDS-soluble (Fig. 2A) or SDS-insoluble aggregates (Fig. 3). Several studies have reported distinct effects of micellar and non-micellar SDS concentrations, with micellar concentrations inhibiting aggregation, and non-micellar concentrations accelerating aggregation\textsuperscript{34; 35; 36}. However with the complexity of the two-stage ataxin-3 aggregation mechanism, there is not such a clear distinction. Only the highly micellar 10 mM SDS inhibits both stages of the pathway, and the other SDS concentrations differentially affect the first and second stages of aggregation.

\textit{SDS modulates the change to \(\beta\)-sheet secondary structure typical of aggregation}

With the intriguing formation of thioT unreactive fibrils by 5 mM SDS, we then went on to characterize the changes in secondary structure occurring during aggregation. SDS induces an increase in \(\alpha\)-helical structure at concentrations above the CMC (Fig. 1), however a key event in fibrillogenesis is the gain of \(\beta\)-sheet structure, and hence far-UV CD was used to follow the impact of SDS upon this structural transition.

As previously reported, we observed that in the absence of SDS ataxin-3(Q64) converts to a \(\beta\)-sheet rich fibrillar species (Fig. 4A). The loss in signal observed over time has been previously suggested to reflect an increase in light scatter\textsuperscript{9}. Incubation in 1 mM SDS (Fig. 4B) accelerates the kinetics of aggregation such that by four hours there has been substantial loss of helical structure and a conversion to \(\beta\)-sheet structure which continues
over time with a loss of signal similar to that seen in the absence of SDS at 100 hours (Fig. 4A).

Incubation of ataxin-3(Q64) in both 5 mM and 10 mM SDS leads to a retention of α-helical structure over 100 hours, and for 10 mM SDS there is a small increase in the minima at 208 nm and 222 nm (Fig. 4D). This is consistent with the lack of aggregation detected with 10 mM SDS throughout this study and suggests that SDS has stabilized the α-helical structure to the extent that the conversion to β-sheet is prevented. With 5 mM SDS present, the retention of α-helical structure over time concurs with the lack of thioT fluorescence observed (Fig. 2) and thus suggests that the SDS-insoluble fibrils being formed (Fig. 3) are more similar to amorphous aggregates than the β-sheet rich amyloid-like fibrils typically formed by ataxin-3. Interestingly, these aggregates are still formed via interactions of the polyQ tract, as addition of QBP1 inhibits their formation (Fig. 3A). The same effects of SDS on the change in secondary structure over time were also observed for ataxin-3(Q15) and the Josephin domain (Fig. S1).

**SDS addition results in changed morphology of fibrils**

The formation of SDS insoluble fibrils which are not thioT reactive suggests that ataxin-3 is forming these fibrils via an atypical pathway. In order to investigate this further, TEM was used to visualize the morphology of the aggregates being formed with 0, 1, 5 and 10 mM SDS present (Fig. 5, Fig. S2). As expected, ataxin-3(Q64) in the presence of 10 mM SDS did not form fibrils (data not shown). In contrast, with 5 mM SDS ataxin-3(Q64) forms amorphous aggregates with a diameter of 20 – 30 nm which although SDS-insoluble,
do not have a fibrillar morphology but appear clumped together to form large amorphous aggregates of up to 0.5 µm in length (Fig. 5C and F).

The fibrils formed by ataxin-3(Q64) in the presence of 1 mM SDS (Fig. 5B) had lengths of 23 - 29 nm and a diameter of 5 nm. These fibrils were initially shorter and smaller in diameter than the curvilinear fibrils formed in the absence of SDS which showed diameters of 12 - 15 nm and lengths of hundreds of nanometers (Fig. 5A) as observed previously. Further incubation resulted in the formation of larger, more rigid SDS-insoluble fibrils which were 40 - 50 nm in width and up to 1 µm in length (Fig. 5D and E).

_PolyQ oligomers interact with acidic phospholipids_

As SDS is a mimetic of acidic phospholipids we then decided to investigate whether ataxin-3 shows a specificity of binding to acidic phospholipids. To test a variety of lipids, we incubated ataxin-3(Q64) from specific stages of the aggregation pathway in a protein-lipid overlay assay. Monomeric ataxin-3(Q64) and the Josephin domain both showed no binding to any of the lipids (data not shown). When a heterogeneous solution taken from the plateau stage of the thioT curve was incubated with the PIP strips, binding to phosphorylated phosphatidylinositols (PtdIns) was observed for both ataxin-3(Q64) and the Josephin domain. Interestingly, there tended to be an additional subset of lipids which bound to the endpoint fibrils of both the Josephin domain and ataxin-3(Q64) (Fig. 6A i - iv), with the endpoint fibrils binding to essentially the same subset of lipids. Ataxin-3(Q64) incubated with QBP1, an inhibitor of polyQ mediated aggregation, showed the same binding pattern as ataxin-3(Q64) thus suggesting that the lipids are predominantly binding to the misfolded Josephin domain (data not shown).
In order to help confirm that the polyQ tract was not involved, we used the model system Staphylococcus protein A (SpA) with an attached polyQ tract of 52 glutamines (SpA(Q52))\textsuperscript{40}. SpA is a membrane-anchored protein and thus native SpA was used as a control. SpA and SpA(Q52) demonstrated similar binding patterns, further suggesting that the polyQ tract is not involved in lipid binding (Fig. 6 v and vi).
Discussion

Effects of SDS on ataxin-3 conformation and aggregation

SDS appears to interact in a common fashion with monomeric fibril-forming proteins, inducing \( \alpha \)-helical structure irrespective of the protein’s initial structure\(^{34; 35; 36; 41} \). It is interesting that this can lead to rapid aggregation, considering the key structural change required for aggregation is the conversion to \( \beta \)-sheet dominated structure. A large number of aggregating proteins, including both disease-causing proteins and the fibril-forming peptide SRC3 show this accelerated aggregation at sub-micellar SDS concentrations, and slowed or inhibited aggregation at SDS concentrations well above the CMC\(^{34; 35; 36; 37; 42} \). The inhibition of aggregation may be due to the highly \( \alpha \)-helical structure stabilized by micellar SDS concentrations (Fig. 1) preventing the transition to \( \beta \)-sheet structure which occurs during aggregation. For ataxin-3 at sub-micellar concentrations, there was significant acceleration of SDS-soluble aggregation, similar to other fibrillogenic proteins, in addition to a significant increase in thioT fluorescence intensity. Pertinhez et al. also show that fibrils in sub-micellar concentrations are shorter than normal, perhaps providing evidence that in this early stage of aggregation SDS directs monomeric protein down similar pathways\(^{34} \).

An interesting result in this study was the differing effects of SDS upon formation of SDS-insoluble ataxin-3 fibrils as opposed to the more commonly observed effects of SDS on SDS-soluble fibril formation. With 1 mM SDS, the lack of an observable lag phase suggests either the lag phase has been accelerated to the extent that it cannot be observed or that ataxin-3 is following an alternate mechanism not involving nucleated elongation. Although the subsequent formation of SDS-insoluble aggregates is slower, the
Effects of SDS on ataxin-3

morphology of the end point aggregates remains unchanged. Thus it appears that with sub-
micellar SDS present, the protein is proceeding through the typical aggregation pathway,
with differing rates of stage 1 and stage 2 aggregation (Fig. 7).

In contrast to 1 mM SDS, the morphology of the aggregates is not fibrillar in the
presence of 5 mM SDS, thus demonstrating that ataxin-3 has the ability to undertake
alternate aggregation pathways to form a range of fibril types (Fig. 7). This is supported by
the lack of both thioT binding and conversion to β-sheet structure. At this micellar
concentration, although there is no formation of thioT reactive SDS-soluble aggregates,
SDS-insoluble aggregates are still formed. These aggregates have a substantially different
morphology, appearing amorphous in structure, however they are still formed via
interactions of the polyQ tract, as formation of these aggregates is inhibited QBP1 (Fig.
3A). The formation of different aggregate morphologies is not unprecedented as
environmental conditions affect the type of aggregate formed by a number of proteins in
vitro\textsuperscript{43; 44}. Within the cell such changes in the intracellular environment could be achieved
by conditions of stress such as elevated temperature or decreased pH in localized
environments caused by specific lipids being clustered in membranes\textsuperscript{31; 45}.

Ataxin-3 oligomers and fibrils displayed a specificity in binding to PtdIns with
varying degrees of phosphorylation. PtdIns are generally located on the cytoplasmic side
of the plasma membrane and are present in specific membranes depending on
phosphorylation, with a higher abundance of these lipids (10 %) in brain tissue\textsuperscript{46}. Although
monomeric huntingtin also bound similar phospholipids\textsuperscript{30}, it appears that this is not a
common polyQ specific effect as only fibrillar species of ataxin-3 showed binding.
Furthermore, when the polyQ-binding peptide QBP1 was added there was no change to the
binding pattern which suggests that binding occurs through the Josephin domain. This is similarly seen in the SDS experiments in this study, where the effect of SDS on the Josephin domain is identical to that on ataxin-3, and unaffected by QBP1.

Phospholipids have been demonstrated to affect aggregating proteins by creating regions which have a local environment with a decreased pH, and through electrostatic interactions which can increase the local concentration of protein at the membrane and induce partial unfolding of proteins. It is interesting that oligomers and fibrillar ataxin-3 bound to the lipid overlay with different specificities as several studies show that oligomers have a generic ability to permeabilize cell membranes by creating pores or single channels within membranes.

Overall, our findings demonstrate the sensitivity of ataxin-3 fibril formation to solution conditions and suggest a possible role for lipid molecules in the development of SCA3. The specificity of binding with only fibrillar species associating with phosphorylated phospholipids provides a link between ataxin-3 and the growing evidence that soluble oligomers disrupt membranes as part of the mechanism of toxicity within amyloidoses.
Materials and Methods

Materials - Phenylmethylsulfonyl fluoride, β-mercaptoethanol and thioT were all obtained from Sigma.

Expression and Purification of Ataxin-3 Variants - All ataxin-3 variants were expressed and purified as previously described\textsuperscript{55}, and the proteins were stored at -80 °C. Following purification the deubiquitinating activity of the proteins was measured\textsuperscript{10} and before use all were analyzed using gel filtration to ensure that no multimeric species were present.

Circular Dichroism - Far-UV CD spectra were measured on a Jasco-810 spectropolarimeter at 37 °C using a path length of 0.1 mm. Scans of monomeric protein in the presence of different concentrations of SDS were carried out. For time course assays, protein aliquots were diluted 1:1 with TBSG (100 mM Tris, 80 mM NaCl, 10 % (v/v) glycerol, pH 7.4), to a final concentration of 30 µM protein. In all scans, spectra were measured from 190 to 260 nm with a scanning speed of 50 nm/min. The CD spectra were analyzed by spectral deconvolution using the CONTINLL algorithm\textsuperscript{56}.

Calculation of the CMC of SDS – To calculate the CMC, the fluorescent probe ANS was used as previously described\textsuperscript{35}. The fluorescence intensity was measured at 475 nm, with an excitation wavelength of 385 nm using a PerkinElmer Life Sciences LS50B spectrofluorometer with a thermostatted cuvette holder and a 1 cm pathlength quartz cuvette. The reaction mix contained 10 µM ANS with 0 – 10 mM SDS in TBSG, pH 7.4 at 37 °C. Two straight lines were fit to the data points with the intersection determining the CMC.

Fibrillogenesis Time Course Assays - Ataxin-3 variants (30 µM) were incubated in TBSG containing 5 mM EDTA, 15 mM β-mercaptoethanol and 2 mM phenylmethylsulfonyl
fluoride, with the addition of 0 – 10 mM SDS as required. Samples were incubated without shaking at 37 °C in air-tight containers to prevent evaporation.

**Thioflavin T Fluorescence** - ThioT fluorescence measurements were recorded on a Gemini platereader, using the buffering conditions described previously with the addition of 30 µM ThT and 0 - 10 mM SDS. Excitation and emission wavelengths of 430 nm and 480 nm with a cut off filter of 455 nm were used, and both excitation and emission were read from the underneath of a black clear bottom plate. All reactions were completed at 37 °C with no shaking.

**Membrane Filter Trap Assay** - Aliquots containing 7.4 µg of protein were taken from the fibrillogenesis reaction, diluted 1:1 with a 4 % (w/v) SDS / 100 mM DTT solution and then boiled for 5 min at 100 °C. 200 µL of 2 % (w/v) SDS was then added to each of the samples, and 5 µg of protein were filtered through a 0.2 µm cellulose acetate membrane (Schleicher and Schuell) using a Bio-Rad Bio-Dot SF microfiltration unit. The membrane was then washed twice by filtering 200 µL of 0.1 % (w/v) SDS, and blotted with a hexahistidine (His6) antibody (Serotec). Densitometry was completed using the program Phoretix 1D Quantifier.

**Transmission Electron Microscopy** - TEM images were obtained using a Hitachi H7500 transmission electron microscope with an accelerating voltage of 80 kV. The samples were adsorbed onto a carbon-coated grid and stained with 1 % (w/v) uranyl acetate.

**Protein-Lipid Overlay** – The protein-lipid overlay assay was performed using fibrillar protein stopped at various timepoints. The protein was incubated with the phospholipids (PIP Strips, Molecular Probes) overnight at 4 °C. Each lipid dot contained 100 pmol of
Effects of SDS on ataxin-3

phospholipids. The membrane was then blotted using His6 primary antibody (Serotec) and peroxidase-conjugated secondary antibody as described\textsuperscript{57}.

Acknowledgements

SPB and RC are Senior Research Fellows of the NHMRC and the work was supported by both the ARC and NHMRC of Australia.
Effects of SDS on ataxin-3

References


Tables

Table 1 Percentage of α-helical content of monomeric protein with SDS present.

<table>
<thead>
<tr>
<th>[SDS] mM</th>
<th>Ataxin-3(64)</th>
<th></th>
<th>Ataxin-3(Q15)</th>
<th></th>
<th>Josephin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% α-helix</td>
<td>Standard Error</td>
<td>% α-helix</td>
<td>Standard Error</td>
<td>% α-helix</td>
</tr>
<tr>
<td>0</td>
<td>27.5</td>
<td>2.3</td>
<td>30.3</td>
<td>2.6</td>
<td>30.8</td>
</tr>
<tr>
<td>1</td>
<td>28.7</td>
<td>2.7</td>
<td>30.4</td>
<td>2.0</td>
<td>29.5</td>
</tr>
<tr>
<td>5</td>
<td>32.0</td>
<td>1.9</td>
<td>37.5</td>
<td>2.0</td>
<td>36.4</td>
</tr>
<tr>
<td>10</td>
<td>32.0</td>
<td>1.8</td>
<td>39.1</td>
<td>2.9</td>
<td>35.4</td>
</tr>
</tbody>
</table>

Table 2 Midpoints of ataxin-3(Q64) aggregation

<table>
<thead>
<tr>
<th>[SDS] mM</th>
<th>SDS-Soluble Aggregation</th>
<th></th>
<th>SDS-Insoluble Aggregation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Midpoint (hrs)</td>
<td>Standard Error (hrs)</td>
<td>Midpoint (hrs)</td>
</tr>
<tr>
<td>0</td>
<td>11.50</td>
<td>1.65</td>
<td>44.7</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0.5</td>
<td>69.0</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>79.9</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
**Figure Legends**

**Figure 1 Far-UV CD spectra of ataxin-3 variants in increasing concentrations of SDS.**
The far-UV CD spectra for each ataxin-3 variant was determined at 37 ºC with increasing concentrations of SDS; 0 mM SDS (*black solid line*), 1 mM SDS (*black dotted line*), 5 mM SDS (*grey solid line*) or 10 mM SDS (*grey dashed line*). The final protein concentration was 30 µM and the spectra measured with a path length of 0.1 mm.

**Figure 2 Aggregation of ataxin-3 in the presence of SDS monitored by ThioT.**
Aggregation of ataxin-3 (30 µM) at pH 7.4 and 37 ºC in the presence of a range of SDS concentrations was monitored by thioT. ThioT fluorescence values were read at 480 nm (λ<sub>ex</sub> = 430 nm) every 30 minutes using a fluorescence plate reader. (A) Ataxin-3(Q64), (B) ataxin-3(Q15) and (C) Josephin domain are shown with the addition of 0 mM SDS (*black solid line*), 1 mM SDS (*dashed line*), 5 mM SDS (*grey solid line*) and 10 mM SDS (*dotted and dashed line*).

**Figure 3 Aggregation of ataxin-3 in the presence of SDS monitored by SDS-insolubility.** Formaition of SDS-insoluble fibrils was followed by taking aliquots from a 30 µM ataxin-3(Q64) timecourse assay at 37 ºC, pH 7.4. (A) A representative filter-trap membrane of ataxin-3(Q64) with 0 – 10 mM SDS is shown. QBP1 was added to a ataxin-3(Q64) containing 5 mM SDS as indicated. (B) Analysis of the filter trap membrane by densitometry. Ataxin-3(Q64) is shown with the addition of 0 mM SDS(■ – ), 1 mM SDS(-▲-), 5 mM SDS (-♦- ), 5 mM SDS with QBP1 (-♦ -) and 10 mM SDS (● – ). Results from three independent experiments were fit to an exponential curve.
Figure 4  Far–UV CD spectra of ataxin-3(Q64) during aggregation in the presence of SDS. Aliquots of ataxin-3(Q64) were taken from a fibrillogenesis time course assay (see Materials and Methods) and their far-UV CD spectra determined. For each indicated SDS concentration aliquots were taken at times of 0 (black), 4 (green), 46 (red) and 100 (blue) hours.

Figure 5 Morphology of fibrils formed by ataxin-3(Q64) with SDS. Transmission Electron Microscopy of ataxin-3(Q64) with 0 mM SDS at 30 hr (A) and 50 hr (D), 1 mM SDS at 5 hr (B) and 30 hr (E) and 5 mM SDS at 30 hr (C) and 50 hr (F). Time course samples from 5 hr, 30 hr or 50 hr were negatively stained using 1 % (w/v) uranyl acetate. Scale bars represent 200 nm.

Figure 6 Binding of polyglutamine proteins to phospholipids. (A) Protein-lipid overlays of ataxin-3(Q64) at 24 hrs (i) and 200 hrs (ii), Josephin domain at 70 hrs (iii) and 200 hrs (iv), and monomeric SpA (v) and SpA(Q52) (vi). A representative membrane is shown. (B) A summary of three independent experiments, with a fully shaded square representing strong binding in all experiments, and a triangle representing weak binding in one or two membranes only. Spot 16 is not included as it is a blank dot.

Figure 7 Summary of effects of SDS on ataxin-3 aggregation
Schematic summarizing the effects of micellar and non-micellar SDS on both stages of ataxin-3 aggregation.
Figure S1. Far-UV CD spectra of ataxin-3(Q15) and the Josephin domain during aggregation in the presence of SDS. Protein aliquots were taken from a fibrillogenesis time course assay and the far-UV CD spectra determined. For each indicated SDS concentration aliquots were taken at times of 0 hr (black), 4 hr (green), 46 hr (red) and 100 hr (blue).

Figure S2. Morphology of fibrils formed by ataxin-3(Q15) and Josephin with SDS. Transmission Electron Microscopy of ataxin-3(Q15) with 0 mM (A), 1 mM (B) and 5 mM SDS (C), and Josephin domain with 0 mM (D), 1 mM (E) and 5 mM (F) SDS. Samples after 100 hr incubation were negatively stained using 1 % (w/v) uranyl acetate. Scale bars represent 200 nm.
Figure 1

A

Ataxin-3(Q64)

B

Ataxin-3(Q15)

C

Josephin
Figure 2

A

Ataxin-3(Q64)

B

Ataxin-3(Q15)

C

Josephin
Figure 3

A

<table>
<thead>
<tr>
<th>[SDS] (mM)</th>
<th>0</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>QBPI</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

B

![Relative Density vs. Time graph]
Figure 4

A. 0 mM SDS

B. 1 mM SDS

C. 5 mM SDS

D. 10 mM SDS
Figure 5

Effects of SDS on ataxin-3
Figure 6

A

(vi) (v)

B

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sample Type</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ataxin-3(Q64)</td>
<td>24 hr Incubation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ataxin-3(Q64)</td>
<td>Endpoint Fibril</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Josephin</td>
<td>70 hr Incubation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Josephin</td>
<td>Endpoint Fibril</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SpA</td>
<td>Monomer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SpA(Q52)</td>
<td>Endpoint Fibril</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 7

1 mM SDS accelerates stage 1, but slows formation of stage 2

Micellar SDS increases helicity of the monomer

5 mM SDS causes formation of SDS-insoluble amorphous aggregates

More Helical Monomer

Amorphous Aggregate
Supplementary Data

Figure S1

A. Josephin

B. Josephin + 1mM SDS

C. Josephin + 5 mM SDS

D. Josephin + 10mM SDS

E. Ataxin-3(Q15)

F. Ataxin-3(Q15) + 1mM SDS

G. Ataxin-3(Q15) + 5 mM SDS

H. Ataxin-3(Q15) + 10mM SDS
Figure S2

Effects of SDS on ataxin-3

Ataxin-3(Q15) and Josephin under different SDS concentrations:
- **0 mM SDS**
- **1 mM SDS**
- **5 mM SDS**

A 0 mM SDS  B 1 mM SDS  C 5 mM SDS

D 0 mM SDS  E 1 mM SDS  F 5 mM SDS
6.1 Role of Flanking Domains in PolyQ Aggregation

The mechanism of toxicity within the amyloid diseases has increasingly been linked to intermediate species formed along the aggregation pathway. Thus there is great interest in the characterization of conformational changes which occur during aggregation. The amyloid diseases are often associated with an amyloidogenic mutation, as exemplified by the polyQ diseases which have a strong correlation between polyQ tract length and disease pathology (269). The only sequence homology between the nine polyQ proteins is the presence of an expanded polyQ tract. Consequently, the host protein dictates factors which vary between polyQ diseases such as the localization of protein deposits and neurodegeneration, in addition to modulating disease progression (179,196,270,271).
Increasingly, domains which flank the polyQ tract have been demonstrated to modulate the aggregation of numerous polyQ proteins (164,191,193,194,196), yet the mechanism via which these flanking domains act remains poorly understood.

A new model was proposed at the commencement of my study whereby the aggregation of pathogenic length ataxin-3 occurs via a multi-stage mechanism involving both the flanking Josephin domain as well as the polyQ tract (191). However Ellisdon et al. did not conclusively demonstrate whether it was the Josephin domain or the C-terminal region surrounding the polyQ tract involved within the first stage of aggregation (192). This thesis has focused to a large degree on the characterization of the conformational changes occurring in the Josephin domain both when isolated and when in the context of the full length protein.

Decreased thermodynamic stability of amyloidogenic proteins is often linked to increased aggregation rates and vice versa (73,77,272,273). This link was exploited in the first part of this study to confirm that the initial stage of the ataxin-3 aggregation pathway is mediated by interactions of the Josephin domain (Ch 3). This study substantiated the role of the Josephin domain within the two-stage aggregation mechanism for ataxin-3, as initially suggested (191). The involvement of the flanking domain in ataxin-3 aggregation is not unique within the polyQ proteins. Concurrent to this study, multi-domain aggregation mechanisms have also been identified for huntingtin exon-1, and the model polyQ protein CRABP I-htt exon-1 (8,204). A multi-domain aggregation mechanism has also been demonstrated for a number of non-polyQ proteins (210-212). Taken together, these studies highlight the importance of characterizing the flanking domains of polyQ proteins, and determining their influence on polyQ protein aggregation.
Different types of fibrils are formed along the multi-stage aggregation pathway of ataxin-3 (191). The conformational transition from stage 1 fibrils to stage 2 fibrils has not been characterized, and it appears that under different conditions there are a number of alternative pathways via which ataxin-3 forms end-stage morphologically-similar fibrils (Fig 6.1). With 1 mM SDS present, ataxin-3 showed an increase in the rate of stage 1 fibril formation yet a decrease in the rate of stage 2 fibril formation (Ch 5). In an alternate pathway, the stabilized ataxin-3(Q64) R103G S81A mutant showed a similar midpoint for both stage 1 and stage 2 fibril formation (Ch 3). Both these results suggest that whilst linked, the formation of stage 2 fibrils is not solely dependent upon the rate of stage 1 fibril formation. In both these pathways however, in order to transition from fibrils mediated via Josephin domain interactions to fibrils formed through polyQ interactions a structural rearrangement must occur within the fibril. This is not unprecedented, with structural rearrangements within a fibril reported for other amyloidogenic proteins (274-276).

A third pathway was characterized in the presence of a stoichiometric excess of αB-crystallin, where stage 2 ataxin-3 fibrils formed without significant formation of stage 1 fibrils, albeit at a slower rate (App 1). The formation of morphologically similar fibrils under these conditions suggests that the Josephin domain is not involved in the core of the stage 2 fibril. Indeed, the presence of flanking domains along the fibril may explain the atypical morphology of the endpoint fibrils formed by polyQ proteins in comparison to most amyloid fibrils (Ch 5)(11,191). It is possible the core fibril structure involves just the polyQ tract, as structural studies using polyQ peptides demonstrate a core cross-β fibril structure can be formed by the polyQ tract alone (32). Proteolysis studies to determine the
core of both the stage 1 and stage 2 type fibrils will provide great insight into how the Josephin domain fits within the architecture of the fibril.

Figure 6.1 Alternative Aggregation Pathways of Ataxin-3. Under native conditions in vitro, ataxin-3 shows the sequential formation of stage 1 then stage 2 fibrils mediated via interactions of the Josephin domain (red arrows) and polyQ tract (green arrows) respectively. Potential arrangements of the Josephin domain and polyQ tract within these fibrils are shown in the grey box. Stage 2 fibrils may involve either just the polyQ tract or the polyQ tract and part of the Josephin domain within the core cross-β structure of the fibril. The alternative pathways do not all show the sequential nature of formation of stage 1 and 2 fibrils seen under native conditions in vitro, whilst the addition of 5 mM SDS demonstrates that ataxin-3 is also able to form aggregates of a non-fibrillar nature.
6.2 Relationship Between the Folding and Aggregation Pathways of Ataxin-3

The folding and aggregation pathways are intrinsically linked, and the relationship between them differs for amyloidogenic proteins. A natively folded protein typically must unfold in order to form an amyloidogenic partially folded intermediate and there have been two models proposed for this mechanism for ataxin-3 (Fig 6.2). Initially, it was considered that ataxin-3 follows the conformational change hypothesis and aggregates via an intermediate along the unfolding pathway (163,201). However a later study proposed that the amyloidogenic intermediate is formed from the native state, separate to the folding pathway (69). This present study has distinguished between these two previously proposed mechanisms and supported the conformational change model whereby the initial misfolding event in ataxin-3 aggregation occurs from a partially unfolded intermediate along the folding pathway (Ch 3).

**Figure 6.2 The Folding and Aggregation Pathways of Ataxin-3.** The two models which have been proposed for ataxin-3 aggregation are depicted. This study has demonstrated that ataxin-3 aggregates according to the conformational change model of aggregation.
Within the family of natively folded amyloid proteins both mechanisms of initial conformational change have been observed. A number of proteins such as transthyretin and lysozyme aggregate according to the conformational change hypothesis of aggregation (68,277), however a growing number of proteins including acylphosphatase, HypF-N and insulin aggregate from an off-pathway intermediate (40,70,278). In addition, studies on the prion protein have shown that different mutations can result in both models of aggregation (279-281). This highlights the importance of using mutations which are spread throughout the tertiary structure of the protein as was done in this study (Ch 3) and thus less likely to push a protein down a specific pathway due to changing other characteristics of the protein.

In a conformational change aggregation model, to increase the propensity of a folded protein to form a partially unfolded intermediate there is typically a change in protein stability or dynamics (74). Previous studies have demonstrated that an expanded polyQ tract does not change the stability of ataxin-3 (69), which thus implicates a potential role for dynamics in initiation of aggregation. The complexity of ataxin-3 aggregation is again highlighted when it is considered that of the two amyloidogenic domains in ataxin-3, the Josephin domain is natively folded whilst the polyQ tract is predominantly unordered (282). Therefore a complete analysis of the early events of ataxin-3 misfolding requires the study of the dynamics of both the Josephin domain and the polyQ tract.

The Josephin domain has been demonstrated to be highly dynamic using hydrogen exchange mass spectrometry (Ch 4). This is consistent with the high amyloidogenicity of this domain under native conditions (206,208). Molecular dynamics simulations have been completed on Josephin in the context of ubiquitin binding (169). These studies correlate with previous NMR dynamics experiments proposing that the helical hairpin of Josephin is
more dynamic than the rest of the domain (166,169). It is yet to be determined which regions of Josephin are important in mediating aggregation, and thus whether the dynamic motion of the helical hairpin may play a role.

The Josephin dynamics study found that native Josephin exists in solution in equilibrium with multiple partially unfolded forms (PUF) (Ch 4, Fig 6.2). These PUF appear to be only slightly unfolded, and interconverting on a timescale which means they are not amenable to detection by NMR spectroscopy. Despite identification of this structural heterogeneity of the folded Josephin domain, connections between these conformations and function or aggregation could not be identified in this study. The presence of PUF in equilibrium with their native state has been identified in a number of proteins, and in some cases, for example the GroEL apical domain, is linked to protein function (248,251,253). To further investigate the specific conformations of the PUF in both Josephin and ataxin-3, single-molecule fluorescence studies could be completed as it is another technique with the capacity to detect multiple conformations in solution.

Having developed a system to determine the hydrogen exchange kinetics of Josephin, future studies can now determine the dynamics of the full length protein and the impact of an expanded polyQ tract. A number of previous studies have focused on the dynamics of an isolated polyQ tract, and found the polyQ tract exists as a heterogeneous ensemble of fluctuating conformations which are in a collapsed state (213,267,283-286). In addition, crystal structures of huntingtin exon-1 suggested that the conformation of the polyQ tract is influenced by flanking regions (285). Taken together, this suggests that for polyQ proteins, determining the impact of flanking domains on polyQ dynamics may provide some insight into the initial conformational change events involved in aggregation.
6.3 Multi-domain Misfolding and PolyQ Disease

The mechanism of toxicity is currently considered to be linked to the formation of prefibrillar oligomers for both polyQ and other amyloid diseases (42,53,139). Together with the strong link between polyQ tract length and toxicity, this suggests there must be a difference in the oligomers formed by pathogenic and non-pathogenic length ataxin-3, regardless of both involving Josephin domain interactions. The structure of oligomers formed by a polyQ protein containing pathogenic and non-pathogenic length polyQ tracts has not been determined. Aggregates have not been detected within animal or cell models with overexpression of non-pathogenic length polyQ proteins (141,207,287). This suggests that either there is no formation of aggregates \textit{in vivo} with non-pathogenic length polyQ proteins, or alternatively non-toxic aggregates form which are cleared from cells.

An assumption of the multi-domain misfolding mechanism is that there is no cleavage of ataxin-3 before aggregation. Conflicting evidence exists as to the presence of ataxin-3 fragments in disease inclusions (141,174-176). However in a study where fragments were detected in diseased brain, there was still an equal proportion of full length ataxin-3 (141). Thus an unsolved question is whether it is the fragment or the full length ataxin-3 responsible for initiation of aggregation and subsequent fibril formation in disease. Yet even if full length ataxin-3 is involved in aggregation, the Josephin domain may not be directly involved in aggregation \textit{in vivo} due to its interaction with other cellular components. Recently, it was demonstrated that non-polyQ domains were involved in nuclear localization of ataxin-3 after a heat shock or oxidative stress was applied (288). With inclusions targeted to the nucleus resulting in increased toxicity for a number of
polyQ proteins (104-108), this is an example of an indirect mechanism via which Josephin may contribute to toxicity.

Despite the currently poor understanding of the structural species forming within the cell, a well characterized in vitro aggregation mechanism provides many opportunities to characterize the interaction between cellular components and aggregation. For example, characterization of the effects of αB-crystallin on aggregation in vitro found that the inhibition of aggregation by this small heat shock protein occurred solely through an interaction with the Josephin domain (App 1). In addition, ataxin-3 with either a pathogenic or non-pathogenic length polyQ tract was found to interact with phospholipids only when fibrillar in structure (Ch 5). In vitro techniques also provide the tools to investigate effects of potential inhibitors of polyQ aggregation such as the polyphenol (-)-epi-gallocatechine gallate (EGCG), which inhibits aggregation of a number of amyloid proteins (289,290). Thus characterization of the in vitro aggregation mechanism provides a platform to both investigate the details of cellular interactions, and determine the effects of potential aggregation inhibitors.

6.4 General Conclusions

The aggregation pathway of multiple polyQ proteins has been demonstrated to be modulated by flanking domains both in cell models and in vitro (8,164,191,193,196). Whilst there is an overriding correlation between disease severity and polyQ tract length, the conformational changes and mechanisms of toxicity in vivo remains unclear for the nine polyQ diseases. There is a trend towards the flanking domain of the polyQ proteins being intrinsically fibrillogenic, as observed in the three polyQ proteins characterized in vitro
Thus investigation of the involvement of flanking domains in both \textit{in vitro} and \textit{in vivo} aggregation is important to determine their role in modulating disease progression.

This study has contributed to this area by the further characterization of the multi-stage aggregation pathway of ataxin-3, in particular the involvement of the flanking Josephin domain. The folding and fibrillogenic pathways and link between them has been described in more detail, with the identification of partially unfolded conformations of Josephin which exist in equilibrium with the natively folded states. In addition, this study has determined that ataxin-3 follows the conformational change hypothesis with aggregation initiating from a partially unfolded intermediate which is on the folding pathway, and then the first stage of aggregation occurring via interactions of the Josephin domain. Analysis across the polyQ disease proteins in the future will allow identification of factors which dictate the characteristics of each disease.

The ability of ataxin-3 to aggregate via alternate pathways under conditions which may be encountered within the cell has also been highlighted by this study. Identification of the conformational space accessible to the polyQ proteins provides insight into the structural changes which occur \textit{in vivo} under both physiological and stress conditions. Furthermore, characterization of the structural changes \textit{in vitro} therefore provides the opportunity to screen inhibitors targeting the specific conformational change events of aggregation.
Chapter 7

REFERENCE LIST


Reference List


Referece List


244. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) Primary structure effects on peptide group hydrogen exchange *Proteins* 17, 75-86


254. Konermann, L., and Douglas, D. J. (1997) Acid-induced unfolding of cytochrome c at different methanol concentrations: electrospray ionization mass spectrometry specifically monitors changes in the tertiary structure Biochemistry 36, 12296-12302


molecules form collapsed structures in aqueous solutions *Proc Natl Acad Sci U S A* **103**, 16764-16769


APPENDICES

Appendix 1:
Small heat shock proteins interact with a flanking domain to suppress polyglutamine aggregation

Appendix 2:
Multi-domain misfolding: understanding the aggregation pathway of polyglutamine proteins
Small heat-shock proteins interact with a flanking domain to suppress polyglutamine aggregation

Amy L. Robertson1, Stephen J. Headey2, Helen M. Saunders3, Heath Ecroyd4, Martin J. Scanlon5, John A. Carver1, and Stephen P. Bottomley1,6

1Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria, 3168, Australia; 2Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria, 3052, Australia; 3School of Biological Sciences, University of Wollongong, Wollongong, New South Wales, 2522, Australia; and 4School of Chemistry and Physics, The University of Adelaide, Adelaide, South Australia, 5005, Australia

Edited by George H. Lorimer, University of Maryland, College Park, MD, and approved April 15, 2010 (received for review December 20, 2009)

Small heat-shock proteins (sHsps) are molecular chaperones that play an important protective role against cellular protein misfolding by interacting with partially unfolded proteins on their off-folding pathway, preventing their aggregation. Polyglutamine (polyQ) repeat expansion leads to the formation of fibrillar protein aggregates and neuronal cell death in nine diseases, including Huntington disease and the spinocerebellar ataxias (SCAs). There is evidence that sHsps have a role in suppression of polyQ-induced neurodegeneration; for example, the sHsp alphab-crystallin (αB-c) has been identified as a suppressor of SCA3 toxicity in a Drosophila model. However, the molecular mechanism for this suppression is unknown. In this study we tested the ability of αB-c to suppress the aggregation of a polyQ protein. We found that αB-c does not inhibit the formation of SDS-insoluble polyQ fibrils. We further tested the effect of αB-c on the aggregation of ataxin-3, a polyQ protein that aggregates via a two-stage aggregation mechanism. The first stage involves association of the N-terminal Josephin domain followed by polyQ-mediated interactions and the formation of SDS-resistant mature fibrils. Our data show that αB-c potently inhibits the first stage of ataxin-3 aggregation; however, the second polyQ-dependent stage can still proceed. By using NMR spectroscopy, we have determined that αB-c interacts with an extensive region on the surface of the Josephin domain. These data provide an example of a domain/region flanking an amyloidogenic sequence that has a critical role in modulating aggregation of a polypeptide and plays a role in the interaction with molecular chaperones to prevent this aggregation.

fibrillogenesis | α-crystallin

S

mall heat-shock proteins (sHsps) are important in the maintenance of cellular homeostasis. sHsps are induced under stress conditions and interact with partially unfolded proteins on their off-folding pathways, thereby providing a protective mechanism against protein misfolding and aggregation (1). They are present in many species, and 10 have been identified in the human proteome (see reviews in refs. 1 and 2). Alpha-crystallin (α-crystallin) is the most well characterized sHsp. The two isoforms of α-crystallin, αA-crystallin (αA-c) and αB-crystallin (αB-c), are a large component of the human eye lens, and αB-c is expressed in many other cell types including neurons (3, 4). Monomeric αB-c is approximately 20 kDa in mass, and under native conditions it forms a heterogeneous array of multimeric complexes ranging from 160 to 1,000 kDa in mass (2). The activity of αB-c is thought to involve hydrophobic interactions between αB-c and the partially folded protein. Under stress conditions, such as elevated temperature, the activity of αB-c is enhanced because of a conformational change within the multimeric complex (5, 6). Cellular stress is induced in several neurodegenerative diseases, such as Alzheimer’s disease (7), prion diseases (8), Parkinson disease (4), and Huntington disease (HD) (4), and there is evidence of increased αB-c activity (4). In addition, αB-c reduces neurotoxicity in a number of disease models (9, 10). In vitro the interaction between αB-c and a number of aggregating proteins has been characterized. Most substrates aggregate through a nucleation-dependent kinetic mechanism; however, the mode of interaction with αB-c appears to be dependent upon the specific target protein and the type of aggregation—i.e., amorphous or fibrillar (11–13). Complex stability is influenced by properties of the target protein, for example, the fibril-forming proteins apolipoprotein-CII (11) and amyloid-β (14) transiently interact with αB-c whereas α-synuclein (6) and e-casein (15) form stable noncovalent complexes. Determining the mechanism by which αB-c interacts with specific aggregation-prone substrates is therefore important and may provide insight into the misfolding pathways leading to toxicity.

Polyglutamine (polyQ) expansion leads to the formation of fibrillar protein aggregates and neuronal cell death in nine diseases, including HD and the spinocerebellar ataxias (SCAs) (1, 12, 13). PolyQ peptides and proteins form fibrillar aggregates by a nucleation-dependent mechanism that is initiated by a monomeric nucleus (16–18). Recent evidence suggests that polyQ protein misfolding involves not only the polyQ tract but other aggregation-prone regions within the polyQ proteins (19–21). There is experimental evidence that the proteins ataxin-3 (in SCA3), ataxin-1 (in SCA1), and huntingtin (in HD) form fibrillar aggregates by a multidomain misfolding mechanism in which non-polyQ regions of the protein self-associate before the polyQ tract (19, 20, 22). Our laboratory has previously shown that ataxin-3 has a minimal two-stage aggregation mechanism (19). The first stage involves aggregation of the globular N-terminal Josephin domain that precedes the self-association of expanded polyQ segments (19). sHsp overexpression decreases neuronal toxicity in HD by nonchaperone mechanisms, including suppression of reactive oxygen species (23) and stimulation of autophagy (24). In disease models overexpression of αB-c modulates toxicity in a context-dependent manner. αB-c did not alter the HD phenotype in a Drosophila model (25); however, it suppressed SCA3 toxicity in a Drosophila model (9). Enhanced suppression was observed when αB-c was coexpressed with full-length ataxin-3 in comparison with a C-terminal fragment not containing the Josephin domain (9). These data suggest that non-polyQ protein regions influence sHsp chaperone activity.

In this study we tested the chaperone ability of αB-c for a polyQ protein (SpAcQ52) that forms fibrillar aggregates by a mechanism involving only the polyQ region. We found that αB-c and a number of aggregating proteins has been characterized. Most substrates aggregate through a nucleation-dependent kinetic mechanism; however, the mode of interaction with αB-c appears to be dependent upon the specific target protein and the type of aggregation—i.e., amorphous or fibrillar (11–13). Complex stability is influenced by properties of the target protein, for example, the fibril-forming proteins apolipoprotein-CII (11) and amyloid-β (14) transiently interact with αB-c whereas α-synuclein (6) and e-casein (15) form stable noncovalent complexes. Determining the mechanism by which αB-c interacts with specific aggregation-prone substrates is therefore important and may provide insight into the misfolding pathways leading to toxicity.

Polyglutamine (polyQ) expansion leads to the formation of fibrillar protein aggregates and neuronal cell death in nine diseases, including HD and the spinocerebellar ataxias (SCAs). There is evidence that sHsps have a role in suppression of polyQ-induced neurodegeneration; for example, the sHsp alphab-crystallin (αB-c) has been identified as a suppressor of SCA3 toxicity in a Drosophila model. However, the molecular mechanism for this suppression is unknown. In this study we tested the ability of αB-c to suppress the aggregation of a polyQ protein. We found that αB-c does not inhibit the formation of SDS-insoluble polyQ fibrils. We further tested the effect of αB-c on the aggregation of ataxin-3, a polyQ protein that aggregates via a two-stage aggregation mechanism. The first stage involves association of the N-terminal Josephin domain followed by polyQ-mediated interactions and the formation of SDS-resistant mature fibrils. Our data show that αB-c potently inhibits the first stage of ataxin-3 aggregation; however, the second polyQ-dependent stage can still proceed. By using NMR spectroscopy, we have determined that αB-c interacts with an extensive region on the surface of the Josephin domain. These data provide an example of a domain/region flanking an amyloidogenic sequence that has a critical role in modulating aggregation of a polypeptide and plays a role in the interaction with molecular chaperones to prevent this aggregation.
αB-c does not inhibit fibrillogenesis in this system. We then tested the effect of αB-c on the aggregation of ataxin-3 harboring a pathological length polyQ tract [at3(Q64)] and a truncated variant comprising only the Josephin domain. Our data show that αB-c potently inhibits the initial Josephin-dependent stage of ataxin-3 aggregation. By using NMR spectroscopy, we determined that αB-c interacts with specific regions of Josephin. These data provide an example and characterization of a domain/region flanking an amyloidogenic sequence that has a critical role in modulation of aggregation by interacting with molecular chaperones.

**Results**

**αB-c Does Not Inhibit PolyQ-Mediated Aggregation.** We first tested whether αB-c can prevent polyQ tract-mediated aggregation. We incubated αB-c with a polyQ protein (SpAcQ52), which forms SDS-insoluble aggregates by a mechanism involving only the polyQ tract (26). The rate at which SpAcQ52 forms ThT-reactive aggregates does not change when incubated with increasing concentrations of αB-c (Fig. 1A). It has been previously shown that molecular chaperones can redirect the aggregation pathway of a protein from amyloid to amorphous aggregation (27, 28). Amorphous aggregates can be differentiated from amyloid by their sensitivity to SDS and appearance when visualized by transmission electron microscopy (TEM). So we tested whether SpAcQ52 incubated with αB-c forms SDS-insoluble aggregates by using a membrane filter-trap assay (29). These data demonstrate that the aggregates formed in the presence or absence of αB-c are insoluble in SDS (Fig. 1B). In addition, TEM analysis of the aggregates indicated that αB-c does not affect their fibrillar morphology. Fibers ranging from 15 to 8 nm in width and 300 to 400 nm in length were observed in both the presence and absence of αB-c (Fig. 1C). Together, these data demonstrate that αB-c does not inhibit or modify the fibrillar aggregation of SpAcQ52.

![Normalized ThT fluorescence vs. time of at3(Q64)](image)

**Fig. 1.** The effect of αB-c on polyQ aggregation. (A) SpAcQ52 aggregation in the presence of increasing concentrations of αB-c monitored by ThT fluorescence. (B) Analysis of the SDS solubility of endpoint aggregates (100 h) in the presence and absence of αB-c. (C and D) TEM analysis of SpAcQ52 aggregates in the absence (C) and presence of αB-c. Ratios of SpAcQ52∶αB-c are given on a molar basis.

**αB-c Inhibits the Josephin-Dependent Stage of the Two-Stage Ataxin-3 Aggregation Pathway.** We have previously reported that full-length ataxin-3 forms aggregates by a two-stage mechanism (18). The aggregates formed in both stages are ThT-reactive, and in the absence of glycerol the two stages are characterized by a biphasic time course (Fig. 2A). The first stage is dominated by interactions of the N-terminal Josephin domain, which is an independently aggregation-prone domain (30). Aggregates formed during the first stage appear as short curvilinear structures and can be solubilized by SDS (19). These aggregates are stabilized in the second stage by interactions of the polyQ regions. Stage-two aggregates can be distinguished from stage-one aggregates because they are SDS-insoluble and morphologically appear as larger fibrillar structures. We determined the effect of αB-c on the two-stage aggregation pathway of at3(Q64). αB-c significantly slows the formation of ThT-reactive at3(Q64) aggregates in a concentration-dependent manner (Fig. 2A and Table 1).

The Josephin domain forms short aggregates with a ribbon-like morphology (30); however, the aggregation kinetics have not been described. Isolated Josephin forms ThT-reactive aggregates at a much slower rate than at3(Q64) (Fig. 2B and Table 1). The half-life of the first ThT-reactive phase of at3(Q64) aggregation represents the formation of Josephin-dependent fibrils (19). The half-life shifts from approximately 9 h in the full-length protein to around 60 h for the isolated Josephin domain (Table 1). These data indicate that the C-terminal polyQ tail in the context of the full-length protein accelerates Josephin domain aggregation. αB-c has potent inhibitory activity on the aggregation of Josephin (Fig. 2B and Table 1), compared to the full-length at3(Q64) variant. Thus, an equimolar concentration of αB-c completely suppresses Josephin aggregation, but at3(Q64) still forms some ThT-reactive aggregates, even when αB-c is present in a 2∶1 molar excess. The reaction time course is too long to determine if the biphasic ThT pattern is altered at stoichiometric concentrations of αB-c. The potent inhibitory effect of αB-c on Josephin aggregation, combined with the data in Fig. 1A, shows that αB-c does not inhibit polyQ-dependent aggregation, suggests that the slow ThT phase in the presence of high αB-c concentrations is dominated by polyQ intermolecular interactions.

We analyzed the Josephin and at3(Q64) fibrils by TEM (Fig. 3A). Josephin forms short curvilinear fibrils, 5–10 nm in width and 10–100 nm long, when incubated alone. However,
Table 1. Half-lives (in hours) of Ataxin-3(Q64) aggregation followed by ThT and SDS insolvability in the presence and absence of α-B-c

<table>
<thead>
<tr>
<th></th>
<th>ThT Stage 1</th>
<th>ThT Stage 2</th>
<th>SDS insolvibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ataxin-3(Q64)</td>
<td>9.3 ± 0.2</td>
<td>31.7 ± 0.9</td>
<td>25.5 ± 1.1</td>
</tr>
<tr>
<td>Ataxin-3(Q64) 10:1 α-B-c</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Josephin</td>
<td>19.7 ± 0.6</td>
<td>49 ± 4</td>
<td>45.1 ± 1.3</td>
</tr>
<tr>
<td>Josephin 10:1 α-B-c</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Josephin 1:1 α-B-c</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

*Aggregation incomplete after 300 h; therefore, data could not be fitted.
††An equimolar amount of α-B-c completely inhibited Josephin aggregation over 300 h.

Table 1: Half-lives (in hours) of Ataxin-3(Q64) aggregation followed by ThT and SDS insolvability in the presence and absence of α-B-c.

By a twofold increase in half-life for at3(Q64) aggregation occurring in the presence of stoichiometric concentrations of α-B-c [at3(Q64) 10:1 α-B-c] (Fig. 3B and Table 1). The inhibitory effect of α-B-c is more pronounced at stoichiometric concentrations where SDS-insoluble fibrils were still formed but at much lower amounts relative to at3(Q64) incubated alone. The equimolar data could not be fit to the classic nucleation-dependent model because aggregation is incomplete over the 300-h time scale.

The Presence of α-B-c Retains Ataxin-3 in a Monomeric Conformation. α-B-c exists as a dynamic oligomeric complex (2). Depending on the target protein, it can irreversibly or transiently bind to target proteins, thereby inhibiting aggregation. To determine the stability of the α-B-c-at3 association we performed size-exclusion chromatography (SEC) of Josephin and at3(Q64) incubated with equimolar concentrations of α-B-c (Fig. 4).

Monomeric Josephin eluted from a Superose 6 column at approximately 18.9 mL. After 24-h incubation, Josephin formed insoluble aggregates that were removed by centrifugation prior to injection onto the column; therefore, the monomeric protein peak eluting at 18.9 mL was diminished (Fig. 4, Top Left). α-B-c displayed a broad peak centered around 14 mL, consistent with α-B-c being present as a polydisperse oligomer. In the presence of α-B-c, most of the Josephin is retained in a monomeric conformation at 49 h, with some recruited into the α-B-c complex, observed by a distinct shift and broadening of the α-B-c peak (Fig. 4, Top Right). We performed SDS-PAGE to examine the composition of the peaks eluting at various times in the SEC experiments (Fig. S1A). A small amount of Josephin coelutes with the multicentric α-B-c complex after a 48-h incubation period, and this is consistent with the shift and broadening of the peak corresponding to multimeric α-B-c (Fig. 4).

At3(Q64) eluted as a single peak with a retention volume of approximately 17.8 mL. After 8 h, most of the at3(Q64) had aggregated, indicated by the presence of a large peak in the void volume (approximately 8 mL) of the Superose 6 column (Fig. 4, Bottom Left). After 8 h in the presence of α-B-c, most of the at3(Q64) is retained in a monomeric conformation, with little change in the elution profile after this time (Fig. 4, Bottom Right). No at3(Q64) coeluted with the multimeric α-B-c (Fig. S1B), consistent with the observation that the α-B-c peak did not shift in the experiments with at3(Q64) (Fig. 4). The SEC data suggest that α-B-c does not form a stable complex with at3(Q64)/Josephin maintaining the proteins in a monomeric conformation.

Addition of α-B-c Affects Josephin/At3(Q64) Fibril Propagation. The observation that α-B-c retains Josephin and at3(Q64) in a monomeric conformation suggests that α-B-c inhibits the conformational changes involved in nucleation of aggregation. α-B-c has also been shown to inhibit fibril propagation in a number of systems (6, 11, 14, 27). We tested the effect of α-B-c on Josephin and at3(Q64) fibril propagation by adding α-B-c at various intervals after the commencement of aggregation.

When α-B-c was added to isolated Josephin at 48 and 96 h, there was no further increase in ThT intensity, suggesting that α-B-c prevents the further propagation of Josephin fibrils (Fig. S2A). Similar behavior is observed when α-B-c is added at 4 and 8 h along the aggregation pathway of α-synuclein (6). When α-B-c was added to the at3(Q64) reactions at various intervals, there was a flattening of the curve for a period of time and then a second slower aggregation phase (Fig. S2B). Because of the absence of this phase in the isolated Josephin reactions, we predict that the second phase is because of polyQ-dependent aggregation. Interestingly, the length of the flattened period was longer when α-B-c was added at earlier time points in comparison to later time points, supporting the hypothesis that the rate at which polyQ interactions occurs is accelerated by the initial formation of stage-one Josephin-dependent fibrils.
αB-c Interacts with an Extensive Region of Josephin. The ability of αB-c to inhibit ataxin-3 aggregation appears to be derived from inhibition of the Josephin-dependent stage (stage one) of the multidomain aggregation pathway. We used NMR spectroscopy to probe the molecular details of the Josephin:αB-c interaction surface. The NMR data show that αB-c interacts with an extensive region of Josephin (Fig. 5). A number of backbone amide cross-peaks in the 15N-HSQC (heteronuclear single quantum coherence) spectrum of Josephin were significantly broadened upon titration of αB-c up to a fivefold stoichiometric excess (Fig. S3). The broadening is consistent with exchange between the free and αB-c bound states of Josephin. Exchange broadening occurs when the time scale of the interaction is commensurate with the difference in NMR frequencies between the bound and free states, which is consistent with the transient interaction mechanism observed by SEC (Fig. 4). The Josephin residues perturbed to the largest degree were Q16, L19, N21, E26, S29, S35, I36, I77, W87, Q100, C114, V123, G127, F131, S135, L137, T138, L148, L155, Q156, E158, I162, Q176, and R182. These residues are confined to the globular subdomain of Josephin and include the active site residue C114.

Because chemical shift perturbations can arise through indirect mechanisms (rather than a direct interaction between αB-c and Josephin), we used the paramagnetic relaxation agent Gd-diethylenetriamine pentaacetic acid-bismethylamide [Gd(DTPA-BMA)] to independently confirm the interaction surface (31, 32). The presence of Gd(DTPA-BMA) in solution causes paramagnetic relaxation enhancement, broadening the protein resonances. By performing these 15N-HSQC experiments in the presence and absence of αB-c, we could determine the αB-c interaction surface of Josephin, because the regions that interact with αB-c were partially shielded from paramagnetic relaxation (Fig. S4). Josephin residues that had >20% stronger signals in the presence of Gd(DTPA-BMA) and a twofold excess of αB-c were S3, G11, L13, Q24, F28, A49, E50, G51, G52, T54, E56, D57, R59, T60, F61, S66, D71, G73, F75, S76, E90, N115, K128, N132, G139, L142, F151, G159, D168, E173, L178, M180, and I181. The identity of these residues correlates well with those broadened in the αB-c-bound state of Josephin (see above, Fig. 5). When mapped to the Josephin structure all of the residues that appeared to be most shielded from the paramagnetic probe by the presence of αB-c were found to be surface-exposed, validating these regions as the αB-c binding interface.

Fig. 4. Ataxin-3 aggregation in the presence and absence of αB-c monitored by SEC. Samples of Josephin and At3(Q64) incubated in the presence and absence of αB-c were removed at indicated time points and analyzed by SEC using a Superose 6 column. In the samples containing αB-c, equimolar amounts of substrate: αB-c were used.

Fig. 5. αB-c interacts with specific regions of Josephin. (A and B) The residues perturbed upon addition of αB-c were mapped onto the Josephin domain (Protein Data Bank ID code 1YZB). Residues in intermediate exchange in the presence of αB-c (Red), residues shielded from broadening by Gd (DTPA-BMA) in the presence of αB-c (Yellow), and orange shows the overlapping residues. A and B show opposing faces of the Josephin structure. (C and D) The electrostatic surface of Josephin. Blue represents positive charge and red indicates negatively charged regions. C and D show opposing faces of the Josephin structure.

Fig. 6. A schematic model describing the effect of αB-c on at3 aggregation. Multimeric αB-c (sHsp) transiently binds to at3 molecules in which Josephin is in an aggregation-prone conformation (Jagg). The Jagg:αB-c complex is unstable, and monomeric Josephin is released from the sHsp.
The significant effect of αB-c on ataxon-3 aggregation is consistent with a *Drosophila* model in which overexpression of αB-c decreased SCA3-induced toxicity (9), suggesting that inhibition of aggregation by chaperones is relevant in vivo. Another *Drosophila* study demonstrated that chaperone-induced reduction of ataxon-3 in a monomeric conformation suppresses toxicity (44). Conversely, αB-c overexpression had no effect on the HD phenotype (25). Htt exon 1 has been proposed to have a similar two-stage aggregation mechanism compared to ataxon-3 (20). The alternative effects of αB-c in the ataxon-3 and Htt systems demonstrate how specific differences in multidomain misfolding pathways could lead to variable cellular responses.

**αB-c Transiently Interacts with MonomERIC Josephin.** The aggregation data clearly show that αB-c suppresses aggregation by associating with species on the Josephin aggregation pathway. Our SEC and NMR data show that inhibition mechanistically occurs by transient association of αB-c with a defined region on the Josephin surface that is partially coincident with the ubiquitin and HHR23B interacting surfaces (45–47). Interestingly, the perturbed residues are not all hydrophobic in character, and it is possible that some surface polar residues appear perturbed in our NMR studies because their amide groups are neighboring interacting hydrophobic groups. Whereas hydrophobic interactions are clearly important for chaperone action, it is also being increasingly recognized that substrate specificity is not purely driven by hydrophobic contacts (48–50).

The SEC data show that the proteins are retained in a monomeric conformation in the presence of αB-c (Fig. 4). These data suggest that the rate at which αB-c interacts with Josephin is more rapid than the rate of misfolding. The NMR data indicate that this is a low affinity interaction; therefore, during this interaction, most of the Josephin is in a monomeric uncomplexed state. These data are consistent with a mechanism whereby αB-c binds and sequesters the free pool of monomeric proteins, thereby preventing fibril growth (Fig. 6).

The interaction between αB-c and monomeric Josephin, encompassing a large region of the target protein, is also observed, in the interaction of αB-c with the fibril-forming proteins α-synuclein (27) and ε-casein (6).

**Conclusions**

Our data demonstrate that inhibition of a non-polyQ aggregation stage by a chaperone can significantly impede protein aggregation involved in polyQ diseases. These data support the two-stage model as a plausible scheme for ataxon-3 aggregation in vitro. Therapeutic up-regulation of chaperone activity may be beneficial in SCA3 and other polyQ diseases that have a multidomain aggregation mechanism. **Experimental Procedures**

**Preparation of Proteins.** SpAcQ52 and the ataxon-3 variants used in this study were expressed and purified as previously described (26, 51). αB-c was expressed and purified as described in ref. 52.

**Protein Aggregation Assays.** Aggregation assays were performed for SpAcQ52 and ataxon-3 variants at a protein concentration of 20 μM in Tris-buffered saline (TBS) (100 mM Tris, 80 mM NaCl, pH 7.4) containing 2 mM PMSF, 5 mM EDTA, and 15 mM β-mercaptoethanol. Variable concentrations of αB-c were added to the samples, ranging from 2 to 40 μM. Samples were incubated at 37 °C.

**ThT Measurements, Membrane Filter-Trap Assay, and TEM.** ThT measurements, the membrane filter-trap assay, and TEM analysis were performed as described in refs. 19 and 26.

**SEC.** SEC of at3(Q64) and Josephin in the presence and absence of equimolar concentrations of αB-c was performed on an Akta-
FPLC equipped with a Superose 6 10/30 GL column (Pharma-
cia). PI injection to the on-column, the samples were centrifuged for 5 min at 13,000 rpm to remove insoluble aggregated material. The absorbance at 214 nm was measured as proteins were eluted in TBS buffer at a constant flow rate of 0.5 mL/min.

26. Robertson AL, et al. (2008) The structural impact of a polyglutamine tract is location-
dependent. Biophys J 95(2):592–599.

NMR Spectroscopy. NMR methods are described in SI Text.

ACKNOWLEDGMENTS. This work was supported by the Australian Research Council and the National Health and Medical Research Council (NHMRC). S.P.B. is a NHMRC Research Fellow. H.E. was a NHMRC Peter Doherty Fellow.
E-mail: steve.bottomley@med.monash.edu.au

aggregation was clearly demonstrated by Nozaki et al which result in their multi-stage aggregation mechanisms. In this review, we will focus on the aggregation of proteins. In this review, we examine the recent data in support of this novel mechanism of protein aggregation.

**Keywords:** aggregation/ataxin-3/polyglutamine repeat/protein misfolding

**PolyQ aggregation and protein context**

The polyglutamine (polyQ) diseases consist of nine neurodegenerative diseases in which a polyQ tract expansion leads to protein misfolding and subsequent aggregation. Even when the causative proteins have the same length polyQ tract, there are differences in the severity and age of disease onset which implicate the polyQ flanking sequences as modulators of disease. Recent studies on the polyQ proteins ataxin-1, ataxin-3 and huntingtin exon-1 have shown that the flanking domains have an intrinsic ability to aggregate. This complex behavior leads to a multi-stage aggregation mechanism which we have termed multi-domain misfolding. In multi-domain misfolding, a flanking domain to the polyQ tract plays an early role in aggregation, before the contribution of the polyQ tract. A number of factors including the stability, dynamics and amyloidogenicity of the flanking domain modulate the impact on polyQ tract aggregation as well as any protein–protein interactions it undertakes. In this review, we examine the recent data in support of this novel mechanism of protein aggregation.

The polyglutamine (polyQ) diseases consist of nine neurodegenerative diseases (Table I) whereby mutation, in the form of a CAG repeat expansion, results in neurodegenerative disease. An inverse exponential correlation exists between glutamine repeat length and age of disease onset and severity (Gaykema and MacDonald, 2000). The same polyQ expansion in different contexts results in varying toxicity and disease pathology, and thus many studies have investigated the role that protein context plays in modulating the extrinsic toxicity of the polyQ sequence. Recently, the flanking domains of three polyQ proteins ataxin-1, ataxin-3 and huntingtin (Fig. 1) have been shown, in vitro, to have intrinsic aggregation propensities which modulate the aggregation of the full-length proteins. In this review, we will focus on the aggregation of these proteins and discuss multi-domain misfolding events which result in their multi-stage aggregation mechanisms.

The ability of polyQ flanking sequences to modulate aggregation was clearly demonstrated by Nozaki et al. (2001) who saw differing aggregation rates when the surrounding 17 amino acids of ataxin-2, ataxin-3, huntingtin and atrophin were added to the same length polyQ tract. In addition to this, they found that mutating four hydrophobic residues in the upstream flanking sequence of ataxin-2 to charged amino acids decreased the aggregation rate, further highlighting the impact of sequence specificity upon aggregation.

Subsequently, a number of studies have investigated the flanking regions to the polyQ tract within huntingtin exon-1. Through the use of peptide studies, the poly proline (polyP) region immediately C-terminal to the polyQ tract in huntingtin (Fig. 1) has been shown to slow aggregation and decrease the stability of aggregates compared with a polyQ peptide alone (Bhattacharyya et al., 2006). This study also found that the protective effects of the polyP region are directionally, having no impact on aggregation when N-terminal to the polyQ tract. The mechanism of inhibition was investigated by Darnell et al. (2007) who show that the polyP region exists as a PPII helix (proline type II helix), inducing a PPII-like structure in the polyQ tract that opposed the aggregation-prone β-sheet conformation formed by this region. In addition to the in vitro study using polyQ peptides, it was shown that the polyP sequence decreases toxicity in a yeast huntingtin exon-1 model system (Duennwald et al., 2006). Duennwald et al. also investigated the impact of two commonly used tags and show that the soluble GST domain prevents toxicity whereas the negatively charged FLAG tag increases toxicity. Therefore, the polyQ surrounding sequence within huntingtin exon-1 has a significant modulating effect on huntingtin exon-1 aggregation kinetics.

The flanking domain also plays a role in determining the molecular species formed during aggregation (Fig. 2). Although pathogenic length pure polyQ peptides only form two species during aggregation, monomer and fibril (Fig. 2) (Chen et al., 2002), in the context of the disease, protein intermediate species such as spherical oligomers and curvilinear prototribilis are also formed (Gales et al., 2005; Thakur et al., 2009). Similar to a number of other amyloidogenic proteins, polyQ oligomers have also been shown to be toxic (Bucciantini et al., 2002; Li et al., 2007; Takahashi et al., 2008), showing the importance of flanking domains in controlling the species formed during fibrillogenesis.

The sequence-specific effects of the flanking domain are unsurprising, given that primary protein sequences vary in their amyloidogenicity. Aggregation is known to depend on factors determined by the primary sequence such as hydrophobicity, charge and propensity to form secondary structure (Chiti et al., 2003), with certain amino acids having specific effects upon aggregation propensity. In keeping with its inhibitory effect on huntingtin exon-1 aggregation, proline is considered to be a ‘β-breaker’ as it is structurally incompatible with β-sheet formation (Monsellier and Chiti, 2007). There is also a group of ‘gatekeeper’ residues which are...
commonly found flanking aggregation-prone regions. These include proline, lysine, arginine, glutamate and aspartate, and all have low hydrophobicity, are charged and show a low propensity to form β-sheet structure (Rousseau et al., 2006), as demonstrated in the mutagenesis study on ataxin-2 (Nozaki et al., 2001).

Analysis of primary sequence can be very informative for unstructured proteins; however, for the natively structured polyQ proteins, an extra level of complexity exists. Evolution selects for burial of hydrophobic amyloidogenic regions, and so predicted amyloidogenic regions may be buried in a stable fold and shielded from forming the interactions required for aggregation. Thus, local destabilization of the native protein fold or increased dynamics is needed before aggregation can occur (Tartaglia et al., 2008). The role of dynamics in the initial misfolding events of polyQ aggregation has not been well characterized. We speculate that as for a number of other amyloid proteins (Wu et al., 2008; Hodkinson et al., 2009), the more dynamic the native state polyQ protein is, the more frequently it will sample amyloidogenic-prone conformations. The question remains as to whether mutation, such as the polyQ expansion in polyQ proteins, can change the protein dynamics and subsequently affect aggregation.

In addition to having a direct impact upon the polyQ tract, the flanking sequences allow for varying protein–protein interactions which also modulate aggregation. Ubiquitin is one protein present in the neuronal intranuclear inclusions (NIIs) of most polyQ diseases; however, typically the subset

<table>
<thead>
<tr>
<th>Disease</th>
<th>Protein</th>
<th>PolyQ tract location (Black) and relative size</th>
<th>Pathogenic Q length</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>HD</td>
<td>Huntingtin</td>
<td></td>
<td>36–121</td>
<td>Group (1993), Nahhas et al. (2005)</td>
</tr>
<tr>
<td>SCA1</td>
<td>Ataxin-1</td>
<td></td>
<td>40–83</td>
<td>Matilla et al. (1993), Ranum et al. (1994)</td>
</tr>
<tr>
<td>SCA3</td>
<td>Ataxin-3</td>
<td></td>
<td>45–84</td>
<td>Kawaguchi et al. (1994) Padiath et al. (2005)</td>
</tr>
<tr>
<td>SCA2</td>
<td>Ataxin-2</td>
<td></td>
<td>35–59</td>
<td>Pulst et al. (2005)</td>
</tr>
<tr>
<td>SCA6</td>
<td>CACNA1A</td>
<td></td>
<td>21–29</td>
<td>Zhu Chen et al. (1997)</td>
</tr>
<tr>
<td>SCA17</td>
<td>TBP</td>
<td></td>
<td>47–63</td>
<td>Koide et al. (1999), Nakamura et al. (2001)</td>
</tr>
<tr>
<td>DRPLA</td>
<td>Atrophin 1</td>
<td></td>
<td>49–88</td>
<td>Komure et al. (1995)</td>
</tr>
<tr>
<td>SBMA</td>
<td>Androgen receptor</td>
<td></td>
<td>40–62</td>
<td>La Spada et al. (1991)</td>
</tr>
</tbody>
</table>

Table I. Characteristics of PolyQ disease proteins

Fig. 1. Domain architecture of huntingtin exon-1, ataxin-3 and ataxin-1. The three polyQ proteins huntingtin exon-1 (A), ataxin-3 (B) and ataxin-1 (C) are shown with an arbitrarily chosen polyQ tract length of 20 glutamines. Huntingtin exon-1 consists of the HttNT domain which comprises the first 17 amino acids of the protein, the polyQ tract (Q20) and the polyP region (P). Ataxin-3 contains the Josephin domain, the polyQ tract and ubiquitin interacting motifs. Ataxin-1 contains the AXH domain and the polyQ tract. The proteins are not illustrated to scale.

DRPLA, dentatorubral and pallidopolysian atrophy; HD, Huntington’s disease; SCA, spinocerebellar ataxia; SBMA, spinal and bulbar muscular atrophy; CACNA1A-alpha. 1A-voltage-dependent calcium channel; TBP, TATA-box-binding protein.
Multi-domain misfolding

The aggregation pathway of ataxin-3 has been well defined and characterized. Ataxin-3, ataxin-1, and huntingtin all undergo multi-domain misfolding and aggregation. Owing to the technical difficulties involved in producing these proteins in vitro, the aggregation kinetics and pathways of only three of the nine polyQ proteins have been characterized. Ataxin-1, ataxin-3, and huntingtin exon-1 all contain a flanking region which is able to aggregate independently of the polyQ tract, albeit with slower kinetics. A key tool in the identification of these polyQ-independent aggregation regions has been the polyQ binding peptide QBP1, which inhibits aggregation by the polyQ tract (Nagai et al., 2000). The ability of flanking regions to independently aggregate accounts for the aggregation seen by non-expanded ataxin-1 and ataxin-3 in vitro (Chow et al., 2004a; Masino et al., 2004; de Chiara et al., 2005a, 2005b; Gales et al., 2005).

The aggregation pathway of ataxin-3 has been well defined in vitro and it has been shown that the folded 24 kDa Josephin domain of ataxin-3 aggregates when isolated (Chow et al., 2004c; Masino et al., 2004). Our group has shown that full-length ataxin-3 aggregates with a two-stage mechanism (Ellisdon et al., 2006, 2007). The first stage of aggregation involves interactions of the Josephin domain results in formation of morphologically similar soluble protofibrils by ataxin-3 with any length glutamine tract. The key evidence these fibrils are formed along the same aggregation pathway is their ability to cross-seed each other (Ellisdon et al., 2006). We know from the addition of QBP1 that the polyQ tract is not directly involved in the first step; however, the expanded ataxin-3 has the fastest aggregation kinetics, suggesting that it still impacts the Josephin domain in some way. The second stage in the polyQ aggregation mechanism involves a subsequent misfolding event, whereby interaction of the polyQ tract results in formation of SDS-insoluble fibrils. Thus, the two domains in ataxin-3, with an intrinsic ability to misfold, aggregate sequentially (Fig. 3).

Although a complete kinetic analysis has not been completed for ataxin-1 due to the limitations caused by its size, the AXH domain of ataxin-1 has similarly been found to aggregate when isolated (de Chiara et al., 2005a, 2005b). Within the context of ataxin-1, the AXH domain has a pro-aggregation effect, as replacement with a homologous non-aggregation-prone sequence decreases the formation of nuclear inclusions within a cell model (de Chiara et al., 2005a, 2005b). The formation of detergent insoluble aggregates, commonly attributed to polyQ tract aggregation, has not been investigated, nor has the use of QBP1 to ascertain whether the polyQ tract misfolds and aggregates subsequently to the AXH domain.

A recent study by the Wetzel group has shown that huntingtin exon-1 also undergoes multi-domain misfolding during aggregation (Thakur et al., 2009). They found the 17 amino acids N-terminal of the polyQ tract accelerated aggregation, whether N- or C-terminal to the polyQ tract. This differs from ataxin-1 and ataxin-3, in that the 17 amino acids are not a globular structured domain, but instead exists in a collapsed state, which an expanded polyQ tract induces into a more aggregation competent extended conformation.

Intriguingly, a kinetic analysis of the huntingtin peptide with the N-terminal sequence resulted in a critical nucleus of minus one, which suggests there is no kinetic barrier to aggregation, and spontaneous aggregation of oligomers occurs in a ‘downhill’ fashion (Thakur et al., 2009). They also found that amorphous aggregates formed in the initial step of aggregation had an exposed polyQ tract that could still bind antibodies, whereas later aggregates masked the polyQ epitope, thus suggesting an additional misfolding event and conformational rearrangement of the polyQ tract to become involved in the fibrils (Thakur et al., 2009). Thus,
The complexity of the effect of the polyQ tract on protein stability is demonstrated by a number of polyQ model systems. On the one hand, the addition and subsequent expansion of a polyQ tract did not result in a change in the thermodynamic stability of a Staphylococcus aureus protein A domain (Robertson et al., 2008). However, this domain is one of the most stable of the model proteins used, so it is unsurprising that the addition of a pathogenic length polyQ tract was not sufficient to disrupt the fold. The CRABP I domain has a similar stability to the Josephin domain of ataxin-3; however, the addition of a non-expanded polyQ tract does not impact upon stability, and it is only when the polyQ tract is expanded that there is a corresponding decrease in the stability of the protein. Furthermore, although in a GST-polyQ model, the stability was not determined, aggregation could only be induced above 50°C, suggesting the stability and/or solubility of the GST domain overrides any effects of the polyQ tract at physiological temperatures (Masino et al., 2002). Therefore, these model systems demonstrate the complexity of the relationship between the polyQ tract and the flanking domains and show that although in some situations, a polyQ tract can decrease the global stability of a protein, it is not necessary in order to induce aggregation.

Effects of flanking domains on aggregation of non-polyQ proteins

Multi-domain misfolding is not exclusive to polyQ proteins, being also seen in several other aggregating proteins such as the poly alanine (polyA) protein Poly A Binding Protein Nuclear 1 (PABPN1) and the yeast prion protein Ure2p. Expansion of the polyA tract occurs at the N-terminal of PABPN1, with an increase from 10 to 17 alanines leading to disease. Similar to ataxin-1 and ataxin-3, aggregation still occurs in the absence of the homo-sequence domain, resulting in aggregates which are smaller and more numerous than full-length PABPN1 (Klein et al., 2008). A second example, Ure2p, consists of a highly amyloidogenic unstructured N-terminal domain which is responsible for the prion phenotype, and a compactly folded C-terminal catalytic domain (Thual et al., 1999). Dimerization of the protein occurs through the C-terminal domain, which is also able to aggregate independently, forming high molecular weight insoluble oligomers, which do not progress to fibrillar aggregates (Thual et al., 2001). Interestingly, the presence of the amyloidogenic N-terminal domain has no impact upon the stability of the full-length protein (Thual et al., 2001).

Conclusions

The aggregation pathways of polyQ proteins characterized to date have involved multiple domains. With the variation seen between polyQ diseases, it is unsurprising that their specific flanking domains, although not the primary cause of aggregation, have the ability to independently misfold and modulate the fibrillogenic mechanism within the context of the full-length protein in vitro. In these studies, the importance of using a number of complementary techniques to follow the multiple aggregation steps has been highlighted. The mechanism via which an expanded polyQ tract results in accelerated aggregation remains unknown, with no clear
relationship between the effects of the expanded polyQ tract and global stability of the protein. Further investigation into polyQ protein dynamics and the effects of the expanded polyQ tract upon local stability will therefore be important for determining the early aggregation events.

Funding
This work was funded by a Program grant from the National Health and Medical Research Council of Australia (NHMRC). SPB is a Senior Research Fellow of the NHMRC.

References

Multi-domain misfolding


Received June 4, 2009; revised June 4, 2009; accepted July 6, 2009

Edited by Sheena Radford