

**EXPRESSION AND  
CHARACTERISATION OF  
CARDIOVASCULAR  
AMYLOID PROTEINS**

Thesis submitted in accordance with the  
requirements of the

**University of Liverpool**

for the degree of

**Doctor in Philosophy**

by

**HANNAH ANGHARAD DAVIES**

June 2013

## **Abstract**

There are currently 30 extracellular proteins known to form pathogenic amyloid within the human body. These proteins result in a wide range of diseases described as amyloidoses. These amyloidoses can be either systemic, affecting multiple organs, or localised to a single tissue. Almost all systemic amyloidoses can have cardiac manifestations and cardiac involvement correlates with poor prognosis. Cardiac infiltration is most commonly associated with AL amyloidosis, with 50 % of patients presenting with heart failure at the time of diagnosis.

In addition to the myocardium, amyloid deposits are frequently reported in the surrounding vasculature. The most common form of localised amyloid – aortic medial amyloid (AMA) - is primarily localised to the internal elastic laminae of the ascending aorta. It is estimated to be present in 97 % of the Caucasian population of 50 years of age but despite its prevalence there remains very little information available about the structural, biophysical and aggregation properties of the main protein component, a 5.5 kDa peptide termed medin. Clinically, it has been suggested that AMA may have a role in aortic aneurysm and dissection through binding to the elastic structures of the aorta and reducing elasticity.

This work aimed to further our molecular-level understanding of amyloid aggregates formed by proteins, known to affect the cardiovascular system; medin, and a model peptide for AL amyloidosis, SMA.

Initially, this work investigated the structural and biophysical properties of three medin derived peptides with a view to identifying key amyloidogenic regions that could be targeted therapeutically. Furthermore, this work established that small

medin peptides were not suitable structural models for full length medin. It was therefore necessary to develop a reliable method for the production of full-length medin. This work describes for the first time a procedure for the expression and purification of soluble medin in both unlabelled and  $^{13}\text{C}/^{15}\text{N}$  labelled forms in *E.coli* for future biophysical and structural investigations. The recombinant medin was then characterised using a variety of techniques commonly used for the characterisation of amyloid proteins. These results indicated that medin aggregation proceeds via a nucleation dependent mechanism with a lag time of 30 hours and forms amyloid-like fibrils visible by transmission electron microscopy. Structural and biophysical studies were coupled with computational techniques to generate a structural model of medin fibrils, stabilised by a salt bridge between residues D25 and K30. Salt bridge mutations gave rise to aggregates with altered aggregation mechanisms and morphology.

It is not known what causes some immunoglobulin light chain fragments involved in AL amyloidosis to become pathogenic. It is possible that structure underlies the differences. This work demonstrates a protocol for producing isotopically labelled immunoglobulin light chains, in particular a model protein SMA, and subsequent NMR analysis to detect inter-nuclear contacts within aggregates.

Collectively this work has identified key regions and features of medin that could be targeted in future for therapeutic or diagnostic purposes. Furthermore, it paves the way for further structural studies of aggregates formed by immunoglobulin light chains.

<b>Abstract</b> .....	<b>i</b>
<b>Table of contents</b> .....	<b>iii</b>
<b>List of figures</b> .....	<b>vii</b>
<b>Acknowledgements</b> .....	<b>x</b>
<b>Abbreviations list</b> .....	<b>xi</b>
<b>1. Introduction</b> .....	<b>1</b>
1.1. Amyloid and disease .....	1
1.1.1. Pathogenicity.....	4
1.1.2. Additional components .....	6
1.2. Protein folding and mis-folding .....	7
1.3. Mechanisms of aggregation .....	9
1.3.1. Destabilisation of the native fold .....	9
1.3.2. Non-pathogenic fibril formation .....	10
1.3.2.1. Functional amyloid .....	10
1.3.3. Nucleation dependant polymerisation .....	10
1.3.4. Seeding.....	12
1.4. Biophysical characteristics of amyloid .....	13
1.4.1. Cross- $\beta$ pattern.....	13
1.4.2. Tertiary structure.....	14
1.4.3. Quaternary structure of amyloid proteins.....	16
1.5. Therapies.....	17
1.5.1. Targeting the source.....	18
1.5.1.1. Stabilising the native fold .....	19
1.5.2. Inhibit fibrillogenesis .....	19
1.5.3. Promote clearance of aggregates.....	21
1.5.4. The importance of structural information in therapeutic design .....	22
1.6. Cardiac amyloid .....	22
1.6.1. Hereditary ATTR amyloidoses .....	23
1.6.2. Acquired ATTR amyloidosis .....	23
1.6.3. AA amyloidosis .....	23
1.6.4. A $\beta$ 2M amyloidosis.....	24
1.6.5. AL amyloidosis .....	24
1.6.5.1. Representative immunoglobulin light chain- SMA.....	27
1.7. Amyloid in the vasculature .....	28
1.7.1. Medin.....	28
1.7.1.1. Lactadherin – the precursor.....	28
1.7.1.2. Prevalence and Distribution .....	32
1.7.1.3. Consequences of medin deposition.....	33
1.8. Aims .....	34
<b>2. Investigating the packing arrangements of medin fibrils</b> .....	<b>36</b>
2.1. Introduction.....	36
2.2. Materials and methods .....	40
2.2.1. Overall strategy .....	40
2.2.2. Amyloid prediction servers .....	40
2.2.2.1. Zyggregator.....	40
2.2.2.2. AGGRESKAN .....	41
2.2.2.3. Prediction of Amyloid STructure Aggregation (PASTA).....	41
2.2.3. Preparation of medin peptide samples for NMR studies .....	41
2.2.4. Characterisation of medin peptide aggregation .....	42
2.2.4.1. Aggregation studies of medin peptides by thioflavin T .....	42
2.2.4.2. Morphology of medin peptides by transmission electron microscopy.....	42
2.2.5. Solid state nuclear magnetic resonance (SSNMR) studies on synthetic medin peptides. 43	
2.2.5.1. Dipolar assisted rotational resonance experiments .....	43
2.2.5.2. Rotational resonance experiments.....	44

2.2.5.3.	Site-specific $^{13}\text{C}$ - $^1\text{H}$ dipolar couplings .....	44
2.2.6.	Solution state NMR studies of Med <sub>1-50</sub> .....	46
2.3.	Results .....	46
2.3.1.	Aggregation prediction servers .....	46
2.3.2.	Kinetic and morphological characterisation of Med <sub>30-50</sub> and Med <sub>1-50</sub> .....	48
2.3.2.1.	Aggregation kinetics of Med <sub>30-50</sub> and Med <sub>1-50</sub> studied by Thioflavin T .....	48
2.3.3.	Morphology of Med <sub>42-49</sub> , Med <sub>30-50</sub> and Med <sub>1-50</sub> aggregates studied by TEM. ....	50
2.3.4.	Peptide NMR .....	52
2.3.4.1.	DARR spectra of Med <sub>30-50</sub> and Med <sub>1-50</sub> aggregates .....	52
2.3.4.2.	Investigating side chain packing arrangements examined by rotational resonance experiments .....	57
2.3.4.3.	Aromatic ring flexibility- CT-DIPSHIFT solid-state NMR experiments .....	59
2.3.4.4.	Solution state NMR studies of Med <sub>1-50</sub> .....	65
2.4.	Discussion .....	68
2.4.1.	Investigating the role of predicted amyloidogenic regions on the kinetics, morphology and structure of medin .....	68
<b>3.</b>	<b>Recombinant expression and purification of medin .....</b>	<b>72</b>
3.1.	Introduction .....	72
3.2.	Materials and methods .....	74
3.2.1.	Recombinant expression of medin in <i>E.coli</i> .....	74
3.2.1.1.	Construct design .....	74
3.2.1.2.	Preparation of chemically-competent cells .....	75
3.2.1.3.	Transformation of vector into chemically competent cell lines. ....	76
3.2.1.4.	Expression tests .....	76
3.2.1.5.	Expression of medin in the pMALC2X vector .....	77
3.2.1.6.	Sub-cloning of medin into pOPINS vector .....	78
3.2.1.7.	Agarose gel electrophoresis .....	81
3.2.1.8.	Expression of pOPINS construct .....	81
3.2.2.	Purification of MBP- medin fusion protein .....	81
3.2.2.1.	Amylose affinity purification .....	81
3.2.2.2.	Cleavage .....	82
3.2.2.3.	Isolating medin .....	82
3.2.2.4.	Purification of SUMO-Medin fusion protein .....	83
3.2.2.5.	Cleavage and removal of the His <sub>6</sub> -SUMO tag. ....	84
3.2.3.	Characterisation .....	84
3.2.3.1.	SDS PAGE gel analysis and staining methods .....	84
3.2.3.2.	Western blot analysis .....	85
3.2.3.3.	Tryptic digest MS-MS. ....	86
3.3.	Results and Discussion .....	86
3.3.1.	Cloning, expression and purification of medin .....	86
3.3.1.1.	pMALC2X construct .....	86
3.3.1.2.	pOPINS construct .....	91
3.3.2.	Characterisation of medin .....	96
3.3.2.1.	Visualisation .....	96
3.3.2.2.	Western blot analysis .....	97
3.3.2.3.	Tryptic digest and mass spectrometry .....	98
3.4.	Discussion .....	98
<b>4.</b>	<b>Biophysical characterisation of medin aggregation .....</b>	<b>101</b>
4.1.	Introduction .....	101
4.2.	Materials and methods .....	105
4.2.1.	Aggregation kinetics studied by Thioflavin T .....	105
4.2.2.	Aggregation kinetics by intrinsic fluorescence .....	106
4.2.3.	Determination of aggregate size distribution by dynamic light scattering .....	107
4.2.4.	Secondary structure analysis by circular dichroism .....	107
4.2.5.	Analysis of morphology by transmission electron microscopy .....	110
4.3.	Results .....	111
4.3.1.	Aggregation kinetics of medin .....	111
4.3.1.1.	Investigating the concentration dependence of medin aggregation .....	111
4.3.1.2.	Investigating the effect of agitation on medin aggregation .....	115

4.3.1.3.	Kinetic analysis by intrinsic tryptophan fluorescence.....	117
4.3.1.4.	Dynamic light scattering .....	119
4.3.2.	Structure and morphology of medin aggregates.....	120
4.3.2.1.	Circular dichroism .....	120
4.3.2.2.	Electron microscopy .....	125
4.3.3.	Conditions for future characterisation.....	125
4.3.4.	Affect of inhibitors on medin aggregation .....	125
4.3.4.1.	$\beta$ -sheet breaking inhibitors.....	127
4.4.	Discussion .....	133
<b>5.</b>	<b>Toward a structural model for medin fibrils.....</b>	<b>138</b>
5.1.	Introduction.....	138
5.1.1.	Stabilising interactions within amyloid fibrils .....	138
5.1.2.	Similarities between medin and other amyloid proteins .....	140
5.1.2.1.	Salt bridge in $A\beta_{1-40}$ : D23 and K28.....	142
5.1.3.	$A\beta$ Iowa mutant .....	143
5.1.4.	Objectives .....	145
5.2.	Materials and methods .....	146
5.2.1.	Overall strategy .....	146
5.2.1.1.	Sequence analysis .....	148
5.2.1.2.	Modelling.....	148
5.2.1.3.	Model validation and distance measurements.....	148
5.2.2.	Solid state NMR studies.....	149
5.2.2.1.	Preparation of medin for NMR studies. ....	149
5.2.2.2.	One-dimensional spectra of isotopically labelled recombinant medin.....	150
5.2.2.3.	Two-dimensional DARR spectra of isotopically labelled recombinant medin. ....	150
5.2.2.4.	Simulated spectra .....	151
5.2.2.5.	Heteronuclear distance measurements using frequency selective REDOR .....	152
5.2.3.	Mutagenesis .....	153
5.2.3.1.	Generation of D25N and K30I mutants by site-directed ligase independent mutagenesis (SLIM). ....	153
5.2.3.2.	Expression of mutants .....	156
5.2.4.	Characterisation of mutants .....	156
5.2.4.1.	Kinetic analysis by thioflavin T assay and intrinsic fluorescence.....	156
5.2.4.2.	Structure and morphology analysis by CD and electron microscopy. ....	157
5.3.	Results.....	157
5.3.1.	Models .....	157
5.3.1.1.	$A\beta_{1-40}$ template.....	157
5.3.1.2.	Medin models.....	159
5.3.2.	Structural studies of wild type recombinant medin by solid state NMR.....	161
5.3.2.1.	One-dimensional solid-state NMR.....	162
5.3.2.2.	DARR spectrum and simulated spectra .....	164
5.3.3.	Testing the salt bridge hypothesis using solid state NMR frequency selective REDOR experiments. 169	
5.3.4.	Mutagenesis .....	170
5.3.4.1.	Cloning.....	170
5.3.4.2.	Mutant expression .....	171
5.3.5.	Aggregation kinetics of mutant peptides D25N and K30I. ....	173
5.3.5.1.	ThT analysis.....	173
5.3.5.2.	Aggregation kinetic of mutants using intrinsic fluorescence .....	176
5.3.6.	Structure and morphology of mutants.....	177
5.3.6.1.	Monitoring secondary structure of D25N over time .....	177
5.3.6.2.	Investigating the stability of D25N in response to UV radiation .....	181
5.3.6.3.	Monitoring the secondary structure of K30I over time. ....	182
5.3.6.4.	Investigating the stability of K30I in response to UV radiation.....	185
5.3.7.	Morphology of mutant aggregates studied by transmission electron microscopy. ..	186
5.4.	Discussion .....	188
5.4.1.	Solid state NMR measurements of medin.....	189
5.4.2.	Mutagenesis studies .....	190
5.4.2.1.	Characterisation of D25N .....	190

5.4.3.	Characterisation of K30I.....	191
<b>6.</b>	<b>Expression and purification of SMA for structural studies .....</b>	<b>193</b>
6.1.	Introduction.....	193
6.1.1.	Immunoglobulin light chain amyloidosis.....	193
6.1.2.	SMA as a model system.....	194
6.1.3.	Previous characterisation .....	194
6.1.4.	Structural information .....	196
6.2.	Materials and methods .....	199
6.2.1.	Vector – pKIV007.....	199
6.2.2.	Expression and purification of pKIV007 .....	199
6.2.2.1.	Anion exchange chromatography .....	200
6.2.2.2.	Size exclusion chromatography .....	201
6.2.3.	Subcloning .....	201
6.2.4.	Expression and purification of SUMO-SMA.....	201
6.2.4.1.	Expression trials and conditions.....	202
6.2.5.	Confirmation of SMA .....	<b>Error! Bookmark not defined.</b>
6.2.5.1.	SDS PAGE analysis.....	202
6.2.5.2.	Tryptic digest and MALDI -MS .....	202
6.2.5.3.	Intact mass analysis using ES-MS .....	202
6.2.6.	Characterisation of SMA aggregation under different buffer conditions .....	203
6.2.6.1.	Secondary structure analysis by circular dichroism .....	203
6.2.6.2.	Morphology analysis by transmission electron microscopy .....	204
6.2.7.	NMR studies .....	204
6.2.7.1.	Preparation of <sup>13</sup> C and <sup>15</sup> N isotopically labelled SMA mixture .....	204
6.2.7.2.	Frequency selective REDOR experiment .....	204
6.2.8.	SMA model.....	205
6.2.8.1.	Model validation and distance measurements.....	205
6.3.	Results.....	206
6.3.1.	Expression and purification of pKIV007 .....	206
6.3.1.1.	Ion exchange chromatography purification.....	208
6.3.1.2.	Size exclusion chromatography purification.....	208
6.3.2.	Sub-cloning of SMA into pOPINS vector.....	209
6.3.3.	Expression and purification of SMA in pOPINS vector. ....	210
6.3.3.1.	Expression tests.....	210
6.3.3.2.	Purification of SMA.....	213
6.3.4.	Confirmation of SMA .....	214
6.3.5.	Characterisation of SMA aggregation under different buffer conditions .....	217
6.3.5.1.	Secondary structure and stability analysis by circular dichroism .....	217
6.3.5.2.	Morphology analysis by transmission electron microscopy .....	221
6.3.6.	NMR experiments .....	223
6.4.	Discussion .....	227
<b>7.</b>	<b>Discussion and future work.....</b>	<b>231</b>
7.1.	Medin .....	231
7.2.	SMA.....	238
	<b>References.....</b>	<b>238</b>
	<b>Appendix 1.....</b>	<b>259</b>
	<b>Appendix 2.....</b>	<b>265</b>
	<b>Appendix 3.....</b>	<b>267</b>
	<b>Appendix 3.....</b>	<b>268</b>

## List of figures

Figure 1.1 .....	5
Figure 1.2 .....	8
Figure 1.3 .....	9
Figure 1.4 .....	12
Figure 1.5 .....	15
Figure 1.6 .....	17
Figure 1.7 .....	18
Figure 1.8 .....	20
Figure 1.9 .....	25
Figure 1.10 .....	27
Figure 1.11 .....	31
Figure 1.12 .....	33
Figure 2.1 .....	37
Figure 2.2 .....	47
Figure 2.3 .....	48
Figure 2.4 .....	49
Figure 2.5 .....	51
Figure 2.7 .....	54
Figure 2.8 .....	58
Figure 2.9 .....	60
Figure 2.10 .....	62
Figure 2.11 .....	64
Figure 2.12 .....	66
Figure 2.13 .....	67
Figure 2.14 .....	69
Figure 3.1 .....	75
Figure 3.2 .....	78
Figure 3.3 .....	80
Figure 3.4 .....	87
Figure 3.5 .....	88
Figure 3.6 .....	90
Figure 3.7 .....	92
Figure 3.8 .....	93
Figure 3.9 .....	94
Figure 3.10 .....	95
Figure 3.11 .....	97
Figure 4.1 .....	106
Figure 4.2 .....	108
Figure 4.3 .....	109
Figure 4.4 .....	112
Figure 4.5 .....	114
Figure 4.6 .....	116
Figure 4.7 .....	117
Figure 4.8 .....	118
Figure 4.9 .....	119
Figure 4.10 .....	120
Figure 4.11 .....	121
Figure 4.12 .....	122
Figure 4.13 .....	123



Figure 4.14.	124
Figure 4.15.	125
Figure 4.16 .	129
Figure 4.17.	130
Figure 4.18.	131
Figure 4.19.	132
Figure 5.1.	140
Figure 5.2.	141
Figure 5.3.	142
Figure 5.4.	145
Figure 5.5.	147
Figure 5.6	149
Figure 5.7.	154
Figure 5.8	155
Figure 5.9.	158
Figure 5.10	161
Figure 5.11.	163
Figure 5.12.	164
Figure 5.13.	167
Figure 5.14.	168
Figure 5.15.	168
Figure 5.16.	170
Figure 5.17.	171
Figure 5.18	172
Figure 5.19.	174
Figure 5.20.	175
Figure 5.21.	176
Figure 5.22.	177
Figure 5.23.	178
Figure 5.24.	179
Figure 5.25.	180
Figure 5.26.	181
Figure 5.27.	182
Figure 5.28.	184
Figure 5.29.	185
Figure 5.30.	186
Figure 5.31.	187
Figure 6.1.	195
Figure 6.2.	197
Figure 6.3.	201
Figure 6.4	207
Figure 6.5.	209
Figure 6.6	210
Figure 6.7 .	212
Figure 6.8.	213
Figure 6.9.	214
Figure 6.10.	216
Figure 6.11 .	218
Figure 6.12	220
Figure 6.13.	221

Figure 6.14. ....	222
Figure 6.15. ....	224
Figure 6.16. ....	225
Figure 6.17. ....	229

## Acknowledgments

My first thanks must go to my supervisor Prof. David Middleton for believing I was capable of doing the work and sticking with me, despite my NMR reluctance!

I owe huge thanks to Jill Madine for her continuous support and friendship – there is no doubt I would not have got this far without her encouragement and expertise. From coffee meetings during maternity leave to keeping me company at 4am on Diamond trips...I appreciate it more than I can say!

I look forward to the next round of adventures for ‘the amyloids’!!

It follows that I owe Marie Phelan thanks too for stepping in when I was a floundering first year and providing much support and guidance while Jill was away, which continued and developed into a firm friendship.

There are too many people to name them all that have provided expertise and guidance that have helped me on my way, in Lab C and throughout the department, but special mention to Dr. Mark Wilkinson for all his help and expertise with the production of medin and for staying late to do mass spec on fresh samples!

I must also thank the late Dr Fred Stevens for providing the SMA clone, Dinu Iuga at the National solid-state NMR facility at Warwick University and the staff at Diamond light source for all their technical assistance with circular dichroism.

I must also thank those who have never seen the lab! My friends and family have been hugely supportive of me throughout, and provided biscuits, tea and wine when required! Mum and Dad deserve a medal for putting up with the ups and downs of research and always pointing out the end goal and the ‘bigger picture’. Thanks also go to my grandparents-they have done everything they can to help including the dreaded proofreading!! Lastly, I must thank Ed for all his support, 4 years of travelling 300+ miles, late night phone calls and amazing holidays-we made it in the end!!!

This work was carried out with funding from a PhD studentship awarded by the British Heart Foundation.

The solid-state NMR data collection and initial analysis was carried out by Prof. D. A. Middleton.

## Abbreviations list

<b>AA</b> serum amyloid A amyloid	<b>kV</b> kilo volt
<b>A<math>\beta</math></b> amyloid beta peptide	<b>LB</b> luria broth
<b>AL</b> immunoglobulin light chain amyloid	<b>MALDI</b> matix assisted laser desorbtion and ionisation
<b>AMA</b> aortic medial amyloid	<b>Med</b> <sub>42-49</sub> Medin peptide comprising residues 42-49
<b>AApoA1</b> Apolipoprotein A1 amyloid	<b>Med</b> <sub>30-50</sub> Medin peptide comprising residues 30-50
<b><math>\alpha</math>-syn</b> $\alpha$ -synuclein	<b>Med</b> <sub>1-50</sub> Full length medin
<b>ATTR</b> TTR derived amyloidosis	<b>MBP</b> maltose binding protein
<b><math>\beta</math>2M</b> $\beta$ 2-microglobulin	<b>NMR</b> nuclear magnetic resonance
<b>BSA</b> bovine serum albumin	<b>O.D.</b> optical density
<b>CD</b> circular dichroism	<b>OPPF</b> oxford protein production facility
<b>CP-MAS</b> magic angle spinning	<b>PCR</b> polymerase chain reaction
<b>CR</b> congo red	<b>ppm</b> parts per million
<b>CT-DIPSHIFT</b> constant time-dipolar and chemical shift correlation experiment	<b>PVDF</b> polyvinylidene fluoride
<b>CV</b> column volume	<b>RR</b> rotational resonance
<b><math>\delta</math></b> chemical shift	<b>SAP</b> serum amyloid P
<b>DAB</b> 3,3'diaminobenzidine tetrahydorchloride	<b>SDS-PAGE</b> sodium dodecyl sulphate polyacrylamide gel electrophoresis
<b>DARR</b> dipolar assisted rotational resonance	<b>SOB</b> super optimal broth
<b>dH<sub>2</sub>O</b> distilled water	<b>SOC</b> super optimal broth with catabolic repression
<b>DLS</b> dynamic light scattering	<b>SRCD</b> synchrotron radiation circular dichroism
<b>DMSO</b> dimethyl sulphoxide	<b>TAE</b> tris-acetate-EDTA
<b>DNA</b> deoxyribose nucleic acid	<b>TB</b> terrific broth
<b><i>E.coli</i></b> Esherisha coli	<b>TCA</b> trichloroacetic acid
<b>EDTA</b> ethylenediamine-tetraacetic acid	<b>TEM</b> transmission electon microscopy
<b>ES-MS</b> electrospray mass spectrometry	<b>TEV</b> tobacco etch virus
<b>GAG</b> glycosaminoglycan	<b>ThT</b> Thioflavin T
<b>IAPP</b> islet amyloid polypeptide	<b>TPPM</b> two-pulse phase-modulated
<b>IF</b> intrinsic fluorecence	<b>TTR</b> transthyretin
<b>IMAC</b> immobilised metal affinity chromatography	<b>UV</b> ultraviolet
<b>HFIP</b> hexafluoroisopropanol	
<b>IPTG</b> isopropyl $\beta$ -D-thiogalactoside	
<b>kDa</b> kilo Dalton	
<b>kHz</b> kilo hertz	

# 1. Introduction

The term amyloid was first used by Virchow in 1857 to describe ‘starch-like’ deposits within tissue that stained with iodine and sulphuric acid (Virchow, 1857). It is now known that this phenomena is due to accessory components not the fibrillar content (Pras, 1968). For approximately the next 100 years the only tools available to study amyloid proteins were light microscopy and a variety of dyes (Buxbaum and Linke, 2012). In the 1920’s it was established that all amyloid deposits bind the dye Congo red (CR) and exhibit apple green birefringence when viewed under polarised light (Puchtler et al., 1962, Divrey, 1927). It remains unclear exactly how CR binds to amyloid deposits, but it still remains a diagnostic characteristic of amyloid. Later developments in biochemistry have greatly increased our understanding of the biophysical and structural properties of amyloid deposits and are now included in the definition of amyloid.

The international society for amyloidosis now defines amyloid as insoluble protein fibrils, primarily found in the extracellular spaces between organs and tissues that result in alternative protein folding and ultimately a condition known as amyloidosis (Sipe et al., 2012). Amyloid fibril proteins have characteristic structural and tinctorial properties, they occur as non-branching fibrils approximately 10 nm in diameter, have a distinguishing cross- $\beta$  pattern when studied by X-ray fibre diffraction (XRFD) and bind CR (Sipe et al., 2012, Divrey, 1927, Puchtler et al., 1962).

## 1.1. Amyloid and disease

Diseases linked with amyloid affect a wide range of tissues and are associated with high morbidity and mortality rates. The amyloidoses can be divided into localised and systemic diseases. Localised amyloid affects a single tissue type whereas systemic amyloid affects multiple tissues and organs. These can be acquired, as seen

in amyloid light chain (AL) and serum amyloid A (AA) amyloidosis, or hereditary in transthyretin (ATTR) variant and apolipoprotein A1 (AApoA1) amyloidosis (Table 1.1).

The signs and symptoms associated with the amyloidoses are highly variable and differ greatly between the different diseases. Several amyloid-associated diseases are localised to the brain and result in neurodegeneration and cognitive decline (Bredesen, 2009). In systemic amyloidoses, such as AL amyloidosis, symptoms include heart failure, hepatomegaly and autonomic or sensory neuropathy (Sanchorawala, 2006).

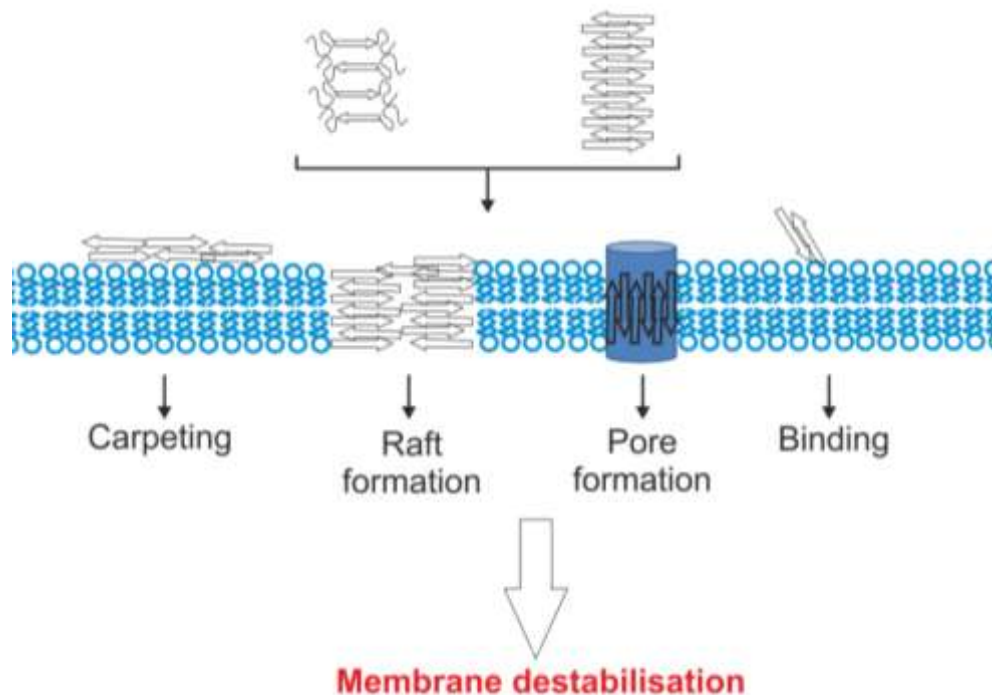
**Table 1.1 Summary of the different types of amyloid with cardiovascular manifestations including examples of specific pathologies and the respective protein components, adapted from (Banyersad et al., 2012). Those diseases and proteins italicised comprise the focus of this thesis.**

<b>Disease</b>	<b>Protein</b>	<b>Cardiovascular involvement</b>	<b>Other organ involvement</b>	<b>Treatment</b>	<b>Prognosis</b>	<b>Prevalence</b>
<i>AL amyloidosis</i>	<i>Monoclonal light chain</i>	40-50%	<i>Kidneys, liver, soft tissue, PNS</i>	<i>Chemotherapy or stem cell transplantaion</i>	<i>48 months or 8 if cardiac tissue involved</i>	<i>Affects 1 in 130000</i>
<b>ATTR</b>	wild type transthyretin	>90%	Carpel tunnel syndrome	supportive	7-8 years	25 % in people over 80 years of age
	mutant transthyretin mutant	Range from rare to >90% depending on mutation	PNS, Carpel tunnel syndrome	Liver transplant or supportive therapy	Variable depending on mutation and treatment	Variable depending on mutation,
<b>AA amyloidosis</b>	serum amyloid A	Rare	Kidneys	Treat underlying inflammatory disease	good	0.5-0.89 % of the population
<b>A<math>\beta</math>2M</b>	$\beta$ 2-Microglobulin	rare	joints	kidney transplant	variable	48 % of patients on long term dialysis but number falling
<i>Aortic medial amyloid (AMA)</i>	<i>medin</i>	<i>100%</i>	<i>unknown</i>	<i>none</i>	<i>unknown</i>	<i>90 % of caucasian population over 50.</i>

### 1.1.1. Pathogenicity

Despite the advances in amyloid research, the pathogenic species and thus the causative agent of these devastating diseases remains unclear. The role of fibrillar deposits in the pathogenicity of amyloidosis continues to be debated. It is still inconclusive whether they represent a causative agent or are merely a secondary by-product of the pathology. This primarily arises from results indicating that there is no correlation between the amount of insoluble plaques of A $\beta$  present in the brain and the degree of cognitive decline (Näslund et al., 2000). Conversely, it appears that higher concentrations of soluble A $\beta$  correlated with cognitive impairment in patients with Alzheimer's disease (Wang, 1999). The hypothesis that deposits are merely by-products is further substantiated by evidence that amyloid plaques can be observed many years before clinical symptoms appear (Perrin et al., 2009). *In vitro* studies have also shown that A $\beta$  pathogenicity is linked to the level of aggregation. Monomeric or highly aggregated fibrils are not toxic, but small oligomeric species are in some cases cytotoxic (Kayed et al., 2003, Madine and Middleton, 2010, Bucciantini et al., 2002, Fändrich, 2012). The toxicity of oligomeric species has focussed attention on the multimeric intermediate species in the amyloidogenesis pathway





**Figure 1.1 Schematic to illustrate some of the possible mechanisms of amyloid toxicity caused by multimeric intermediates within the amyloid pathway. (Berthelot et al., 2013).**

Many of the proposed cytotoxic mechanisms of these oligomers involve disruption and destabilisation of the lipid membrane (Figure 1.1) (Stefani, 2007, Berthelot et al., 2013).

Many studies have reported accelerated aggregation in the presence of lipids and membranes, resulting in carpeting and raft formation, whereby amyloid proteins assemble on or within the membrane preventing normal membrane activity. (Necula et al., 2003, ChooSmith et al., 1997, Knight and Miranker, 2004). Although these interactions are far from fully understood it seems that hydrophobic interactions are a key determinant in how different amyloid proteins behave in the presence of lipids (Berthelot et al., 2013).

Amyloid proteins are also capable of interfering with normal membrane physiology at different stages of aggregation (Williams and Serpell, 2011). For example,

monomeric A $\beta$  can impede the diffusion of metabotropic glutamate receptors through binding of the synaptic plasma membranes which results in abnormal Ca<sup>2+</sup> mobilization (Renner et al., 2010). A further mechanism involves the formation of pores within the membrane resulting in cytotoxicity (Demuro et al., 2011, Butterfield and Lashuel, 2010, Milanesi et al., 2012)

Notwithstanding the confusion surrounding the cytotoxic mechanisms of oligomeric species, these are more commonly observed in neurodegenerative amyloidoses.

The primary pathology associated with systemic diseases is the deposition of amyloid within the visceral organs. It remains evident that large quantities of amyloid deposits, in sometimes kilogram amounts, can impair organ function and damage tissue structure (Pepys, 2001).

### **1.1.2. Additional components**

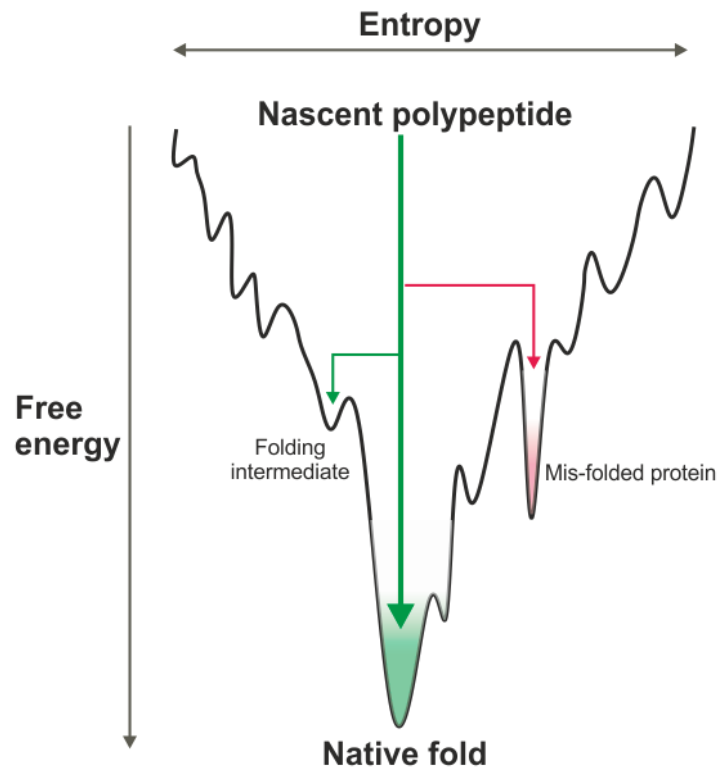
Amyloid deposits are always found associated with other proteins, glycosaminoglycans (GAGs) such as heparan sulphate, and proteoglycans (Papy-Garcia et al., 2011, Pepys, 2001, Ariga et al., 2010). GAGs and other additional components are important in the process of fibrillogenesis and understanding their role in plaque formation may provide ways of modulating amyloid formation. Furthermore they can be useful tools for the development of imaging techniques or diagnostics. Several components in particular are well documented to be found in amyloid plaques; one such protein is serum amyloid P (SAP) protein (Pepys et al., 1997). This protein has been shown to bind to amyloid fibrils *in vitro* in a calcium dependent manner (Pepys et al., 1979). Later work has exploited this interaction and led to the development of new imaging techniques that use SAP, labelled with iodine isotopes 123 or 131, to image deposits within the body (Hawkins et al., 1998, Lovat et al., 1998b). This is a very effective method for imaging deposits in patients with

systemic amyloidoses but it is unclear whether it is sufficiently sensitive to image smaller deposits in the brain (Lovat et al., 1998a).

Many studies have documented the presence of glycosaminoglycans (GAGs) within amyloid deposits *in vivo* e.g. (Stevens and Kisilevsky, 2000) GAGs are large unbranched polysaccharides consisting of repeating disaccharide units. Cardiac tissue containing amyloid deposits have been shown to have 5-fold increase in GAGs (Ohishi et al., 1990) The presence of GAGs has also been shown to increase the rate of aggregation (Cohlberg et al., 2002, Madine and Middleton, 2010). The intimate relationship between amyloid fibrils and GAGs is currently being exploited for new therapeutic options. This is discussed further in Section 1.5.2.

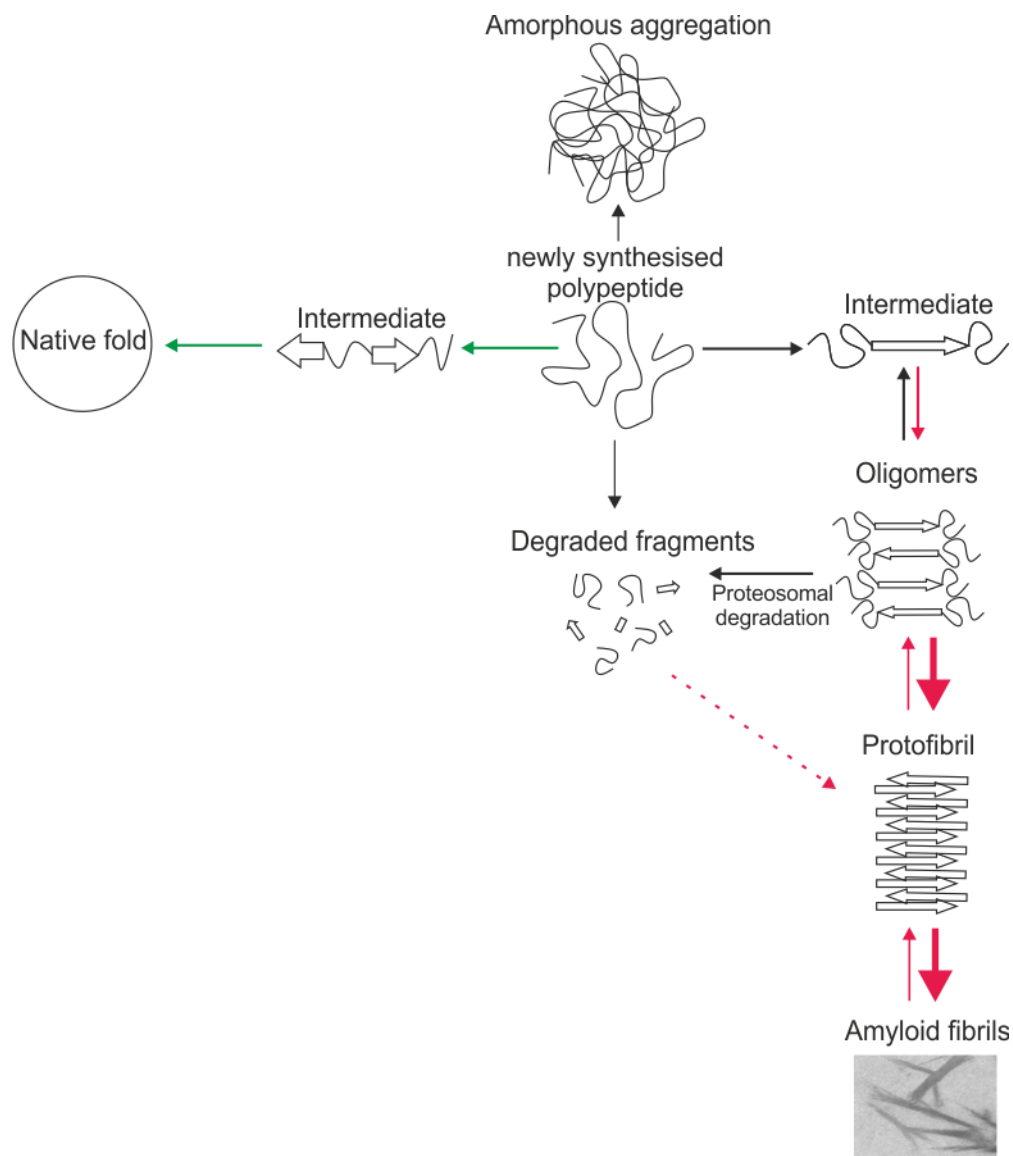
## **1.2. Protein folding and mis-folding**

Protein folding is a process whereby polypeptide chains transform into specific three-dimensional structures. This conversion is prerequisite to the biological function of a protein and is an important means of biological control. The current theory describing this process utilises the concept of a funnelled energy landscape (Onuchic and Wolynes, 2004). This theory assumes that the native fold of a protein corresponds to the structure with the lowest free energy. Initially, a polypeptide chain samples many possible conformations but as interactions start to occur, lower energy structures are selected, and ultimately this results in the native fold. The landscape is often described as rugged to account for the multiple non-native local minima and may act as kinetic traps along the route to the native fold (Figure 1.2).



**Figure 1.2 Schematic of the rugged funnel energy landscape of protein folding. As a protein interacts and starts to fold, free energy is reduced. The pathways highlighted in green indicates possible routes to the native fold, including small kinetic traps that can be overcome with the help of chaperones. The landscape also features local energy minima that can act as kinetic traps for mis-folded proteins (red) that may result in amyloid formation.**

Some of these traps can be overcome with the help of chaperone proteins but, in some cases the mis-folded protein can aggregate (Dobson, 2003, Dobson, 2006, Chiti and Dobson, 2006). The mis-folded protein may form well ordered aggregates that act as the precursor to amyloid formation (Figure 1.3)



**Figure 1.3 Schematic illustrating some possible routes to amyloid formation. Green arrows indicate the correct route to the native fold and red arrows indicate disadvantageous routes. The large arrows indicate the most kinetically favourable route. Degraded fragments can become incorporated into amyloid fibrils, indicated by the dashed red line.**

### 1.3. Mechanisms of aggregation

#### 1.3.1. Destabilisation of the native fold

There are many different factors that can cause proteins and polypeptides to mis-fold and aggregate. In some cases aggregation follows cleavage from a precursor protein as in the case of A $\beta$  (Tomita et al., 1998). In other examples, aggregation results from destabilisation or conformational change of the intermediates or native fold.

These can occur as a result of mutation (Steward et al., 2008), membrane binding (Williamson et al., 2009), molecular crowding (Hatters et al., 2002), binding of metal ions (Mantyh et al., 1993) or the presence of a seed (further details in section 1.3.4) (Paravastu et al., 2009). It is thought that this destabilisation exposes aggregation prone regions, which are then able to self-assemble and form amyloid-like structures.

### **1.3.2. Non-pathogenic fibril formation**

It has been suggested that fibril formation is a generic property of all polypeptides and that under the right destabilising conditions all proteins are capable of forming fibrillar species (Pertinhez et al., 2001). This was first noticed fortuitously but has since been observed with many plasma proteins with no known disease links (Chiti et al., 2001, Sunde and Blake, 1998) . It should be noted that the conditions required to destabilise some of these proteins, such as low pH, are frequently non-physiological and thus these proteins differ from their pathogenic counterparts which are capable of aggregating *in vivo*.

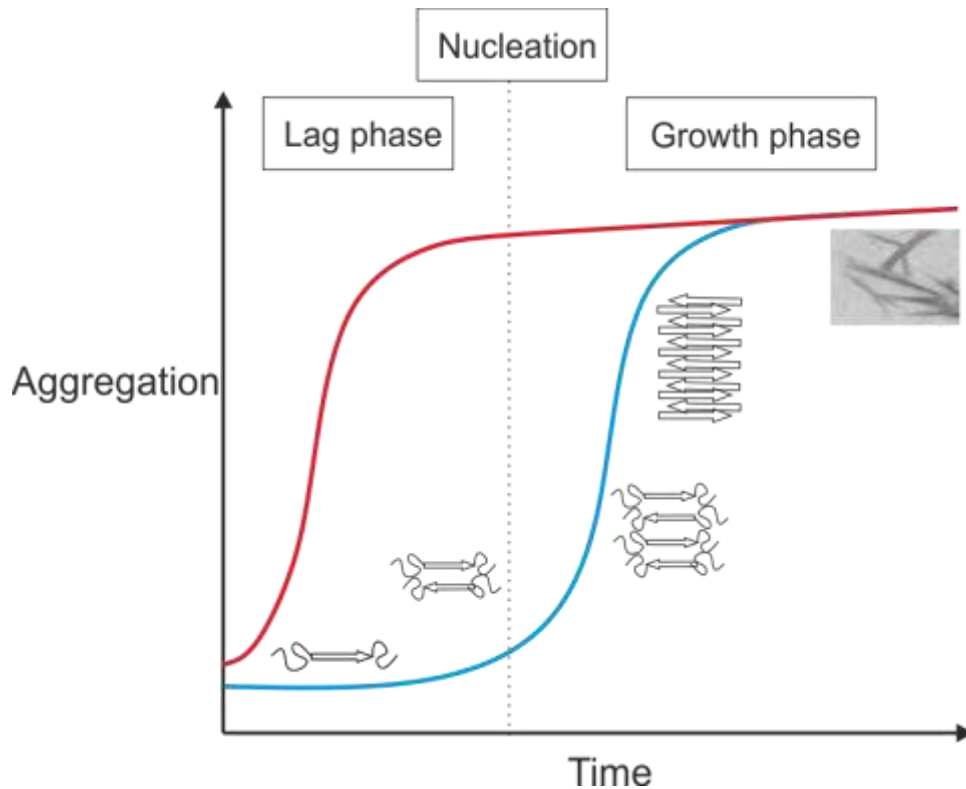
#### **1.3.2.1. Functional amyloid**

There is a further sub-class of proteins utilising an amyloid-like protein fold to perform a function. There are many examples in nature, in a wide variety of organisms, such as: curlin in *E.Coli* (Chapman et al., 2002), spidroin in *N. edulis* (Kenney et al., 2002), the intra-luminal domain of Pmel17 (Berson et al., 2003) and the organisation of hormones stored in secretory vesicles (Maji et al., 2009) in humans.

### **1.3.3. Nucleation dependant polymerisation**

Nucleation dependent polymerisation is the current leading theory for amyloid formation and growth. The assembly process is characterised by three distinctive features, a lag phase, a nucleation step, and a growth and elongation phase (Figure

1.4). Initially a lag time is observed, during which monomeric protein and small soluble oligomers associate to form a critical nucleus. During this stage aggregates are usually undetectable by electron microscopy (Chiti and Dobson, 2006). The process of nucleation is kinetically unfavourable and is highly dependent on protein concentration and other variables such as pH and salt concentration (Jarrett and Lansbury Jr, 1993). It is this nucleus formation that represents the rate limiting step in amyloid formation. Once the critical nucleus is formed, amyloidogenesis enters a growth and elongation phase. This process is thermodynamically favourable and proceeds rapidly resulting in the formation of fibrillar species with amyloid characteristics, detectable by electron microscopy (Figure 1.4). The majority of aggregation kinetic studies monitor aggregation using the benzothiazole salt Thioflavin T (ThT). Upon binding to  $\beta$ -rich structures, ThT exhibits enhanced fluorescence and a red shift in its emission spectra. Monitoring the fluorescence increase allows researchers to observe aggregation, often in real time and subsequently calculate kinetic parameters.



**Figure 1.4 Schematic of amyloid nucleation dependent growth theory. The blue growth curve indicates normal amyloid formation and the red curve highlights the change upon addition of a seed. The vertical dotted line designates the point of nucleation.**

#### 1.3.4. Seeding

Seeding is the term given to the process of adding pre-formed aggregates or a ‘nucleus’ to monomeric proteins to eliminate the lag phase (Figure 1.4 red growth curve) (Come et al., 1993). This phenomenon has been observed both *in vitro* (Kim et al., 2007, Petkova et al., 2005a, Jarrett and Lansbury Jr, 1993) and *in vivo* (Prado and Baron, 2012, Rosen et al., 2012, Jucker and Walker, 2011). Seeding usually requires a high degree of similarity between the seed and the elongating species (i.e. homologous seed) but there is evidence that cross-seeding can occur between protein species. The most notable example of heterologous seeding can be observed in the relationship between bovine spongiform encephalopathy and Creutzfeldt-Jacob disease, where the bovine protein (93 % sequence identity) has been shown to be



able to seed the human protein (Mabbott and MacPherson, 2006). More recently it has been shown that one amyloid protein is capable of seeding a protein with a different origin (Larsson et al., 2011). A study investigating the effect of sequence identity on seeding kinetics revealed a high level of sequence identity gave rise to optimal seeding. However mutant studies have shown that primary sequence similarity is not as important as structural similarity between the fibril growth face and the seed (O'Nuallain et al., 2004, Krebs et al., 2004). Studies investigating cross-seeding between A $\beta$  and islet amyloid polypeptide (IAPP) demonstrated that A $\beta$  was able to seed IAPP but not *vice versa* implying that heterologous seeding can be one-directional (O'Nuallain et al., 2004). It is unknown whether these cross-seeding interactions occur *in vivo* and what the potential implications could be with regards to toxicity and therapeutics.

#### **1.4. Biophysical characteristics of amyloid**

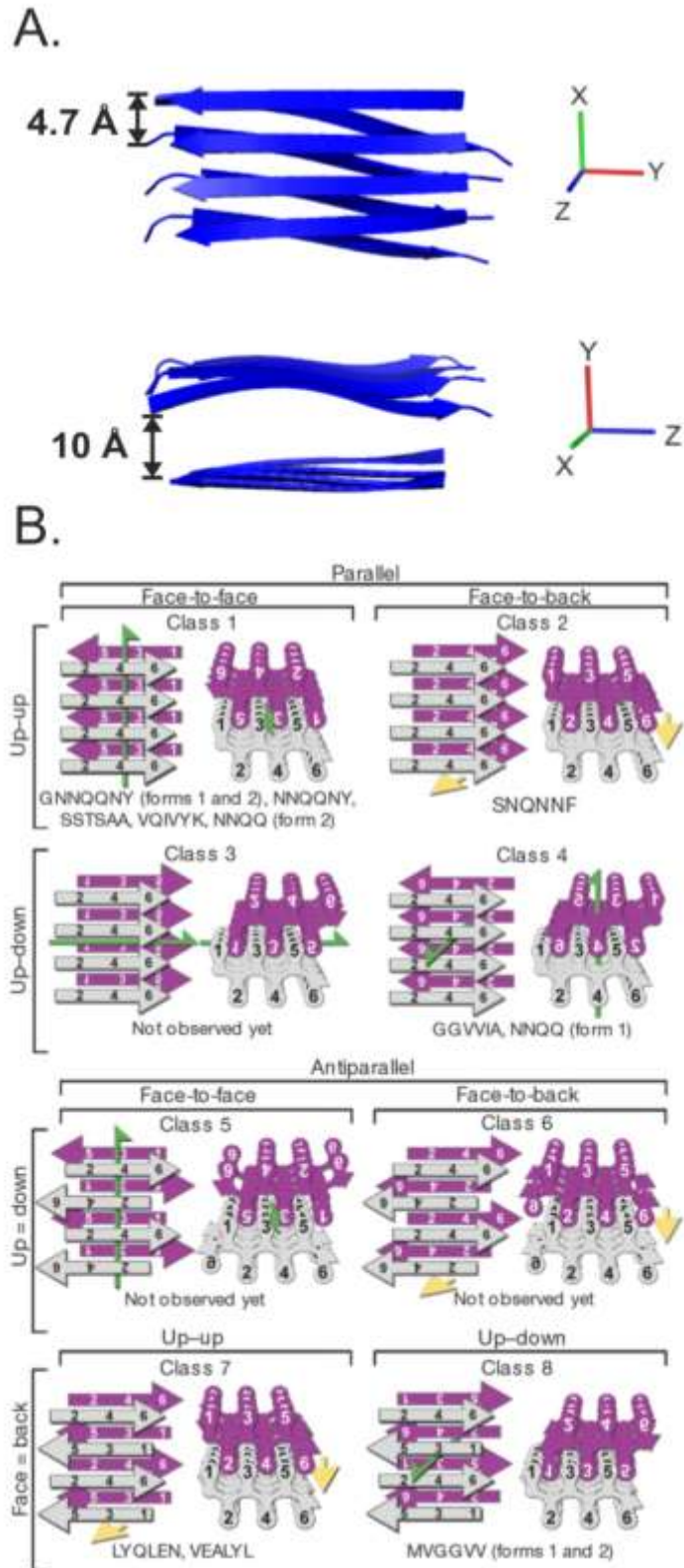
The recent development of structural and biophysical techniques has led to the elucidation of structural and biophysical characteristics.

##### **1.4.1. Cross- $\beta$ pattern**

Amyloid fibrils typically comprise rigid non-branching fibrils with an approximate diameter of 6-12 nm and exhibit a typical cross- $\beta$  pattern when analysed by X-ray fibre diffraction (XRFD) (Jahn et al., 2010, Serpell et al., 2000). The cross- $\beta$  pattern consists of a diffuse reflection at 4.7 Å along the meridian, suggestive of elongated protein chains running approximately perpendicular to the fibril long axis. Also a more diffuse reflection at 10 Å along the equator, indicative of sheets spaced 10 Å apart (Sunde et al., 1997). This cross- $\beta$  structure refers to the spine of the fibril and is common to all amyloid proteins, regardless of the primary sequence (Panel A, Figure 1.5).

### **1.4.2. Tertiary structure**

Within this cross- $\beta$  structure the  $\beta$ -sheet pairs can be orientated parallel or anti-parallel to each other (Balbach et al., 2000, Petkova et al., 2004). The structure is formed and stabilised by hydrogen bonding through the backbone amide groups (Makin, 2005). Furthermore, in the majority of structures obtained, the  $\beta$ -sheets lie 'in-register' with one another, i.e. identical side chains lie directly on top of one another. The structure is therefore further strengthened by hydrogen bonding through side chains containing amine groups. In such assemblies, the side chains are tightly inter-digitated into a 'steric zipper' (Sawaya et al., 2007). Eight classes of steric zipper have been described that denote: the face orientation and side chain orientations and whether the strands are parallel or anti-parallel (Panel B, Figure 1.5) (Sawaya et al., 2007).



**Figure 1.5 Amyloid cross- $\beta$  structure and steric zipper orientations. Schematic of the amyloid cross- $\beta$  structure (A). The eight classes of steric zipper taken from Sawaya, Sambashivan et al. 2007 (B).**

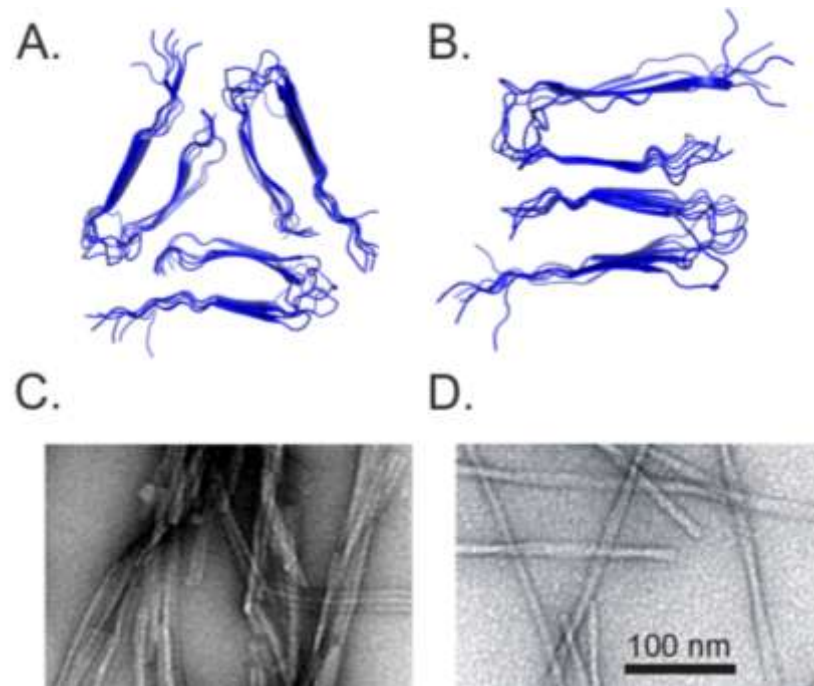
The most common of these arrangements is face to face (Panel B, class 1 Figure 1.5) (Eisenberg and Jucker, 2012) but other arrangements have been observed, in some cases, multiple steric zippers from the same protein (Lewandowski et al., 2011).

In accordance with the classifications described, 30 amyloid-forming proteins are defined as amyloid (Sipe et al., 2012) The field is further complicated by the study of peptide fragments, synthetic peptides and recombinant peptides which according to the current classifications should be referred to as ‘amyloid-like’ (Sipe et al., 2012). Accordingly, for the remainder of this thesis, fibril assemblies observed from recombinant or synthetic protein will be referred to as ‘amyloid-like’.

### **1.4.3. Quaternary structure of amyloid proteins**

Despite the uniformity of cross- $\beta$  structure in amyloid forming proteins there is variation in the quaternary structure. This quaternary structure is generated through intra molecular contacts between amino acid side chains that project out from the  $\beta$ -sheet assemblies. It has been shown that single amyloid proteins can have multiple quaternary structures, characterised *in vitro*. A good example of such a protein is amyloid- $\beta$  (A $\beta$ ), associated with Alzheimer’s disease. A $\beta$  is arguably the best characterised amyloid protein studied to date and much of our understanding about amyloid structure, aggregation and pathogenicity is derived from studies on A $\beta$ . Under different conditions A $\beta$  is capable of forming at least two very different morphologies. When incubated under quiescent conditions, A $\beta$  forms a triangular assembly with three-fold symmetry, however, when incubated with agitation it forms a structure with two-fold symmetry (Panels, A and B respectively, Figure 1.6) (Petkova et al., 2005b, Paravastu et al., 2008). Furthermore, a single point mutation in A $\beta$  results in a third quaternary assembly unlike those from the wild type

structures (Qiang et al., 2012). These structures will be discussed in more detail in chapter 5.

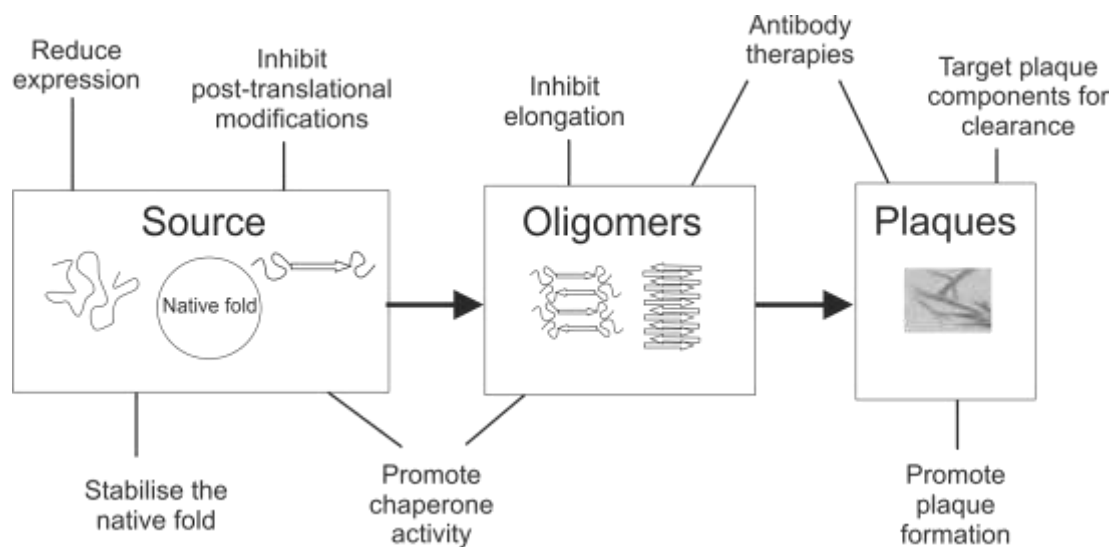


**Figure 1.6** Examples of quaternary structures of A $\beta$  determined by solid state NMR viewed down the fibril long-axis. A $\beta$  formed under quiescent (A) and agitated (B) conditions. TEM micrographs of the corresponding morphologies shown below. (C and D)

In turn these differences at the molecular level result in different fibril morphologies. Panels C and D (Figure 1.6) shown TEM images of two different A $\beta$  morphologies formed from under quiescent and agitated conditions respectively.

## 1.5. Therapies

Despite the prevalence and growing awareness of the amyloidoses there is no available cure and current treatments are often toxic or invasive and carry a high level of risk. Currently, finding new and alternative therapeutics capable of effectively targeting these devastating diseases is being prioritised. Figure 1.7 details some of the different strategies currently being explored.



**Figure 1.7 Schematic illustrating various strategies for targeting amyloid protein species therapeutically. Several approaches aim to reduce the amount of source protein that could go on to aggregate. Alternatively, it may be possible to prevent oligomer association in turn preventing fibril formation. Additionally, some therapeutics target the mature plaques. There is a further class of therapeutics that target the effects of amyloid deposition rather than the protein component.**

### 1.5.1. Targeting the source

The majority of current treatments against the systemic amyloidoses in particular, are centred on the removal of the precursor protein. The precursor can be either a wild type protein or a protein from which the amyloidogenic peptide is cleaved such as immunoglobulins, which are cleaved to generate free light chains. Patients suffering from AA amyloidosis are treated with anti-inflammatory drugs to reduce the levels of the causative protein serum amyloid A (Gillmore and Hawkins, 2006). Current treatments for AL amyloidosis involve administering highly toxic chemotherapeutic agents, such as melphalan and dexamethasone, and can cause severe side effects. Patients with hereditary amyloidoses often require organ transplantation to remove the source of the aberrant transthyretin protein (Dubrey et al., 2011).

It is clearly desirable to develop new therapeutic strategies that are less toxic or invasive.

### **1.5.1.1. Stabilising the native fold**

As described in section 1.3.1, amyloid formation can occur following the destabilisation of a protein's native fold. Therefore a possible strategy for drug design is to stabilise of the native fold, preventing monomer dissociation and amyloid formation. Bulawa *et. al.* (2012) have developed a small molecule to stabilise the native tetrameric fold of TTR, and reduce the number of available monomers for amyloidogenesis. This drug, Tafamadis™, is currently undergoing clinical trials and early results are promising (Bulawa et al., 2012).

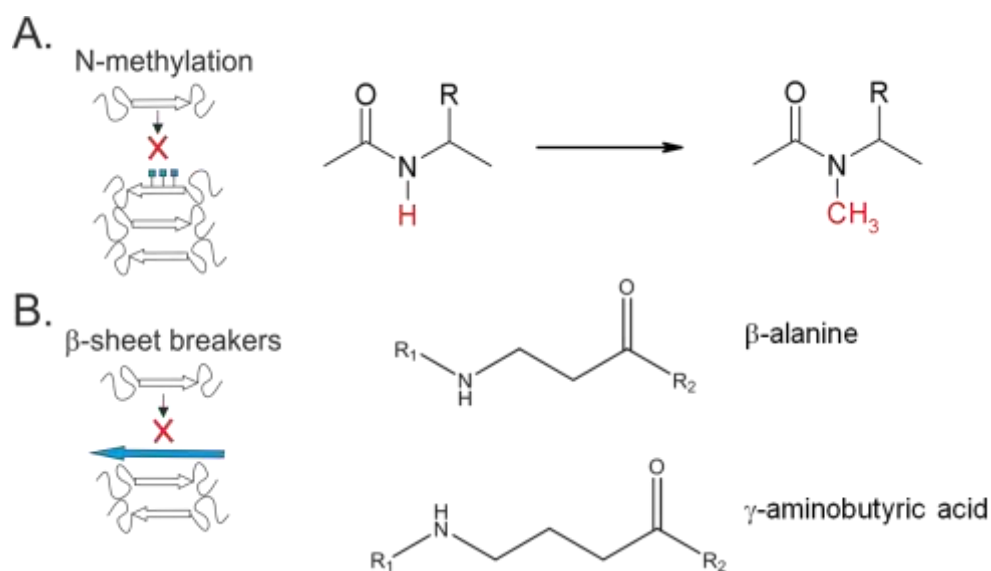
### **1.5.2. Inhibit fibrillogenesis**

A second strategy for reducing amyloid toxicity is to inhibit fibrillogenesis. This can take the form of targeting co-localising factors such as GAGs or the peptide assemblies themselves. As discussed previously, GAGs are often found to accelerate amyloid formation. The GAG mimetic agent eprodisate (Fibrillex™) binds to the GAG binding site on the amyloid protein serum amyloid A and prevents GAG binding and halts amyloidosis. Eprodisate (trade name Fibrillex™) reached stage II/III clinical trials, although without meeting the endpoint criteria, patients did show signs of improvement and delay in the onset of disease (Manenti et al., 2008). Further studies are being carried out on Fibrillex™ and GAG mimetics remain a plausible option for amyloid therapies.

A rational element can be introduced into the design of amyloid inhibitors, by targeting a sequence or region driving aggregation, known as the self recognition element (SRE). Following identification, peptide or small molecules can be designed to bind to this region and perturb or prevent further association of peptides, thus limiting elongation (Amijee et al., 2009). This strategy has been previously shown to be effective in inhibiting A $\beta$  fibrillogenesis and disassociating preformed fibrils *in*

*vitro*. A five residue peptide (LPFFD) was designed based on the amyloidogenic region (LVFF) to include a proline residue with the intention of disrupting  $\beta$ -sheet formation. Furthermore, it reduced deposition of A $\beta$  *in vivo* and blocked fibril formation in a rat brain model of Alzheimer's disease (Soto et al., 1998). Problems have occurred when some peptide based inhibitors have been identified as capable of forming amyloid.

An alternative strategy is to disrupt the hydrogen bonding network. Two methods of achieving this are outlined in Figure 1.8, N-methylation and peptide backbone extension. N-methylation prevents hydrogen bonding through steric inhibition, and blocking fibril extension. A schematic of this process is shown in panel A, Figure 1.8. N-methylated peptides to date have been successfully designed to inhibit A $\beta$  (Hughes et al., 2000, Amijee et al., 2012), IAPP (Kapurniotu et al., 2002) and  $\alpha$ -synuclein (Madine et al., 2008b) aggregation.



**Figure 1.8 Two amyloid inhibition strategies. Illustration of the N-methylation method and demonstration of the altered amide bond (A). Illustration of  $\beta$ -sheet breakers and example of  $\beta$ -sheet breaking amino acids (B).**

A different approach is to limit hydrogen bonding by extension of the peptide backbone. One such method described by Madine *et. al.* (2009), involves substitution



of a residue within the SRE with a  $\beta$ -sheet breaking alternative such as  $\beta$ -alanine or  $\gamma$ -aminobutyric acid (Panel B, Figure 1.8). This methodology has been successfully employed to generate inhibitory peptides to  $\alpha$ -synuclein and medin (Madine et al., 2009b, Madine and Middleton, 2010). Furthermore, one of the  $\alpha$ -synuclein inhibitors was also capable of inhibiting A $\beta$  and IAPP (Madine et al., 2009b).

Caution must be taken when inhibiting aggregation not to stabilise the formation of oligomeric species. As discussed earlier, the toxicity of oligomeric species has now been established and increasing the oligomer population may result in poorer prognosis. Results from Madine et al 2009 however demonstrated that inhibition of medin by extension of the backbone reduced the cytotoxicity of medin aggregates (Madine and Middleton, 2010).

### **1.5.3. Promote clearance of aggregates**

Given the variety of clinical signs and symptoms associated with the amyloidoses, diagnosis is frequently delayed and only occurs after large amyloid deposits have already formed. Consequently there is a drive to develop methods of clearing aggregates within the body. Currently, a plethora of antibodies have been developed that selectively bind different species along the aggregation pathway (Kayed et al., 2010, Yanamandra et al., 2011). This can result in the activation of phagocytic degradation and clearance of the amyloid deposits (Emadi et al., 2007).

Significant progress has been made recently targeting co-localising factors such as non-fibrillar plasma glycoprotein, serum amyloid P component (SAP) (Pepys et al., 2002).

Administering anti-human-SAP antibodies in mice with human SAP in association with fibrillar plaques, gave rise to a potent, complement-dependent, macrophage-

derived giant cell reaction that effectively eliminated amyloid plaques (Bodin et al., 2010). These therapies are currently undergoing further development.

#### **1.5.4. The importance of structural information in therapeutic design**

Recent progress in amyloid therapies has been significantly enhanced by the advances in structural techniques e.g. the rational design of Tafamadis™ (Bulawa et al., 2012). Techniques such as nuclear magnetic resonance (NMR) and X-ray crystallography provide atomic resolution detail of amyloid precursors and more recently fibrils which are invaluable in helping design effective small molecule modulators (Tycko, 2006a, Madine et al., 2008b).

### **1.6. Cardiac amyloid**

Virtually all systemic amyloidoses have, to some extent, cardiovascular manifestations. Cardiac involvement is most common in immunoglobulin-derived and senile systemic amyloidoses (Kholová and Niessen, 2005) and leads to a much poorer prognosis when compared with diseases with different organ involvement (Dubrey et al., 2011). The symptoms of cardiac amyloidosis include: restrictive cardiomyopathy, bilateral dilation of the atria, arrhythmia, and frequently congestive heart failure (Falk and Dubrey, 2005). Due to the non-specific nature of the symptoms, cardiac amyloidosis is often mis-diagnosed and, as a result, patients are often in the late stages of the disease when a final diagnosis is made (Kholová and Niessen, 2005). Diagnosis can be then further complicated by the need to accurately determine the amyloidogenic protein precursor responsible in order to establish a treatment programme. Once a patient is suspected of cardiac amyloidosis a definitive diagnosis is typically obtained by a combination of methods such as echocardiography, electrocardiography, cardiac biomarker screening and

radionuclide imaging, followed by histological immunoanalysis and increasingly mass spectrometry (Hassan et al., 2005, Banypersad et al., 2012).

There are principally five amyloidoses that affect the heart: a group of hereditary amyloidoses that result from mutations in TTR, ATTR, AA amyloidosis, A $\beta$ 2M amyloidosis and AL amyloidosis (Hassan et al., 2005, Banypersad et al., 2012).

### **1.6.1. ATTR amyloidosis**

ATTR amyloidosis is caused by aggregation of transthyretin (TTR). ATTR amyloidosis can be either hereditary or acquired.

#### **1.6.1.1. Hereditary ATTR amyloidoses**

There are multiple forms of familial amyloidosis, many of which are the result of single point mutations in the transthyretin (TTR) gene. There are over 100 mutations of TTR currently understood to form amyloid (Connors et al., 2000). The most prevalent variation is a V122I mutation which is present in 3.9 % of the African American population (Jacobson et al., 1996). Cardiac amyloidosis is most associated with V122I and T60A mutations (Dungu et al., 2012).

#### **1.6.1.2. Acquired ATTR amyloidosis**

ATTR amyloidosis is caused by the aberrant aggregation of wild-type TTR (Westermarck et al., 1990). It almost exclusively affects men over 70 years of age. Deposits are almost always limited to the cardiac tissue, only occasionally presenting with carpal tunnel syndrome 3-5 years preceding cardiac symptoms (Falk and Dubrey, 2005).

### **1.6.2. AA amyloidosis**

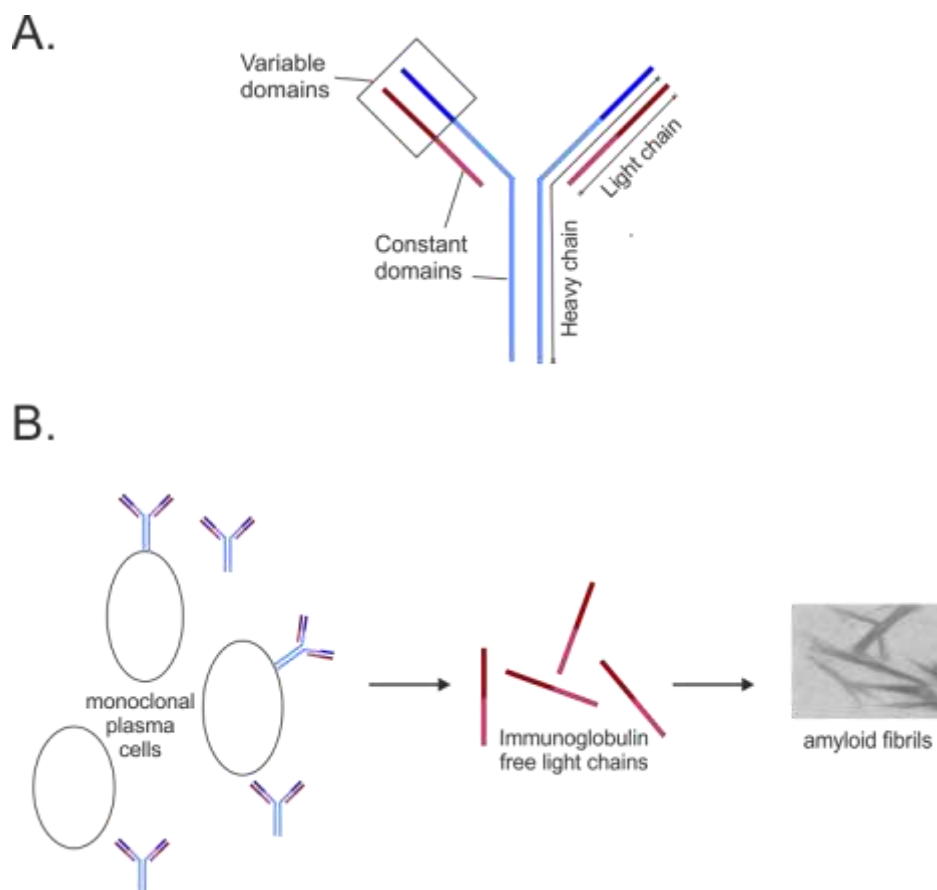
AA amyloidosis is a relatively rare complication of chronic inflammatory conditions such as inflammatory bowel disease and rheumatoid arthritis. It occurs when fragments of serum amyloid A protein are deposited in tissues as amyloid plaques. Deposits are found in the heart in only 2 % of cases, and the incidence is continually reducing, probably reflecting the improved treatments of the inflammatory diseases themselves.

### **1.6.3. A $\beta$ 2M amyloidosis**

A $\beta$ 2M amyloidosis is experienced by few patients receiving long term dialysis treatment and results in the accumulation of amyloid plaques composed of  $\beta$ 2-microglobulin in the joints and , very rarely, the cardiac tissue (Gal et al., 1994, Gejyo et al., 1985).

### **1.6.4. AL amyloidosis**

AL amyloidosis, also known as primary amyloidosis, is the most common form of systemic amyloidosis with an incidence of 8.9 cases per million of the population (Kyle et al., 1992). This statistic, however, is probably an underestimate given the difficulties described earlier in obtaining a correct diagnosis. In the majority of patients, more than one organ is affected. Heart involvement portends the worst prognosis (Kyle and Gertz, 1995). Death is attributed to heart involvement in approximately 50 % of cases, the leading causes of death being congestive heart failure and arrhythmia (Kyle and Gertz, 1995). AL amyloidosis is caused by the deposition of the immunoglobulin light chains produced by clonal plasma cells in the visceral organs, primarily the kidneys, heart and peripheral nervous system (PNS) (Kholová and Niessen, 2005).



**Figure 1.9 Schematic of immunoglobulin structure (A) and the AL amyloidosis pathway (B).**

A characteristic of AL amyloidosis is the predominance of the  $\lambda$  over  $\kappa$  light chain isoforms in a ratio of 3:1, the opposite of that observed in multiple myeloma (Falk, 1998). AL amyloidosis is a very heterogeneous disease; both symptoms and organ involvement vary greatly between patients. One explanation for this variety can be found in the plaque composition; no patient has yet been found to have the same light chain component as another (Enqvist et al., 2009). The light chain protein is usually cleaved, but each patient typed to date has varying fragment size. The majority of light chain proteins consist of the N-terminal variable region and different amounts of the constant region. On rare occasions plaques have been observed that consist of full length light chains and those containing only constant regions (Olsen et al., 1998, Buxbaum, 1992). It has been suggested that the C-terminal region needs to be

cleaved to initiate fibrillogenesis, but this does not explain the inclusion of constant regions in amyloid plaques in some patients. In a recent study Enqvist *et. al.* (2009) analysed the proteolytic pattern of light chains from six patients in an attempt to pinpoint the origin of the fragmented species. They concluded that proteolytic cleavage is likely to occur after fibril formation and at the site of deposition. These findings were based on further evidence that all light chain components could be identified within deposits and fragmentation patterns observed were conserved within a patient, but differed between patients (Enqvist et al., 2009).

### 1.6.4.1. Representative immunoglobulin light chain- SMA

The heterogeneity of AL amyloidosis makes it especially difficult to study and characterise. In an attempt to clarify the differences between non-amyloidogenic and amyloidogenic light chains, Stevens *et al* (1995) isolated immunoglobulin light chain variable region proteins from three patients (Stevens *et al.*, 1995b). Two of these proteins, REC and SMA were extracted from patients with amyloid deposits. The third protein, LEN was isolated from the urine of a patient with no apparent amyloid deposits.

```
SMA  DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNRNYLAWYQQKLGQPPKLLIYWASTR 60
LEN  DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNSKNYLAWYQQKPGQPPKLLIYWASTR 60
REC  DIVMTQSPDSLAVSPGERATINCKSSQNLLDSSFDNTLAWYQQKPGQPPKLLIYWASSR 60
*****  *****.:* ** . * ***** *****.:*

SMA  ESGVPDRFSGSGSGTDFTLTISSLAEDVAVYYCHQYYSHPQTFGQGTKLELKR 114
LEN  ESGVPDRFSGSGSGTDFTLTISSLAEDVAVYYCQYYSTPYSFGQGTKLEIKR 114
REC  ESGVPDRFSGSGSGTDFTLTISSLAEDVAVYYCQYYSTPPTFGGKVEIKR 114
*****:***** * :** ***:***
```

**Figure 1.10 Sequence alignment of the three proteins extracted by Stevens *et. al*. The differences between sequences are highlighted in red.**

Synthetic genes were generated for the three variable light chain portions of LEN, REC and SMA. Subsequently the genes were incorporated into pASK40 vector for recombinant expression in *E.coli*. Recombinant production of these proteins has provided a means to compare and contrast the amyloidogenic REC and SMA proteins against the non-amyloidogenic LEN. SMA differs from LEN by 8 residues. Arguably the most significant of these is at position 40, where a proline in LEN and REC is replaced by the hydrophobic residue leucine in SMA. It is possible that this substitution of a  $\beta$ -sheet breaking residue, for a hydrophobic residue, marks the difference in amyloid-forming capabilities. The vast majority of light chain sequences examined have a proline at position 40. Furthermore, there are currently



no non-amyloidogenic light chains that possess a hydrophobic residue at that position. These peptides can be used to identify the differences between amyloidogenic and non-amyloidogenic light chains. SMA is used, in this thesis and elsewhere, as a model system for AL amyloidosis (Meng et al., 2008). Further discussion about the aggregation and structural properties of SMA can be found in chapter 7.

## **1.7. Amyloid in the vasculature**

In addition to the myocardium, amyloid deposits are frequently reported in the surrounding vasculature. Transthyretin and serum amyloid A have been found in the aorta and less frequently in small vessels (Mucchiano et al., 2001, Pitkänen, 1984, Cornwell and Westermark, 1980). Several groups have identified a distinct form of amyloid residing predominately in the medial layer of the aorta (Iwata et al., 1978, Cornwell, 1982). Mucchiano *et. al.* (1992) examined aortic tissue samples obtained at autopsy from 84 individuals and noted the level, distribution and type of amyloid deposits (Mucchiano et al., 1992). Haggqvist *et. al.* (1999) then went on to characterise the isolated protein using a variety of techniques, resulting in the identification of a 5.5 kDa protein derived from an internal cleavage fragment of the glycoprotein lactadherin. This protein, which comprises the focus of this thesis, was named medin (AMed), based on its location in the arteries (Haggqvist et al., 1999).

### **1.7.1. Medin**

#### **1.7.1.1. Lactadherin – the precursor**

A combination of sequence analysis and alignment has revealed that medin is likely to be an internal cleavage product of lactadherin (Haggqvist et al., 1999). Lactadherin was first described as a breast antigen (BA46) against the human milk

fat globule membrane (Ceriani et al., 1977). Antibodies to BA46 were raised and developed to target human breast tumours (Ceriani and Blank, 1988, Ceriani et al., 1977).

Lactadherin is a 46 kDa protein divided into three domains spanning 364 amino acids (Panel B, Figure 1.11). The C-terminal portion of the protein has been shown to have strong homology with serum factors VIII and V, C1 and C2 domains (Larocca et al., 1991). These domains are predominantly responsible for membrane binding and in particular binding of phosphatidylserine though the C2 domain (Andersen et al., 2000). The N-terminal section of lactadherin was later characterised as an epidermal growth factor (EGF)-like domain which contained the cell adhesion motif, arginine-glycine-aspartic acid (RGD) (Couto et al., 1996).

The medin sequence is located within the C2 domain of lactadherin. The process by which medin is excised from lactadherin remains unknown but it seems likely that it is a result of enzymatic cleavage rather than alternative splicing. This conclusion arises from the isolation of medin of different lengths with a ragged N-terminus (Häggqvist et al., 1999). The 50 amino acid peptide represents the core medin sequence.

The structure of the bovine C2 domain has been solved to 1.7 Å and 2.4 Å using X-Ray crystallography (Shao et al., 2008, Lin et al., 2007). It consists of a distorted  $\beta$ -barrel core made up of 8 major  $\beta$ -strands and a further 11 short minor  $\beta$ -strand regions which pack above and below the barrel. Additionally there are three relatively large loops at one end described as spikes 1, 2 and 3. These spikes contain hydrophobic amino acids that share similarity with the membrane interacting regions (Shao et al., 2008). Medin (highlighted red, Figure 1.11) encompasses 2 major strands and part of a third, and 11 minor strands separated by spikes 2 and 3 (Shao et

al., 2008). The C2 domains of serum factors VIII and V exhibit conformational changes upon membrane binding (Macedo-Ribeiro et al., 1999). The similarity between the C2 domain of lactadherin and these serum factor domains may suggest that it too undergoes a conformational change upon contact with lipids. This may result in partial unfolding of the domain structure, exposing medin to enzymatic cleavage.

The specific function of lactadherin is relatively unknown, it is expressed in many cell lines and seems capable of many different functions including: carcinoma antigen (see above) (Ceriani et al., 1977), phagocytosis (Hanayama et al., 2002), rota virus protection (Newburg et al., 1998), anti-coagulation (Shi and Gilbert, 2003) and elastin interactions (Larsson et al., 2006). The most important of these functions for the context of this work are arguably its anti-coagulation mechanisms and elastin binding properties.

Lactadherin has been shown to be expressed in smooth muscle cells and co-localise to the elastic structures of arteries (Häggqvist et al., 1999). Further studies by Larsson *et. al.* (2006) demonstrated that lactadherin can bind to elastin and that the binding is mediated through the medin (C2) domain (Larsson et al., 2006). Binding to elastin and a subsequent conformational change may represent a second mechanism by which medin is exposed and cleaved from lactadherin.

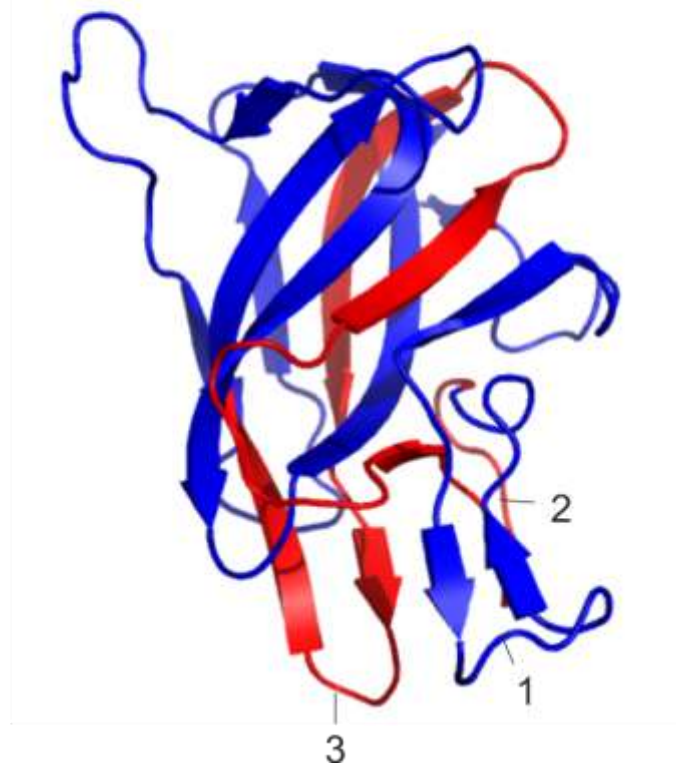
A

```
LDICSKNPCH NGGLCEEISQ EVRGDVFPSY TCTCLKGYAG NHCETKCVEP
LGMENGNIAN SQIAASSVRV TFLGLQHWVP ELARLNRMGM VNAWTPSSND
DNPWIQVLL RRMWVTGVVT QGASRLASHE YLKAFKVAYS LNGHEFDPIH
DVNKKHKEFV GNWNKNAVHV NLFETPVEAQ YVRLYPTSCH TACTLRFELL
GCELNGCANP LGLKNSIPD KQITASSSYK TWGLHLFSWN PSYARLDKQG
NFNAWVAGSY GNDQWLQVDL GSSKEVTGII TQGARNFGSV QFVASYKVAY
SNDSANWTEY QDPRTGSSKI FPGNWDNHSK KKNLFETPIL ARYVRILPVA
WHNRIALRLE LLGC
```

B



C

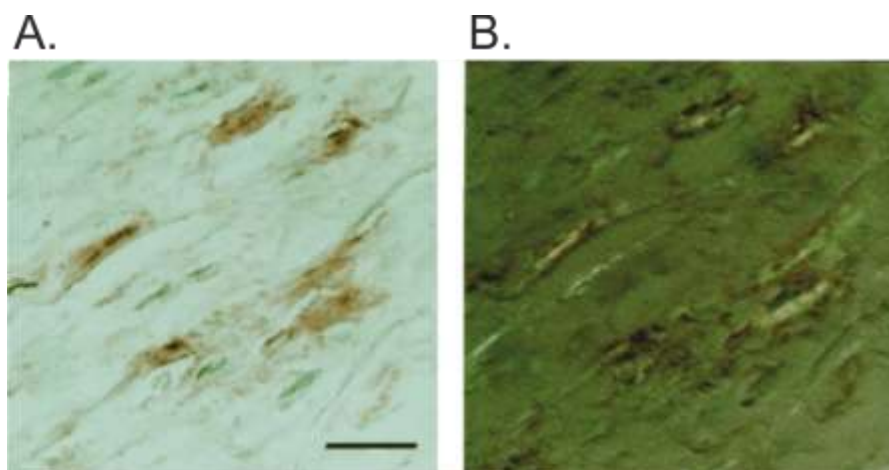


**Figure 1.11** Sequence of lactadherin with medin sequence highlighted in red (A) Schematic of lactadherin domain structure highlighting the location of medin within the C2 domain (B) Crystal structure of lactadherin C2 domain from *Bos taurus* adapted from Shao *et. al.* (2007) (PDB I.D. 3BN6) Red highlights the location of the medin sequence within the domain. The hydrophobic loop regions (spikes) hypothesised to interact with membranes are numbered 1-3 (C).

### **1.7.1.2. Prevalence and Distribution**

Medin amyloid is the most common form of human amyloid identified to date. Multiple studies have reported on the incidence, type and location of medin (or senile aortic amyloid) deposits in the vasculature. Similar to senile systemic amyloidosis it is more prevalent with age. A report by Schwartz in 1970 found medin in all aortas from donors over the age of 50 (Schwartz, 1970). Subsequently similar figures have been described. Mucchiano *et. al.* showed that 97% of aortas from persons over 50 tested positive for medin, opposed to 1 out of 21 aortas from younger donors (Mucchiano et al., 1992). Further reports have set the prevalence lower at 79 % and 68 % by Iwata *et. al.* and Wright *et. al.* respectively. Allowing for these differences, medin remains the most common form of human amyloid.

Medin deposits are predominately found in the aorta but it has been shown to localise to other large arteries, mainly in the upper half of the body (Peng et al., 2005, Peng et al., 2002). Outside the aorta, medin was most commonly found in the temporal, carotid and basilar arteries, and less frequently in the superior mesenteric and the left coronary artery (Peng et al., 2005). Medin has never been found to be present in veins.



**Figure 1.12** Section of aortic media stained with antibodies specific to residues 19-26 of medin. Immunoreactive areas stained with diaminobenzidine (A) and subsequently stained with Congo Red and visualised under polarised light (B). Taken from Häggqvist, Näslund et al. 1999.

When stained with Congo Red, medin displays the typical apple green birefringence associated with amyloid formation, although it stains weakly in comparison with some other amyloid deposits (Peng et al., 2002). This may account for some of the variability in the reported incidences. It commonly appears as nodules and thin streaks in all three layers of the artery but predominantly in the inner half of the tunica media. Intimal deposits were found in 35 % of persons over the age of 50 compared with 2 % for adventitial deposits (Mucchiano et al., 1992). Medin is frequently described as being closely associated with the internal elastic laminae (Cornwell, 1982). While the majority of medin deposition is extracellular, some small deposits have also been identified within smooth muscle cells (Cornwell Iii et al., 1983, Cornwell, 1982).

### **1.7.1.3. Consequences of medin deposition**

To date no definitive consequences of medin deposition have been confirmed. The location of medin deposition has led researchers to investigate possible links with other vascular diseases. Sporadic thoracic aneurysm and dissection are life-

threatening conditions with unknown pathogenesis. They commonly present with signs of degeneration of the elastic structures and GAG deposition. This deterioration of the tissue structure has been observed in other amyloid pathologies and therefore led researchers to investigate a possible role for medin in these conditions. Specimens were taken from the thoracic ascending aorta of patients with sporadic thoracic aneurysm or dissection and from a control group and analysed for the amount of amyloid and medin immunoreactivity. Congoophilic material was observed in all three groups, demonstrating medin immunoreactivity, with further granular deposits that were negative for Congo Red but immunoreactive for medin. These granular deposits were more numerous in the disease samples. It was, therefore, concluded that these deposits represented oligomeric/protofibrillar medin amyloid, consistent with theories suggesting that the oligomeric intermediate species are more toxic than the mature fibrils. Furthermore, the study investigated the relationship between gelatinase expression and medin, showing that proteases were up regulated in presence of medin.

## **1.8. Aims**

The overall aim of this thesis is to gain a detailed understanding of the amyloidogenic and structural properties of medin, the main component of aortic medial amyloid, and SMA, a monoclonal amyloid protein underlying cardiovascular amyloid disease.

This will be achieved by addressing five key objectives:

1. To examine the use of model peptides within amyloid research and establish whether the medin peptide comprising residues 42-49 (Med<sub>42-49</sub>) is a good

model system for studying the aggregation properties of full length medin. Furthermore, chapter 2 investigates the role of predicted aggregation ‘hot spots’ within medin and how they could be targeted for therapeutic design

2. To establish a protocol for the expression and purification of full length medin in *E.coli*. This will provide a replenishable source of peptide for biophysical studies. Moreover, it remains necessary to develop reliable methods of producing isotopically labelled medin for NMR structural studies.
3. To characterise the aggregation properties of medin using biophysical techniques. Very little is known about the aggregation properties of medin and the aim of this work is to establish some of the basic characteristics of medin aggregation that can be used to enhance further studies of this peptide.
4. To investigate the possibility of a stabilising salt bridge in medin fibrils using both direct measurements using SSNMR and characterisation of salt bridge mutants.
5. To establish, for the first time, a protocol of the expression of isotopically labelled SMA to enable preliminary structural studies of SMA and form the basis for more detailed future investigations.



## 2. Investigating the packing arrangements of medin fibrils

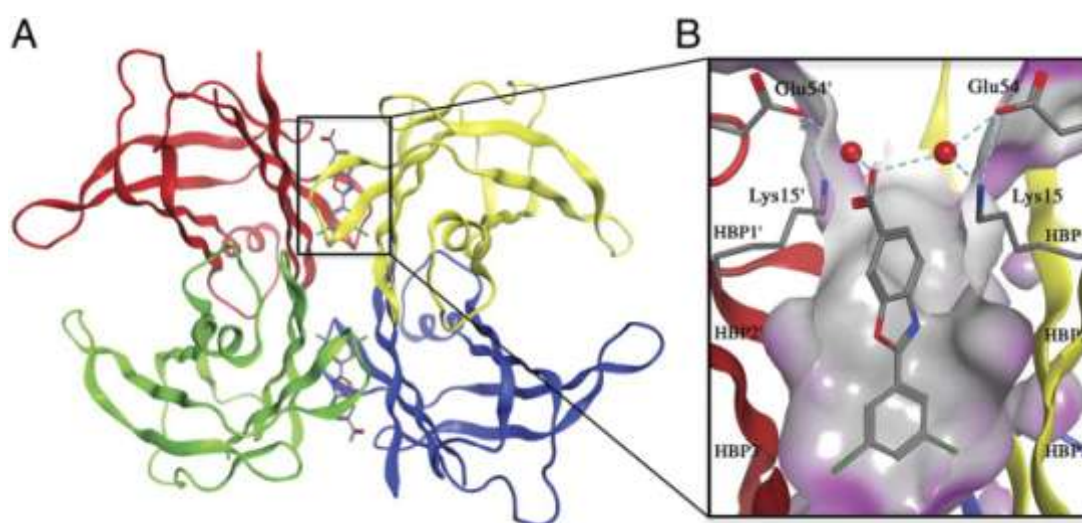
### 2.1. Introduction

Gaining a greater understanding of amyloid molecular structures is vitally important for a variety of reasons including identification of self recognition elements which could be targeted by therapeutics (as discussed in section 1.5.2), understanding the relationship with amyloid toxicity (Van Nostrand et al., 2001) and to elucidate the nature of intermediate species along the aggregation pathway (Fändrich, 2012).

Amyloid fibrils, by their insoluble, non-crystalline nature, are difficult to study by conventional methods of structural biology. Advances in high-resolution structural techniques such as solid-state nuclear magnetic resonance (SSNMR) and X-ray crystallography have greatly enhanced our knowledge of amyloid structures but nevertheless there remain only a handful of full length molecular models available, these are namely: polymorphs of A $\beta$ <sub>1-40</sub> fibrils (Petkova et al., 2002, Paravastu et al., 2008) and fibrils from a mutated variant (D25N) (Qiang et al., 2012), IAPP fibrils (Luca et al., 2007) and the native tetramer of transthyretin (TTR).

Structural information has been integral to the design and refinement of a therapeutic agent against the acquired amyloidosis ATTR (Klabunde et al., 2000). Elucidation of the native tetramer structure of TTR led to the rational design of a stabilising small molecule that prevents the dissociation of monomers and thus amyloid formation (Figure 2.1) (Bulawa et al., 2012). TTR is a major carrier of thyroxine T<sub>4</sub> in rodents however the binding sites in human TTR are rarely occupied and therefore available for small ligand binding. The TTR tetramer has two dimer-dimer interfaces; the thyroxine binding site is located on at the *z*-axis interface. The binding pocket

consists of two subsites. Many rounds of optimisation resulted in a compound comprising two differentially substituted aryl rings. Each of the rings occupies one of the subsites. Tafamidis™ or 2(3,5-dichloro-phenyl)-benzoxazole-6-carboxylic acid is capable of stabilizing the dimer-dimer interface through a combination of specific hydrophobic residues and electrostatic interactions. Tafamidis™ is now in clinical trials and early results are promising (Bulawa et al., 2012). This therapeutic strategy would not have been possible without high resolution molecular detail.



**Figure 2.1** Crystal structure of Tafamidis bound to TTR (PDB ID: 3TCT). 3D ribbon diagram depiction of the TTR tetramer with tafamidis bound (A). The four TTR monomers are individually coloured. Magnified image of tafamidis bound in one of the T<sub>4</sub>-binding sites (B). Taken from (Bulawa et al., 2012).

This therapeutic strategy targeted the source of amyloidogenic protein (the wild type tetramer). There is however a great interest in targeting fibrils themselves, in order to break up existing aggregates as well as preventing formation.

In order to overcome some of the challenges surrounding structural studies of these large insoluble protein aggregates, many studies use small peptides or fragments as model systems (Balbach et al., 2000, Madine et al., 2008c). Previous studies of such model systems have shown that aggregate structure is highly sensitive to alterations

in sequence, particularly via truncation or mutation, and may be completely different from the parent molecule from which they are derived. For example, fibrils formed by the 40-residue A $\beta$  (A $\beta_{1-40}$ ) comprise parallel and in-register hydrogen bonded  $\beta$ -strands (Balbach, 2002, Petkova et al., 2002). This organisation is reproduced in fibrils of the 26-residue peptide fragment A $\beta_{10-35}$  (Burkoth et al., 2000), whereas the shorter peptides A $\beta_{16-22}$  and A $\beta_{11-25}$  adopt an anti-parallel  $\beta$ -strand arrangement within fibrils, with a pH-dependent registry in the case of A $\beta_{11-25}$  (Balbach et al., 2000, Petkova et al., 2004). Hence, it cannot be assumed that a smaller fragment is a suitable model of the parent peptide.

The first study on medin protein was performed on protein extracted from aortic tissue (Häggqvist et al., 1999). During analysis it was noted that an octapeptide, generated as a result of tryptic digest, with the sequence NH<sub>2</sub>-NFGSVQFV-COOH was capable of forming amyloid-like aggregates in isolation. It was later noted that this octapeptide shares sequence similarity with fragments of the islet amyloid polypeptide (IAPP) also known as amylin, NH<sub>2</sub>-NFGAIL-COOH and NH<sub>2</sub>-NFLVH-COOH. Further work was then carried out to identify the minimum motif required for fibril formation, which was NH<sub>2</sub>-NFGSV-COOH (Reches and Gazit, 2004).

A more comprehensive study on medin aggregation was later carried out by Larsson *et al.* (2007). Using a variety of synthetic and recombinant fragments and full length medin, they tested them for their ability to bind Congo Red, Thioflavin T and to form fibrillar structures visible by electron microscopy. They indicated that the C-terminal 18-19 residues comprised the amyloid promoting region (Larsson et al., 2007).

Recent work by Madine *et al.* (2009) investigated the structural properties of a peptide representing the C-terminus of medin, comprising residues 42-49 (Med<sub>42-49</sub>).

This peptide has previously been shown to be highly amyloidogenic in isolation

(Larsson et al., 2007). Madine *et al* (2009) showed that this peptide aggregates within two days to form bundles of microcrystalline-like needles displaying a high degree of order. Using  $^{13}\text{C}$  selectively labelled peptides coupled with solid-state nuclear magnetic resonance (SSNMR) methods they investigated the molecular architecture of the resultant amyloid-like aggregates (Madine et al., 2009a). They used rotational resonance (RR) experiments to determine constraints on distances between selective atomic sites within the fibrils. The data was consistent with aggregates formed from unbroken  $\beta$ -strands hydrogen bonded in a parallel in-register arrangement within  $\beta$ -sheets. Moreover, additional RR experiments identified close ( $> 6.5 \text{ \AA}$ ) contacts between residues F43 and V46 and between S45 and V46. The constraints determined from SSNMR experiments were combined with X-ray fibre diffraction data to generate a structural model for the cross- $\beta$ -spine of Med<sub>42-49</sub> (Madine et al., 2009a).

These data have led to a much greater understanding of the structure of Med<sub>42-49</sub> aggregates but it is not known how relevant the structure of the model peptide fibrils is to the structure of full length medin.

The aim of this work is to investigate how the length of the medin affects the rate of aggregation, morphology and structure, and thus whether Med<sub>42-49</sub> is a good model for studying the structure of full length fibrils.

## **2.2. Materials and methods**

### **2.2.1. Overall strategy**

The characteristics and fibril structure of Med<sub>43-49</sub> investigated by Madine *et al* (2009) was compared with a longer medin sequence and full length medin with the aim of identifying any differences in their aggregation kinetics, morphology and structure.

### **2.2.2. Amyloid prediction servers**

Amyloid prediction servers predict aggregation prone regions based on the physicochemical properties of a peptide's constituent amino acids. They were used here to identify amyloidogenic regions in the medin sequence which, in turn guided the design of the intermediate length medin peptide. Three different servers were used to obtain a consensus prediction and provide more reliable results. If two or more servers predicted a particular area sequence it was considered to be amyloidogenic.

#### **2.2.2.1. Zyggregator**

The Zyggregator server was developed following the work of Chiti *et al* (2003) which demonstrated that the effect of mutations on aggregation propensity was largely down to the physicochemical properties of the substituted residues themselves. Properties such as hydrophobicity, secondary structure propensity and charge all help to determine the aggregation potential of a sequence (Chiti *et al.*, 2003). This work led to the development of the Zyggregator server which calculates an aggregation profile of a sequence based on the properties of its constituent amino acids (Tartaglia and Vendruscolo, 2008). An input sequence is entered into the server in FASTA format and the results are displayed in a variety of formats. Any residues

with a score above 1 are thought to promote aggregation of the sequence. This method allows the user to calculate aggregation profiles of an input sequence at different pH values. The aggregation profile of medin was calculated at pH 7.4.

#### **2.2.2.2. AGGRESCAN**

The AGGRESCAN server also calculates an aggregation propensity value for each residue within the input sequence and provides a graphical output of the results. Areas of the profile above a pre-calculated threshold are termed aggregation ‘hot spots’(Conchillo-Solé et al., 2007).

#### **2.2.2.3. Prediction of Amyloid Structure Aggregation (PASTA)**

Unlike Zyggregator and AGGRESCAN, the PASTA algorithm predicts which portions of a given input sequence stabilises the corresponding cross- $\beta$  structure and hydrogen bond networks (Trovato et al., 2007). The server is based on the assumption that the same stabilising hydrogen bonds found in globular proteins, stabilise amyloid structures (Trovato et al., 2006).

#### **2.2.3. Preparation of medin peptide samples for NMR studies**

Synthetically produced medin peptides were purchased from Protein Peptide Research (Cambridge, U.K.) with selective uniform  $^{13}\text{C}$  and  $^{15}\text{N}$  isotope labels (indicated with asterisk in Figure 2.4).

Lyophilised medin peptides were initially dissolved in hexafluoroisopropanol (HFiP), bath sonicated for 10 minutes and then the HFiP removed under a stream of nitrogen. The samples were then placed on a high vacuum overnight to remove the last remaining traces of HFiP. This process was repeated three times prior to resuspension in 10 % dimethylsulphoxide (DMSO) 90 %  $\text{H}_2\text{O}$  to induce aggregation. Medin was incubated at a concentration of 200  $\mu\text{M}$  with agitation at room

temperature for up to 21 days. These conditions were selected to mirror the conditions used by Madine *et al* (2009) thus enabling direct comparison of the three peptides. Furthermore, at lower DMSO concentrations both peptides became insoluble and precipitated. Moreover, the *in vivo* concentration of medin is unknown and therefore a high concentration was utilised to ensure optimal conditions for NMR experiments. The peptide suspension was then centrifuged at 21,000 g for 1 hour to generate a tightly packed pellet before being transferred to a zirconium 4 mm rotor with a Kel-F cap (Bruker, U.K.) in preparation for NMR studies.

#### **2.2.4. Characterisation of medin peptide aggregation**

##### **2.2.4.1. Aggregation studies of medin peptides by thioflavin T**

Both peptides were incubated at a concentration of 200  $\mu\text{M}$  with agitation at room temperature for up to 21 days. Aliquots (25  $\mu\text{L}$ ) of the suspensions were removed at 0, 2, 4, 7 and 14 days and frozen. Following completion of the time course aliquots were defrosted and added to 180  $\mu\text{L}$  of 10  $\mu\text{M}$  ThT solution. Emission spectra were recorded on a Cary Eclipse Varian fluorescence spectrometer operating at room temperature. Spectra were recorded from 450-600 nm with an excitation wavelength of 450 nm and a 5 nm band pass. Values at 482 nm were recorded and each time point represented as a percentage of the end point fluorescence.

##### **2.2.4.2. Morphology of medin peptides by transmission electron microscopy.**

The resultant peptide aggregates were visualised using TEM. Peptide suspensions (10  $\mu\text{L}$ ) were loaded onto carbon coated copper grids held between tweezers. Excess solution was removed by placing the grid perpendicular to blotting paper. The grid was then inverted onto a droplet of  $\text{dH}_2\text{O}$  to wash of excess peptide aggregates. The grid was blotted again before being inverted onto a droplet of 4 % uranyl acetate to

negatively stain the aggregates. Excess stain was removed by blotting before allow the grid to air dry prior to visualisation. The grids were observed using a Technai 10 electron microscope operating at 100 kV located in the EM Unit in the physiology building at the University of Liverpool.

### **2.2.5. Solid state nuclear magnetic resonance (SSNMR) studies on synthetic medin peptides.**

All cross-polarization magic-angle spinning (CP-MAS)  $^{13}\text{C}$  SSNMR experiments on the synthetic medin peptides were carried out on a Bruker Avance 400 spectrometer operating at a magnetic field of 9.3 Tesla located in the NMR Centre for Structural Biology at the University of Liverpool. The peptide samples were rotated at the magic angle at rates between 6 kHz and 13 kHz ( $\pm 1$  Hz). All experiments utilized an initial 4.0- $\mu\text{s}$   $^1\text{H}$   $90^\circ$  excitation pulse length, 1-ms Hartmann-Hahn contact time at a  $^1\text{H}$  nutation frequency of 65 kHz, two pulse phase modulated (TPPM) proton decoupling (Bennett et al., 1995) at a field of 85 kHz during signal acquisition and a 2-s recycle delay.

#### **2.2.5.1. Dipolar assisted rotational resonance (DARR) experiments**

Two-dimensional  $^{13}\text{C}$ - $^{13}\text{C}$  broadband dipolar re-coupled spectra were obtained using the dipolar-assisted rotational resonance (DARR) experiment (Takegoshi et al., 2001). Spectra were recorded with 128 hypercomplex points in the indirect dimension ( $t_1$ ), with a mixing time of 20 ms during which the proton nutation frequency was adjusted to the spinning frequency of 8 kHz. Chemical shift values were referenced to external adamantane at 37.8 ppm.

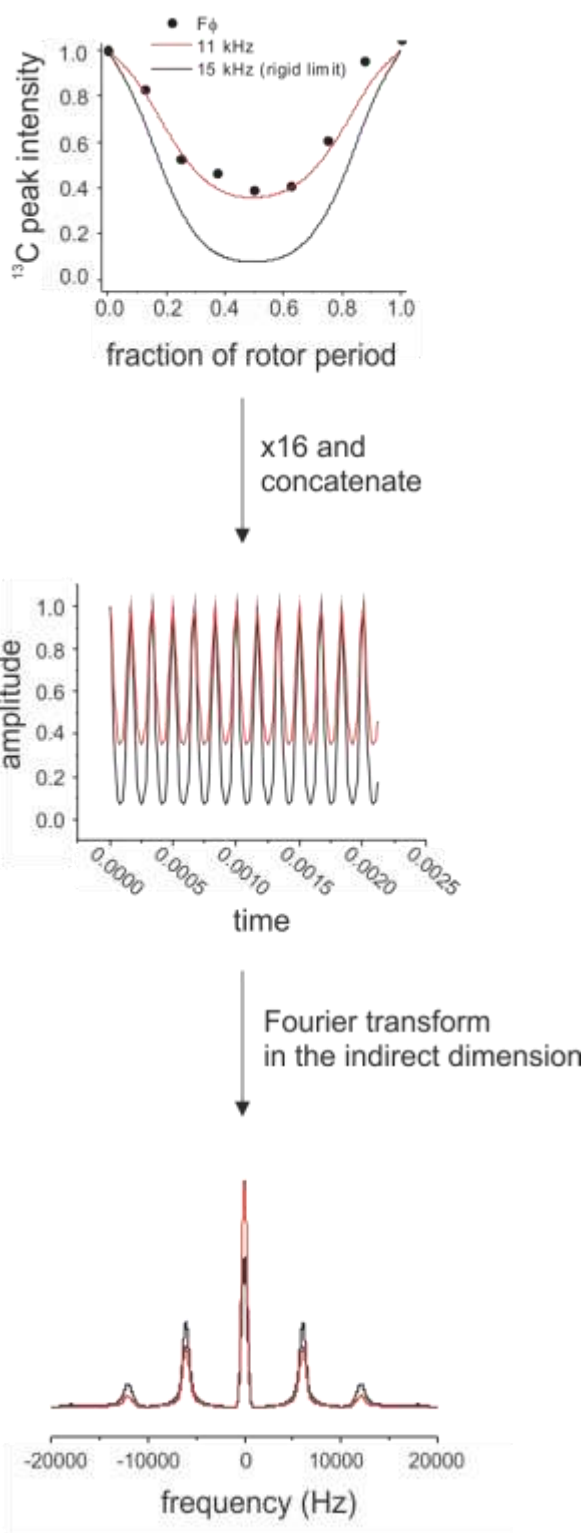


### 2.2.5.2. Rotational resonance experiments.

A two-dimensional version of the rotational resonance (RR) exchange experiment (Madine et al., 2009a) was used to selectively observe the dipolar interactions between the aromatic ring ( $F\phi$ ) and valine methyl ( $V\gamma V\gamma'$ )  $^{13}\text{C}$  spins. Spectra were recorded with 64 hypercomplex points in the indirect dimension ( $t_1$ ) and the spinning frequency ( $\omega_R$ ) was adjusted to the difference in the resonance frequencies ( $\Delta\Omega_{CC}$ ) for the  $F\phi$  and  $V\delta$  spins (i.e., the  $n = 1$  RR condition). Protons were decoupled at 100 kHz during a mixing time of 30 ms.

### 2.2.5.3. Site-specific $^{13}\text{C}$ - $^1\text{H}$ dipolar couplings

Site-specific  $^{13}\text{C}$ - $^1\text{H}$  dipolar couplings, which are sensitive to local dynamics on the  $\mu\text{s}$ - $\text{ms}$  time scale, were measured using the constant time dipolar and chemical shift (CT-DIPSHIFT) correlation experiment (Hong et al., 1997). In a series of 8 experiments, peak intensities were measured for 8  $t_1$  intervals, where  $t$  was incremented from zero to one rotor period ( $2\pi/\omega_R$ ) at a spinning frequency of 6 kHz. The 8 spectra ( $S_n$ ), each defined by 1024 points, were then replicated 16 times and concatenated ( $S_1, S_2 \dots S_8, S_1, S_2 \dots S_8$ , etc), to give a  $1024 \times 128$  matrix, which was Fourier transformed in the indirect dimension to produce the pseudo-two dimensional spectra, with the  $^{13}\text{C}$  chemical shift information in the direct dimension and C-H dipolar coupling information in the indirect dimension.



**Figure 2.2 Schematic to illustrate how the dipolar spectrum of a spin system can be Fourier transformed to give a pure dipolar spectrum. Peak intensities of at 8 different rotor period fractions were recorded and fitted to curves of simulated coupling strengths. These data were then replicated 16 times and concatenated. This is then FT in the indirect dimension to give C-H dipolar coupling information.**

### **2.2.6. Solution state NMR studies of Med<sub>1-50</sub>**

1D proton NMR spectra were recorded at 298 K on Bruker Avance 600 MHz spectrometer equipped with <sup>1</sup>H, <sup>13</sup>C, <sup>15</sup>N triple resonance gradient cryoprobe. Spectra (64 scans) were recorded on freshly prepared synthetic medin samples (see section 2.2.3) and after 6 days incubation at 25 °C with no agitation. The large resonance due to the water protons was suppressed by the WATERGATE pulse sequence (Sklenar et al., 1993). The spectra were processed using TopSpin (Bruker) and plotted in Origin Pro 8.5.

## **2.3. Results**

### **2.3.1. Aggregation prediction servers**

Aggregation prediction servers identify aggregation prone regions based on the physicochemical properties of the constituent amino acids, and pinpoint areas that can be tested experimentally. Prediction servers were used here to direct the design of the intermediate medin sequence and in turn investigate the role of different aggregation ‘hot spots’ on the structure and aggregation properties of medin. The sequence of medin was analysed using three different predictive methods, Zyggregator (Tartaglia and Vendruscolo, 2008), AGGRESCAN (Conchillo-Solé et al., 2007) and PASTA (Trovato et al., 2007) to obtain more reliable prediction results. All three servers identified the region encompassing residues NH<sub>2</sub>- F<sup>46</sup>GSV-COOH, as a region that promotes aggregation (Figure 2.3). The known amyloidogenicity of the peptide Med<sub>42-49</sub> (Larsson et al., 2007, Madine et al., 2009a, Reches and Gazit, 2004) indicates the reliability of the predictions in this case.

### Zyggregator

RLDKQGNF**NAWVAG**SYGND**QWLQ**VDLGSSKEVTGIITQGARNFG**SVQFVA**  
10 20 30 40 50

### AGGRESCAN

RLDKQGNF**NAWVA**GSYGNDQWLQVDLGSSKE**VTGIIT**QGARNFG**GSVQFVA**  
10 20 30 40 50

### PASTA

RLDKQGNFNAWVAGSYGNDQWLQVDLGSSKE**VTGIIT**QGARNFG**SVQFVA**  
10 20 30 40 50

**Figure 2.3 The aggregation prone regions of Med<sub>1-50</sub> (shown in red) predicted by three independent servers: Zyggregator, Aggrescan, and PASTA.**

Zyggregator and AGGRESCAN also predicted that residues NH<sub>2</sub>-N<sup>9</sup>AWVA-COOH are prone to aggregation (Figure 2.3). A third region with the sequence NH<sub>2</sub>-V<sup>32</sup>TGII-COOH was predicted by AGGRESCAN and PASTA (Figure 2.3).

From the results of the predictions, two peptides were produced synthetically, a peptide encompassing residues 30-50 (Med<sub>30-50</sub>) and full length medin (Med<sub>1-50</sub>) (Figure 2.4). Med<sub>30-50</sub> encompasses two of the aggregation prone regions and was investigated alongside selectively labelled full length medin (Med<sub>1-50</sub>). In order to compare results for both of these peptides to data previously collected for Med<sub>42-49</sub>, both peptides were synthetically produced, incorporating the same isotope labelling scheme.

**Med**<sup>42-49</sup>

RLDKQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQGARNFGSVQFVA\*\*\*\*\*

**Med**<sup>30-50</sup>

RLDKQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQGARNFGSVQFVA\*\*\*\*\*

**Med**<sup>1-50</sup>

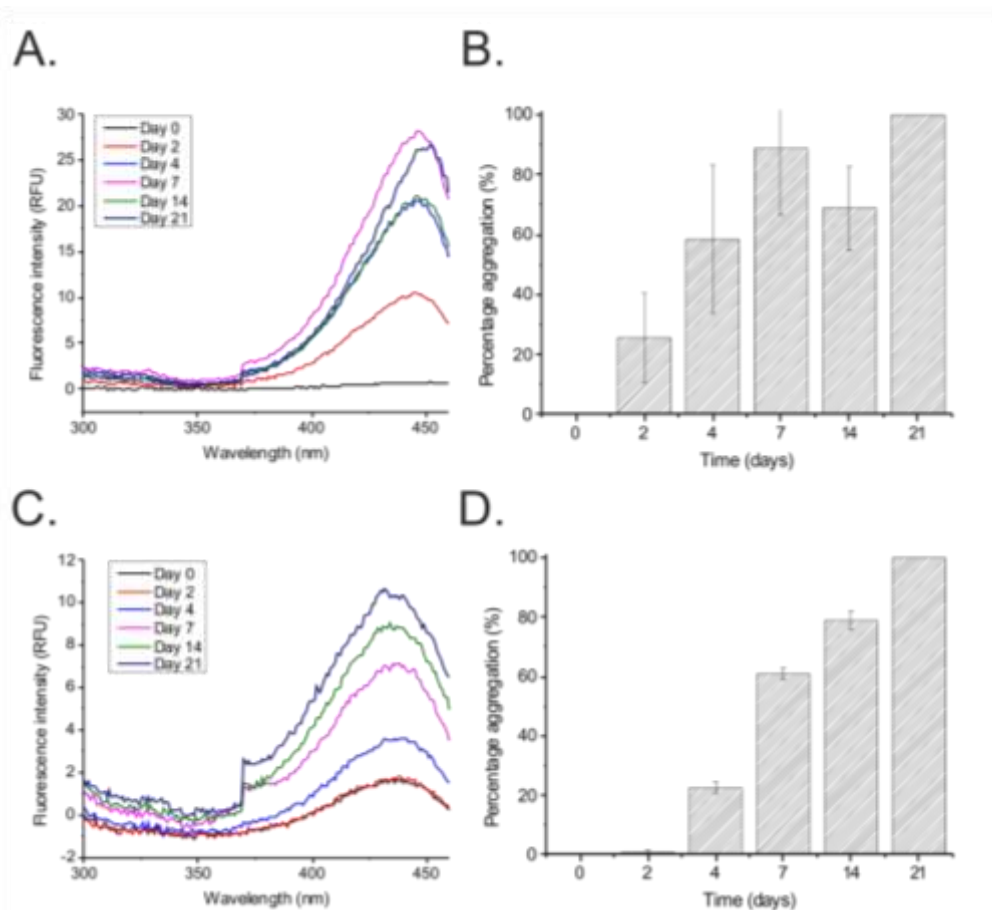
RLDKQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQGARNFGSVQFVA\*\*\*\*\*  
10 20 30 40 50

**Figure 2.4** Medin peptide sequences. Med<sub>42-49</sub> is shown in red, Med<sub>30-50</sub> in blue and full length wild type medin is shown in black. Asterisks indicate the presence of uniform <sup>13</sup>C and <sup>15</sup>N isotope labels. Residues underlined are predicted to form amyloid.

### 2.3.2. Kinetic and morphological characterisation of Med<sub>30-50</sub> and Med<sub>1-50</sub>

#### 2.3.2.1. Aggregation kinetics of Med<sub>30-50</sub> and Med<sub>1-50</sub> studied by Thioflavin T

The aggregation kinetics of both Med<sub>30-50</sub> and Med<sub>1-50</sub> were investigated using ThT fluorescence assay. Both Med<sub>30-50</sub> and Med<sub>1-50</sub> give rise to the characteristic rise in ThT fluorescence associated with amyloid formation (Figure 2.5).



**Figure 2.5** ThT fluorescence measurements of Med<sub>30-50</sub> (A and B) and Med<sub>1-50</sub> (C and D) incubated at a concentration of 200  $\mu$ M in 10 % DMSO and 90 % H<sub>2</sub>O. Samples were taken discrete time points over a three week period. Data were recorded on a Cary Eclipse Varian fluorescence spectrometer at 25 °C. Data represent mean data for 2 replicates. Panels A and C represent raw excitation spectra for one of the replicates. Panels B and D show the averaged data for the two replicates converted to percentage aggregation; fluorescence intensity at T21 represents 100 % and T0 represent 0%.

Aggregation of the peptides was carried out over a period of 21 days and samples were removed from the solution at discrete time points during the incubation. The length of the time intervals between samples make it difficult to draw conclusions about the exact length of the lag phase and the maximum fluorescence. Med<sub>30-50</sub> aggregation may exhibit a small lag phase before a period of growth and elongation (Panel A, Figure 2.5). There is also a dip at time 14 days, however this decrease is within the margin of error and therefore the relevance cannot be determined. In

contrast Med<sub>1-50</sub> exhibited a longer lag phase and did not reach equilibrium within 21 days (Panel B, Figure 2.5). Data for Med<sub>1-50</sub> in particular, was consistent with nucleation dependent growth with a lag phase followed by a period of growth and elongation (Cabriolu et al., 2010).

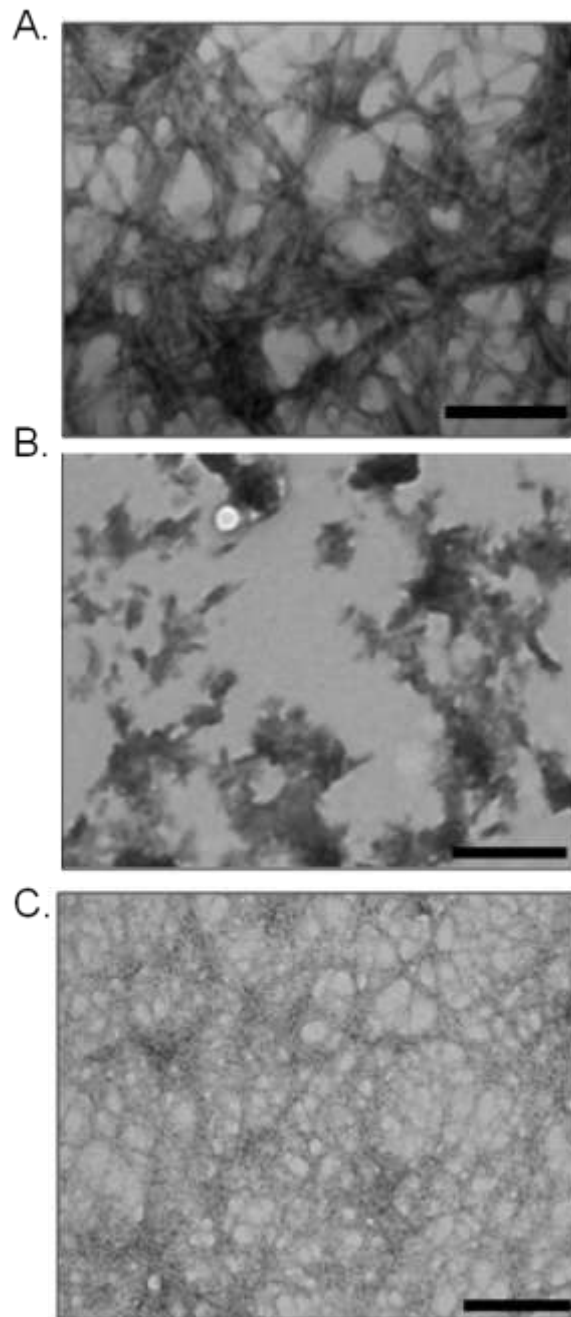
Previous *in situ* SSNMR analysis of Med<sub>42-49</sub> aggregation indicated that large or insoluble species occurred after monitoring for 6 hours, with aggregation reaching completion by 24 hours (Madine et al., 2009a). Med<sub>42-49</sub> aggregation is not accompanied by an increase in ThT fluorescence and so could not be studied by this method. ThT insensitivity for short amyloidogenic peptides is often observed (Groenning, 2010, Goldsbury et al., 2000) although the reasons are not well understood. It is possible that the length of Med<sub>42-49</sub> strand and/or the properties of the exposed residues do not favour ThT binding or fluorescence enhancement.

Together these data indicate that although truncation of the medin sequence alters the aggregation kinetics, both peptides demonstrate the rise in thioflavin T fluorescence associated with amyloid formation. In order to directly compare the different kinetic parameters such as lag time and growth and elongation rate, the aggregation of Med<sub>30-50</sub> and Med<sub>1-50</sub> would need to be followed in real time.

### **2.3.3. Morphology of Med<sub>42-49</sub>, Med<sub>30-50</sub> and Med<sub>1-50</sub> aggregates studied by TEM.**

The morphology of Med<sub>30-50</sub> and Med<sub>1-50</sub> was studied by TEM after 21 days and compared to previous images of Med<sub>42-49</sub>. Each of the three peptides gave a distinct morphology. Med<sub>42-49</sub> forms large numbers of rod-like fibres or needles (Panel A, Figure 2.6), Med<sub>1-50</sub> shows a dense network of slender, amyloid-like fibrils (Panel C, Figure 2.6) and Med<sub>30-50</sub> aggregates appear to be granular in character (Panel B, Figure 2.6). Hence, the resulting morphology of the medin-derived peptides appears

to be highly sensitive to their length and/or sequence. These results also indicate that not all predicted regions are required for aggregate formation to occur.



**Figure 2.6** Electron micrographs of Med<sub>42-49</sub> (A) reproduced from (Madine et al., 2009a) depicting rod-like fibrils or needles. Electron micrographs of Med<sub>30-50</sub> (B) indicate the presence of amorphous aggregation whereas and Med<sub>1-50</sub> (C) demonstrates the presence of slender amyloid-like fibrils. Scale bars represent 1000 nm.



### 2.3.4. Peptide NMR

Solid state NMR experiments on Med<sub>42-49</sub> aggregates provided information about their structure and packing arrangement (Madine et al., 2009a). Here the same SSNMR experiments are applied to Med<sub>30-50</sub> and Med<sub>1-50</sub> to investigate the structure and molecular architecture of these longer peptides and compared to Med<sub>42-49</sub>.

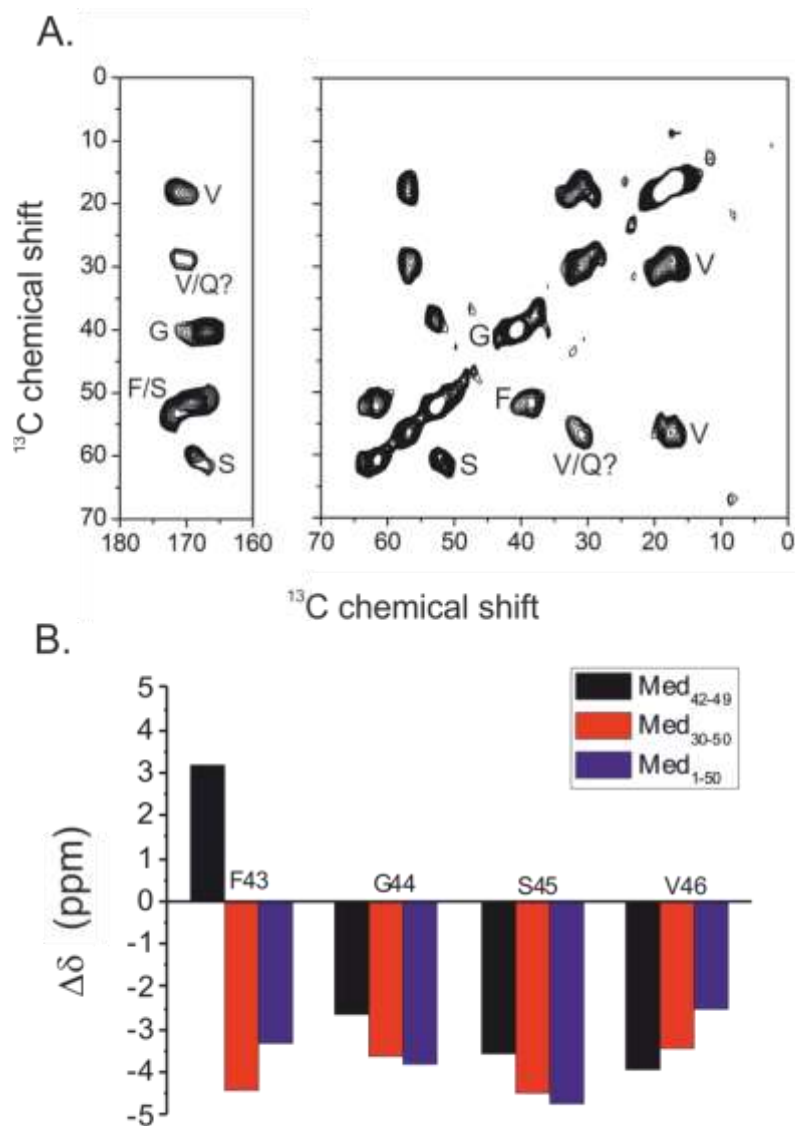
#### 2.3.4.1. DARR spectra of Med<sub>30-50</sub> and Med<sub>1-50</sub> aggregates

As shown in Figure 2.4, both Med<sub>30-50</sub> and Med<sub>1-50</sub> are <sup>13</sup>C/<sup>15</sup>N labelled at 5 positions: NH<sub>2</sub>-FGSVQ-COOH. Residue selective labelling is achieved by solid-phase synthesis and simplifies otherwise complex spectra and increases sensitivity and resolution. The DARR experiment is a broadband recoupling experiment which allows the detection of inter-residue contacts that can then be assigned. DARR spectra of Med<sub>30-50</sub> and Med<sub>1-50</sub> were collected and used to assign the selectively labelled residues in preparation for further structural analyses. It was however not possible to assign Q<sup>47</sup> in either Med<sub>30-50</sub> or Med<sub>1-50</sub>. The chemical shift values for Q C $\alpha$  reportedly range from 54.1 - 57 ppm for a  $\beta$ -sheet conformation and between 54.55 and 58.23 for a random coil arrangement (Wang and Jardetzky, 2002). This region of the DARR spectrum also corresponds to typical C $\alpha$ -C $\beta$  valine shifts. It is therefore not possible to distinguish between C $\alpha$ , C $\beta$  of V46 and Q47 in this data set. However, the C $\gamma$  resonance of V46 can clearly be identified from its distinct chemical shifts at around 19 ppm, hence allowing the resonance assignment of this residue (Figure 2.7). Furthermore, the chemical shift values obtained can provide information about the secondary structure of the C-terminus of these peptides (Wishart and Sykes, 1994). It has been shown previously that C $\alpha$  carbons

experience a downfield shift when located within an  $\alpha$ -helix and an upfield shift when found in a  $\beta$ -stand. Change in chemical shift values for the labelled residues of each of the peptides were calculated using the following equation:

$$\Delta\delta = (\delta C\alpha - \delta C\beta)_{obs} - (\delta C\alpha - \delta C\beta)_{RC} \quad \text{Equation 2.1}$$

where  $\delta C\alpha$  and  $\delta C\beta$  are the chemical shift values for  $C\alpha$  and  $C\beta$ , and subscripts RC and obs denote the mean random coil (Wang and Jardetzky, 2002) and observed values, respectively (Wishart and Sykes, 1994) and  $\Delta\delta$  is the chemical shift index.



**Figure 2.7** Regions of a representative  $^{13}\text{C}$  DARR spectrum for Med<sub>30-50</sub> with dashed lines showing peak assignments to the F43, G44, S45 and V46 spin systems (A).  $\text{C}\alpha$  chemical shifts for Med<sub>42-49</sub>, Med<sub>30-50</sub> and Med<sub>1-50</sub> are summarised in Table 2.1. Chemical shift index ( $\Delta\delta$ ) for residues F43, G44, S45 and V46 of Med<sub>42-49</sub>, Med<sub>30-50</sub> and Med<sub>1-50</sub> (Wishart and Sykes, 1994). Positive values of  $\Delta\delta$  are consistent with a helical structure, and negative values are consistent with a  $\beta$ -strand structure.

A representative  $^{13}\text{C}$  DARR spectrum of Med<sub>30-50</sub> is shown in panel A, Figure 2.7, and the measured chemical shift values for the  $\text{NH}_2$ -FGSV-COOH sequence of all three peptide aggregates are given in Table 2.1. Although the accuracy of the

measurements is compromised to some extent by the NMR line widths (typically 1.0-2.0 ppm), the chemical shift values nevertheless show some interesting site-specific and peptide-specific differences. The values for C $\alpha$  of F43 range from 55.8 to 52.6, a difference of >2.5 ppm (Table 2.1). Values for C $\alpha$  and C $\beta$  of G44 and V46 also differ by 1.0-1.5 ppm (Table 2.1). This distribution of values amongst the three peptides may reflect differences in local conformation within the NH<sub>2</sub>-FGSV-COOH sequence.

Despite these differences, secondary structure chemical shift, relating to random coil values ( $\Delta\delta$ ), are largely negative (Panel B, Figure 2.7) and thus broadly consistent with a  $\beta$ -sheet structure across the C-terminal region. Conversely, F43 of Med<sub>42-49</sub> gives a positive  $\Delta\delta$  value, which is consistent with a more  $\alpha$ -helical geometry (Wang and Jardetzky, 2002). In contrast to the longer peptides, F43 within Med<sub>42-49</sub> is close to the N-terminus, and its geometry may therefore be more weakly constrained by neighbouring residues.

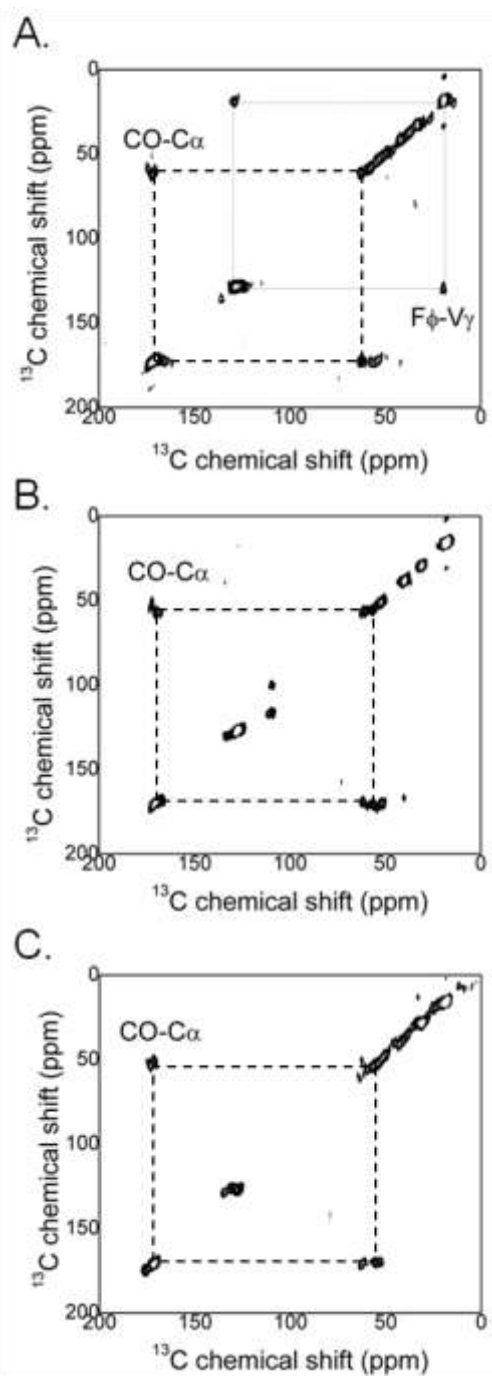
Interestingly, chemical shift values for the F43 ring (Table 2.1) are quite similar for Med<sub>30-50</sub> and Med<sub>1-50</sub> but somewhat different for Med<sub>42-49</sub> implying that the shorter peptide may have a different side chain packing arrangement to the longer peptides.

**Table 2.1**  $^{13}\text{C}$  chemical shift values for the amide carbon (C),  $\text{C}\alpha$ , side chain ( $\beta, \gamma, \gamma'$ ) and all aromatic ( $\phi$ ) positions of medin peptide aggregates. Values for each site corresponding (top to bottom) to Med<sub>42-49</sub>, Med<sub>30-50</sub> and Med<sub>1-50</sub>. Line width values shown in parentheses.  $\Delta\delta$  is defined as  $\Delta\delta = (\delta\text{C}\alpha - \delta\text{C}\beta)_{obs} - (\delta\text{C}\alpha - \delta\text{C}\beta)_{RC}$

Residue	Chemical shift (ppm)					
	C	$\text{C}\alpha$	$\text{C}\beta$	$\Delta\delta$	$\text{C}\gamma(\gamma')^a$	$\text{C}\phi^b$
<b>F43</b>	170.4	55.8 (1.5)	35.1 (1.8)	3.19	136.9 (2.1)	128.5
	170.3	52.6 (1.6)	39.5 (2.0)	-4.41	136.0 (2.2)	129.1
	170.6	53.1 (1.6)	38.9 (1.7)	-3.31	136.0 (2.2)	129.3
<b>G44</b>	169.5	42.7 (1.3)	-	-2.64	-	-
	167.9	41.7 (1.4)	-	-3.64	-	-
	168.2	41.5 (1.5)	-	-3.84	-	-
<b>S45</b>	171.0	53.6 (2.0)	62.7 (1.0)	-3.57	-	-
	169.7	52.9 (1.6)	62.9 (1.3)	-4.47	-	-
	170.2	53.1 (1.6)	62.1 (1.5)	-4.94	-	-
<b>V46</b>	173.4	58.8 (1.5)	33.6 (1.4)	-3.92	18.3 (1.2)	-
	172.5	57.7 (1.5)	32.0 (1.4)	-3.42	18.6 (1.1)	-
	172.9	58.7 (1.6)	32.1 (1.5)	-2.52	18.5 (1.2)	-

### 2.3.4.2. Investigating side chain packing arrangements examined by NMR rotational resonance experiments.

The side chain packing of F43 in the three peptide aggregates was examined further using selective dipolar recoupling NMR measurements. Two-dimensional exchange spectra of Med<sub>42-49</sub>, Med<sub>30-50</sub> and Med<sub>1-50</sub> were collected using MAS rates of 11,126, 11,116 and 11,146 Hz respectively. The rate corresponds to the frequency difference between the  $F\phi$  and  $V\gamma/V\gamma'$  (i.e.  $n = 1$  rotational resonance). At rotational resonance, pairs of nuclear spins are recoupled under MAS, provided the distance between them is  $< 6.5 \text{ \AA}$ . If this condition is met, cross-peaks corresponding to the resonances of the two nuclei may be observed in the 2D spectrum. Figure 2.8 shows cross-peaks correlating the  $^{13}\text{C}$ - $^{13}\text{C}$  dipolar coupling between  $F\phi$  and  $V\gamma/V\gamma'$  of Med<sub>42-49</sub> (Panel A, Figure 2.8). For  $F\phi$  and  $V\gamma/V\gamma'$  to be  $< 6.5 \text{ \AA}$  apart pairs of  $\beta$ -sheets must be packed together in a face-to-back arrangement (Nelson et al., 2005, Nelson and Eisenberg, 2006) (taking into account the additional SSNMR structural restraints from (Madine et al., 2009a)). Coincidentally, the MAS rate also coincides with the  $n = 1$  rotational resonance condition for the directly bonded amide C O and C $\alpha$  spins (dashed lines) giving strong cross-peaks which can be used as an internal reference. It is not possible to identify which CO-C $\alpha$  spins are recoupled due to overlap in that region of the spectra. Equivalent spectra of Med<sub>1-50</sub> and Med<sub>30-50</sub> show cross-peaks correlating the strongly coupled amide CO and C $\alpha$  spins, but cross-peaks between  $F\phi$  and  $V\gamma/V\gamma'$  are absent (panels B and C, Figure 2.8). Hence for the longer peptides, the  $F\phi$ - $V\gamma/V\gamma'$  distance is greater than  $6.5 \text{ \AA}$ , and the molecular packing is thus different in some respect from that of Med<sub>42-49</sub>.

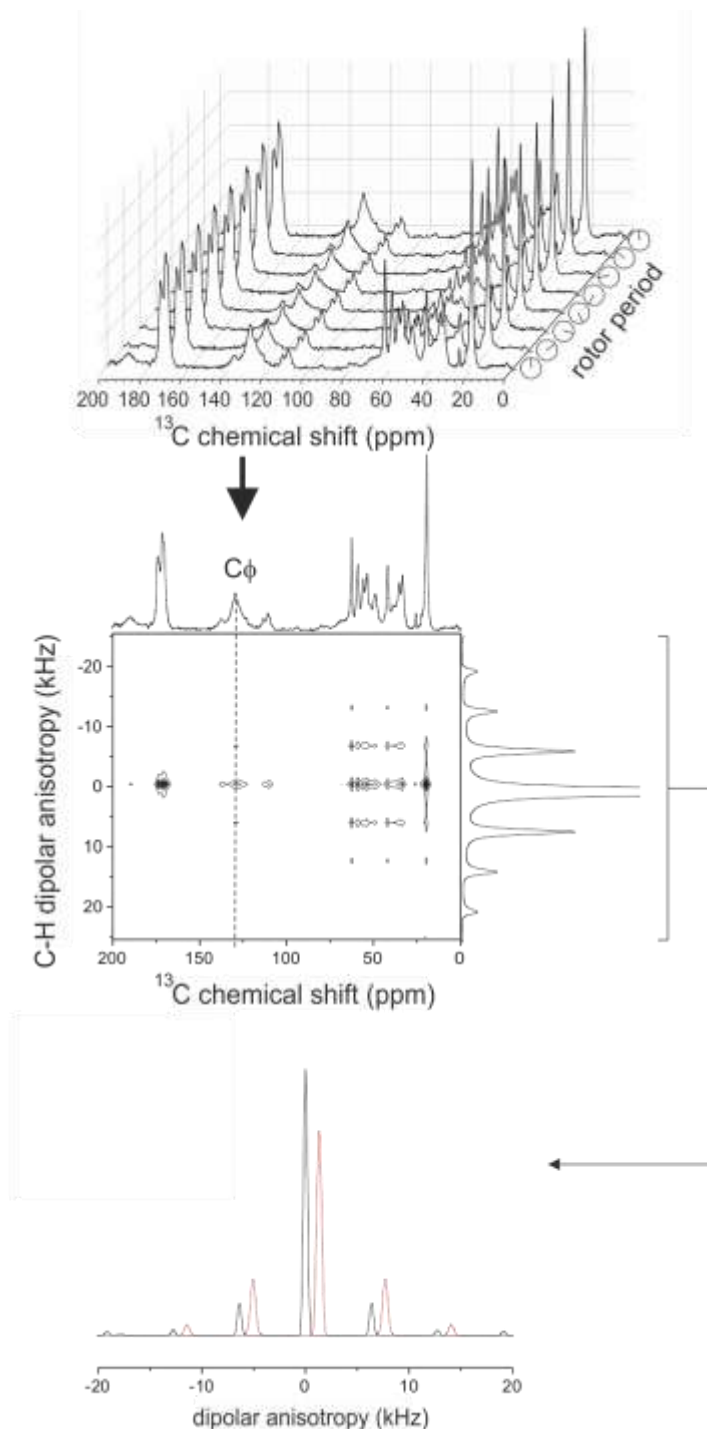


**Figure 2.8** Rotational resonance spectra for Med <sup>42-49</sup> (A) Med <sup>30-50</sup> (B) and Med <sup>1-50</sup> (C). MAS rates correspond to  $n = 1$  rotational resonance with respect to the frequencies of  $F\phi$  and  $V\gamma/\gamma'$  in each case (11126 Hz, 11116 Hz and 11146 Hz respectively). The  $F\phi-V\gamma/\gamma'$  cross peak is highlighted by the dotted line in (A) but is absent from (B) and (C). The dashed lines highlight cross-peaks arising from strongly coupled CO-C $\alpha$  spins. In each spectrum the intensities were normalised between 0 and 1 and peaks are represented by the same 24 contour levels.

#### **2.3.4.3. Aromatic ring flexibility- CT-DIPSHIFT solid-state NMR experiments**

Local dynamics within amyloid fibrils have until recently been neglected in structural studies yet potentially carry useful information about the fibril architecture (Helmus et al., 2011, Scheidt et al., 2011, Sackewitz et al., 2008). Here local dynamics of NH<sub>2</sub>-FGSV-COOH region were investigated to determine whether measurements within this area could provide clues about differences in the side chain packing in the three peptide aggregates. The <sup>13</sup>C-detected DIPSHIFT SSNMR experiment was used to measure <sup>13</sup>C-<sup>1</sup>H dipolar couplings within the peptide backbone and side chains.





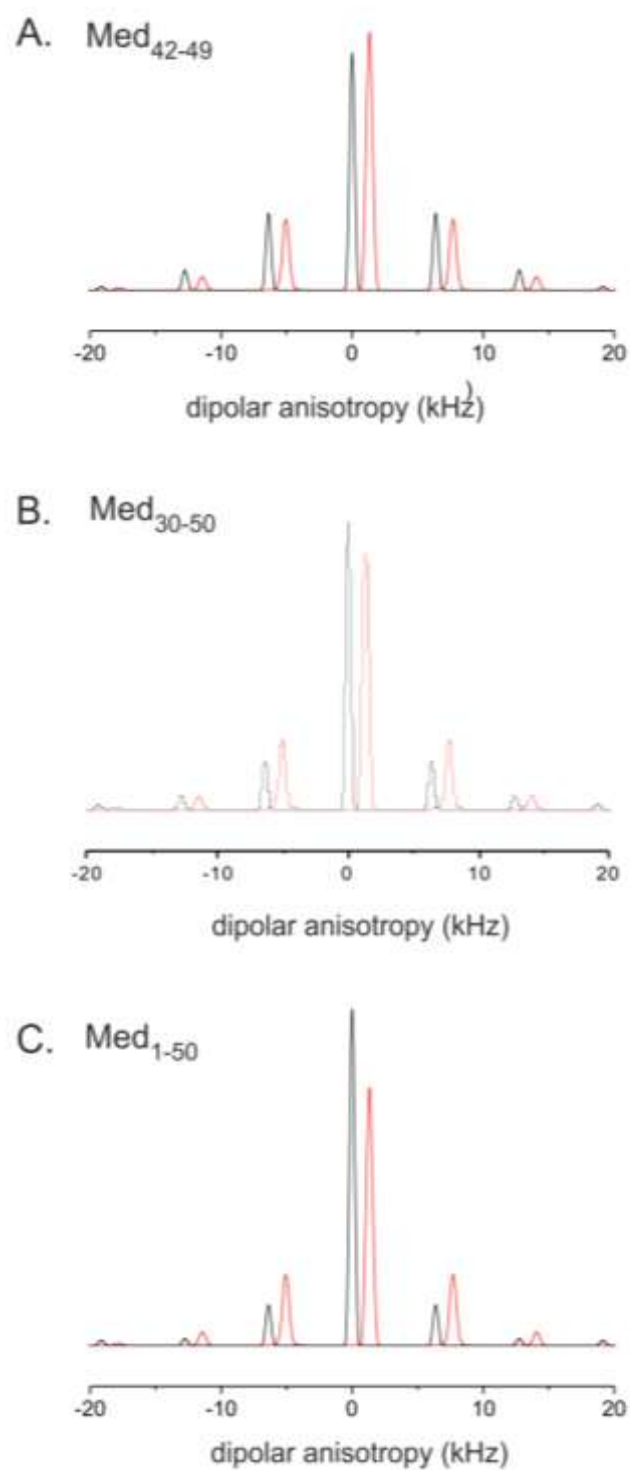
**Figure 2.9 (Top)** Stack plot showing the 2D CT-DIPSHIFT spectra for Med<sub>42-49</sub>. **(Middle)** A corresponding pseudo-2D spectrum. A slice through the spectrum representing the C—H dipolar anisotropy spectrum at the frequency of  $\text{C}\phi$  (dotted vertical line) for F43 shown on the right vertical axis. This dipolar anisotropy spectrum consists of a centre band flanked by a series of side bands separated by the sample spinning frequency. **(Bottom)** C-H dipolar anisotropy spectrum of F43  $\text{C}\phi$ , experimental data is shown in black and simulated rigid-limit dipolar spectra are shown in red (the data are offset for clarity).

Figure 2.9 is an example of a two dimensional CT-DIPSHIFT spectrum for Med<sub>42-49</sub> and illustrates how the spectra shown in Figure 2.10 are taken from a slice at the frequency indicated by the dashed line.

Spinning side band patterns extracted from the indirect dimension of the two-dimensional spectrum carry site-specific information about the <sup>13</sup>C-<sup>1</sup>H dipolar anisotropy, which is sensitive to motions on the μs–ms timescale (Figure 2.9).

Different relative side band intensities correspond to different motional averaging. The larger the central band intensities with respect to the side bands, the greater the motional averaging. For each of the three peptide aggregates, side band patterns were extracted at the <sup>13</sup>C frequencies of the protonated carbon sites.

As an example, panels A-C Figure 2.10, show the side band patterns for C $\phi$  of F43 for Med<sub>42-49</sub>, Med<sub>30-50</sub> and Med<sub>1-50</sub> alongside simulated side band patterns calculated for dipolar couplings assuming complete rigidity of the C $\phi$  of F43.



**Figure 2.10 Dipolar anisotropy spectra for the F43 aromatic carbons ( $F\phi$ ) of Med<sub>42-49</sub>, Med<sub>30-50</sub> and Med<sub>1-50</sub> aggregates (B-D). Experimental spectra are shown in black, and simulated rigid-limit dipolar spectra are shown in red. The spectra are offset to improve clarity.**

The amplitudes of motion at each site can be expressed as an order parameter  $S$  (Huster et al., 2001), which is calculated from the observed dipolar anisotropy ( $\bar{\delta}$ ) divided by the maximum anisotropy (i.e. rigid limit) ( $\delta$ ) (Equation 2.2).

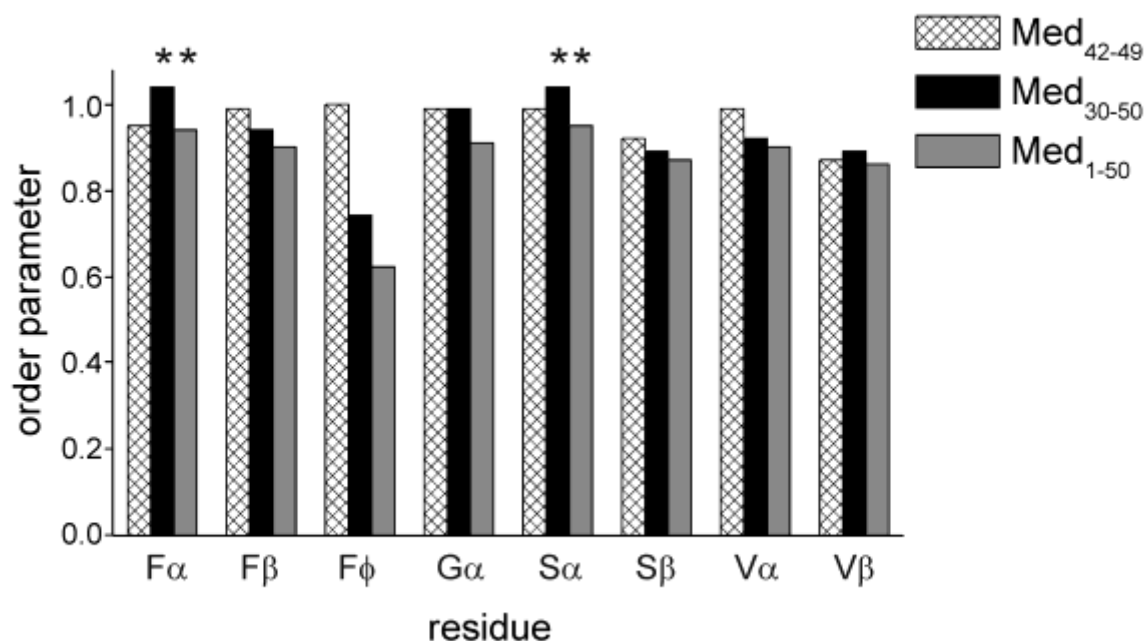
$$S = \frac{\bar{\delta}}{\delta} \quad 2.2$$

Complete isotropic motional averaging gives an order parameter of zero, and a completely rigid site gives an order parameter of 1. The dipolar coupling rigid limit is calculated using equation 2.3 and setting the inter-nuclear distance to correspond to the C-H bond limit:

$$b^{I_j S_k} = -\left(\frac{\mu_0}{4\pi}\right)\gamma_I\gamma_S\hbar(r^{I_j S_k})^{-3} \quad 2.3$$

The dipolar coupling constant ( $b^{I_j S_k}$ ) for a hetero-nuclear dipolar coupling between two spins ( $I_j$  and  $S_k$ ) is related to the inter-nuclear distance ( $r^{I_j S_k}$ ) by the gyromagnetic ratio of the two nuclei  $\gamma_I$  and  $\gamma_S$ , reduced Planck's constant ( $\hbar$ ) and the molecular moment  $\mu_0$ .

For Med<sub>42-49</sub>, all of the sites have high-order parameters (>0.9), consistent with no motional averaging on the  $\mu$ s-ms timescale of the peptide backbone and side groups, including F43 (Figure 2.11).



**Figure 2.11 Site-specific order parameters for the medin peptide aggregates. Lower values are consistent with increased amplitudes of motion on the  $\mu$ s–ms timescale. Asterisks indicate that the peaks for C $\alpha$  of F43 and S45 overlap in the DIPSHIFT spectra, and the values thus represent averages for the two sites.**

Residues within amyloid cores tend to be rigid in both the backbone and side chain sites, as observed for prion fibrils (Helmus et al., 2011) and also for stabilised A $\beta$  protofibrils (Scheidt et al., 2011), whereas N-terminal and C-terminal flanking residues are more dynamically disordered.

It appears that the N-terminal segment of Med<sub>42-49</sub> is highly ordered and may therefore be stabilised by a full complement of inter-strand hydrogen bonds and inter-sheet side chain interactions. For Med<sub>30-50</sub> and Med<sub>1-50</sub>, the backbone order parameters are also consistent with a rigid amyloid core (Figure 2.11). Order parameters for F43 however, are substantially lower (~0.6) indicating higher amplitudes of motion of the aromatic ring compared with Med<sub>42-49</sub> (Figure 2.11).

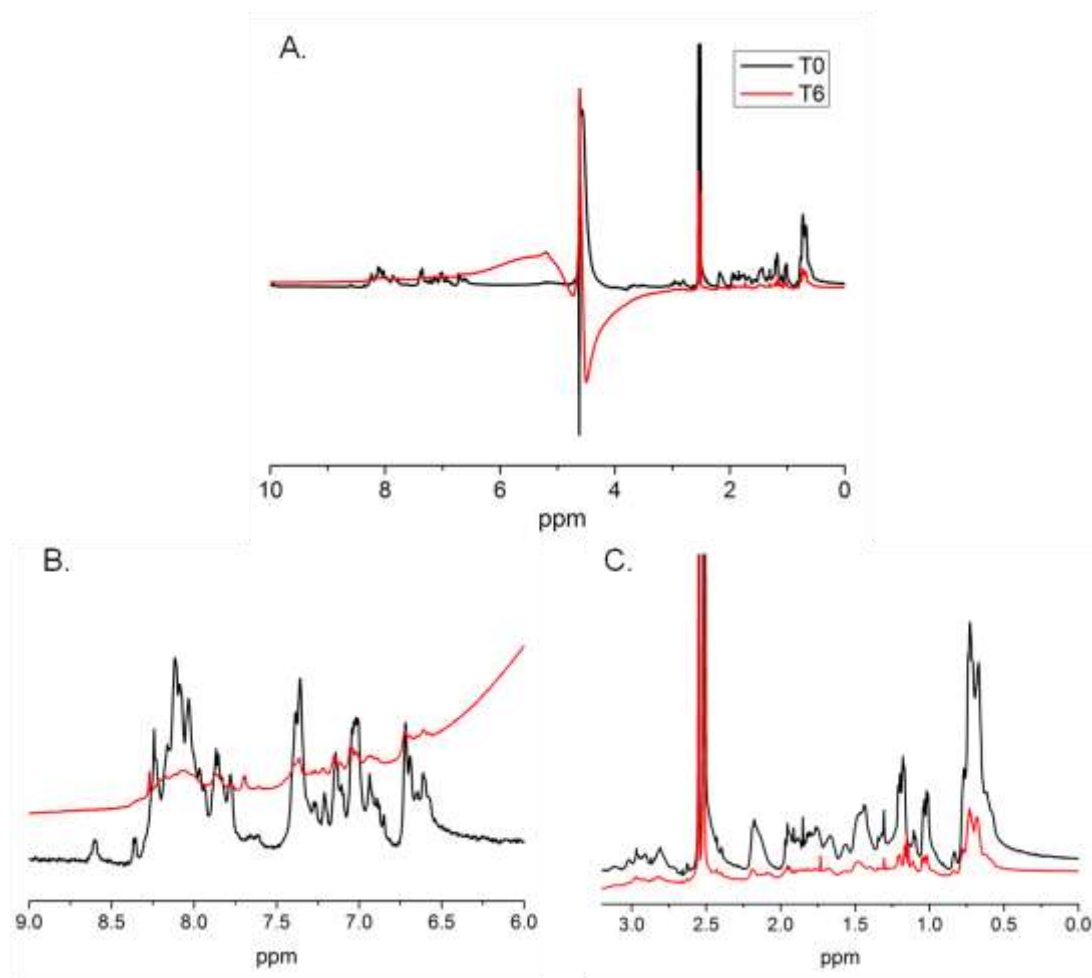
Thus, the dynamics again suggest that the chain packing of Med<sub>42-49</sub> is different from the two longer peptides.

#### **2.3.4.4. Solution state NMR studies of Med<sub>1-50</sub>**

A key area of current amyloid research is the study of intermediate species along the amyloid pathway such as oligomers. These oligomers, unlike the mature fibrils, are often soluble and may therefore be amenable to study by solution NMR. In order to investigate whether this would be a viable avenue of investigation, preliminary solution state experiments were carried out. As the solid state NMR data of the three peptides indicate that they each have distinct morphologies and aggregate structures, thus only freshly dissolved Med<sub>1-50</sub> was prepared for solution state NMR investigation.

Furthermore, the selective labelling scheme would reduce the complexity of the spectra and would be a good sample to test the possibility of characterisation of the early oligomerisation of medin using solution state NMR.

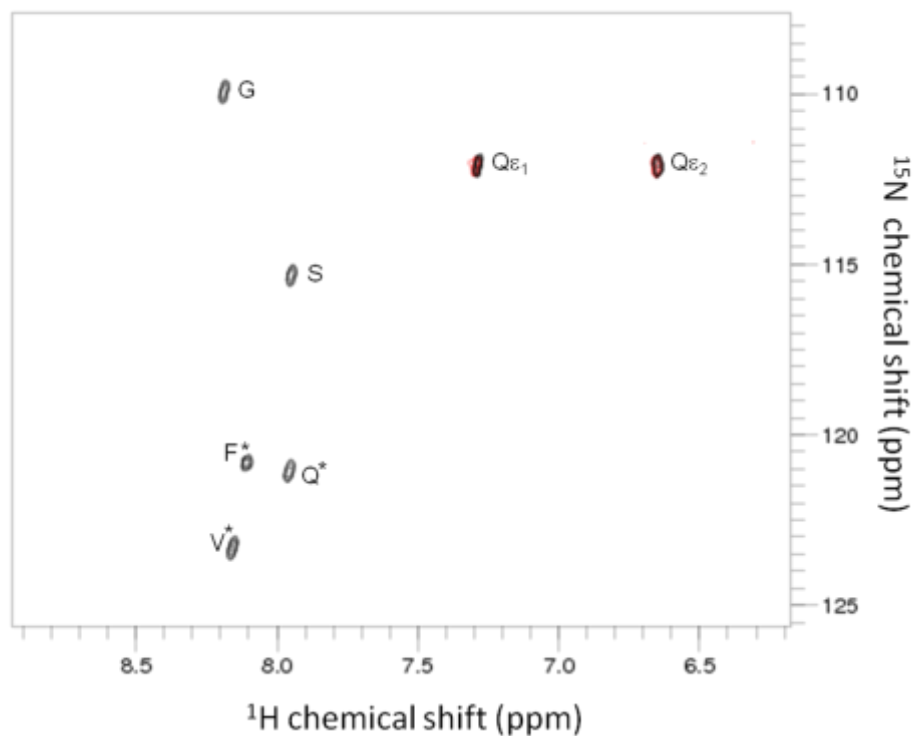
<sup>1</sup>H-NMR spectra of synthetic medin (20 μM) were collected on freshly prepared medin (see section 2.2.3 for details) and following 6 days incubation in the NMR tube. Interestingly, even freshly prepared medin showed evidence of oligomerisation giving broad peaks and uncharacteristic of a monomeric 50 amino acid peptide (black spectrum, Panel A, Figure 2.12). After 6 days incubation, precipitate was visible at the bottom of the NMR tube, the signals are considerably broader and there is a loss in intensity (red spectrum Figure 2.12). This suggests that medin has formed large aggregates that have precipitated and are therefore no longer visible by solution state NMR.



**Figure 2.12**  $^1\text{H}$ -NMR spectra of synthetic medin ( $200\ \mu\text{M}$ ) dissolved in  $90\%$   $^1\text{H}_2\text{O}$  +  $10\%$  deuterated DMSO. Data were collected on a Bruker Avance 600 spectrometer operating at a static magnetic field of 14 T. Black spectra represents freshly prepared medin sample and red spectra were recorded after 6 days incubation in the NMR tube at  $25\ ^\circ\text{C}$ . Full  $^1\text{H}$ -NMR spectra are shown in A, expanded amide and carbon regions are shown in B and C respectively.

Two-dimensional HSQC experiments were carried out on the same sample. These data show attenuation of the peaks over the time course, with the exception of the glutamine side chain (Figure 2.13). This suggests that the side chain group retain flexibility whereas the backbone amides are restrained. There is no chemical shift change suggesting that these peaks represent multiple steady-state oligomer populations.

The broad linewidths observed in the  $^1\text{H}$ -NMR spectra of the freshly dissolved peptide and the lack of shift changes in the 2D indicate that it would be very difficult to study medin aggregation in this way. As such, this approach was abandoned in favour of further biophysical characterisation and additional solid-state NMR experiments.



**Figure 2.13** Heteronuclear single quantum coherence (HSQC) spectra overlay of full length synthetic medin isotopically labelled in the region  $\text{NH}_2\text{-FGSVQ-COOH}$  in 90 %  $\text{H}_2\text{O}$ , 10 % DMSO. Spectrum of freshly prepared sample (black) overlaid with a spectrum recorded following 6 days incubation (red). Data were collected on a Bruker Avance 600 spectrometer operating at a static magnetic field of 14 T. Asterisks represent ambiguous assignment.



## 2.4. Discussion

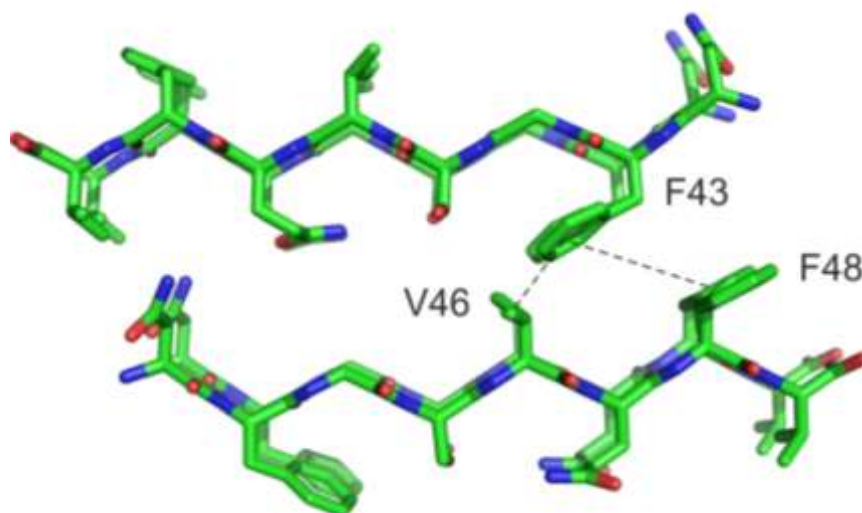
### 2.4.1. Investigating the role of predicted amyloidogenic regions on the kinetics, morphology and structure of medin

Kinetic analysis of the three peptides suggests that they have distinct aggregation kinetics. Previous data indicate the Med<sub>42-49</sub> aggregates rapidly within 48 hours (Madine et al., 2009a), and data shown here (Figure 2.5) imply that Med<sub>30-50</sub> and <sub>1-50</sub> have not reached completion after 21 days of incubation. There are also differences in the lag times between Med<sub>30-50</sub> and <sub>1-50</sub> further demonstrating the variation in the aggregation kinetics of the three peptides. TEM analysis revealed striking morphological differences between the three peptides (Figure 2.6). Together these data indicate that truncation of the medin sequence results in distinct aggregation kinetics and morphologies.

There is an increasing body of detailed structural information on the fibrillar aggregates of A $\beta$ , IAPP and several other amyloidogenic polypeptides associated with disease, as well as for numerous shorter peptides derived from these molecules. Solid-state NMR has played a key role in providing experimental constraints from which models of fibril architecture have been constructed (Paravastu et al., 2008, Luca et al., 2007, Van Melckebeke et al., 2010). Although medin is the most common form of localised amyloid, virtually nothing is known about its fibrillar assembly at the molecular level.

Investigation of two peptides Med<sub>30-50</sub> and Med<sub>1-50</sub> together with previous data on Med<sub>42-49</sub> by Madine *et al* (2009) provided a means to examine the effect of three predicted amyloidogenic regions on the aggregation properties of medin.

Several methods shown here suggest that Med<sub>42-49</sub> has a different fibril packing arrangement to Med<sub>30-50</sub> and Med<sub>1-50</sub>. The properties of F43 appear to be particularly diagnostic of differences in the molecular organisation of the Med<sub>42-49</sub>  $\beta$ -spine compared with Med<sub>30-50</sub> and Med<sub>1-50</sub>. In the structural model of the Med<sub>42-49</sub>  $\beta$ -spine described by Madine *et al* (2009) (Panel A, Figure 2.14), F43 and F48 of adjacent  $\beta$ -sheet layers are positioned close enough together to allow a stabilising  $\pi$ - $\pi$  interaction between them. The fibril architectures of Med<sub>30-50</sub> and Med<sub>1-50</sub> may preclude such a  $\pi$ - $\pi$  interaction because cross-peaks between F43-V46 are not visible in the RR experiments suggesting that the inter-nuclear distance between F43-V46 is longer in Med<sub>30-50</sub> and Med<sub>1-50</sub> than in Med<sub>42-49</sub> (Panel B, Figure 2.14). This might also explain the higher mobility of F43 ring seen for the two longer peptides. It is possible that the anomalous chemical shifts seen for F43 of Med<sub>42-49</sub> (Panel B, Figure 2.7) reflect local distortions in the backbone geometry that are needed to accommodate the  $\pi$ - $\pi$  interaction.



**Figure 2.14** A model of the Med<sub>42-49</sub> cross- $\beta$ -spine comprising two  $\beta$ -sheet layers in a face-to-back arrangement (Nelson *et al.*, 2005). The model satisfies the close through-space distance between F43 and V46 and allows a  $\pi$ - $\pi$  interaction between F43 and F48 of adjacent  $\beta$ -sheet layers as seen in Med<sub>42-49</sub>.

Until further structural constraints are obtained, it is only possible to speculate on the origins of the different F43 ring position and dynamics for Med<sub>30-50</sub> and Med<sub>1-50</sub> compared with Med<sub>42-49</sub>. Structural models of longer amyloidogenic polypeptides often indicate the presence of one or more turn regions within the fibril structure (Paravastu et al., 2008, Petkova et al., 2002, Luca et al., 2007) whereas fibrils of shorter peptides, including Med<sub>42-49</sub>, tend to be assembled from linear  $\beta$ -strands (Madine et al., 2008c, Madine et al., 2009a).

The different aggregate morphologies for Med<sub>1-50</sub> and Med<sub>30-50</sub> observed using TEM suggest that truncation of the 19 residues at the N-terminus significantly alters the structure of the fibrils. The similar SSNMR results for Med<sub>30-50</sub> and Med<sub>1-50</sub> may therefore not necessarily translate into similar fibril architectures for the two peptides at the molecular level, but further SSNMR measurements may confirm this.

Previous studies on amyloid fragments have shown that they rarely have the same fibril structure as their full length counterparts (Paravastu et al., 2006, Bu et al., 2007, Paravastu et al., 2008). This research emphasises the need to investigate the full length peptide alongside fragments, to correctly identify key structural and amyloidogenic features for future investigation.

Preliminary solution state NMR investigations of the early stages aggregation indicate that even freshly dissolved Med<sub>1-50</sub> shows signs of aggregation. Although the study of early amyloid species is important, given the known toxicity of oligomeric species, the structural characterisation of these species by solution state NMR was not pursued within this thesis.

Taken together, these results indicate that at least two of the three predicted amyloidogenic regions are required for the formation and elongation of medin fibres

observed in the disease state. These results provide the motivation for further, more detailed structural studies of this poorly characterised amyloidogenic polypeptide. This data also demonstrates that Med<sub>42-49</sub> is not a good structural model for full length medin. Consequently, full length medin will be studied for the remainder of this thesis.

## 3. Recombinant expression and purification of medin

### 3.1. Introduction

As the previous chapter has demonstrated, it is important to study full length medin. To date there are no detailed published protocols for the production of recombinant medin. The primary objective of this work was to establish a method for the production of soluble medin, in sufficient quantities for aggregation studies and structural analysis by solid state NMR. NMR experiments, in particular solid state NMR experiments, require large quantities of isotopically labelled protein. This chapter will discuss the methods used to produce, purify and prepare medin for aggregation and structural studies.

Isotope labelling is an important part of sample preparation for NMR spectroscopy. The use of stable isotope labels ( $^{13}\text{C}$  and  $^{15}\text{N}$ ) improves resolution and sensitivity, and can reduce the complexity of spectra allowing for better interpretation of the data (Lian and Middleton, 2001). The majority of macromolecular NMR studies employ  $^{13}\text{C}$ ,  $^{15}\text{N}$  and  $^2\text{H}$  either in isolation or combination. There are multiple methods and strategies used to label proteins of interest. These include: selective labelling of single residues using cell free or synthetic peptide production (Torizawa et al., 2004), selective or inverse labelling where metabolic precursors or unlabelled amino acids are incorporated into the media (Hiroaki et al., 2011) or uniform labelling (McIntosh and Dahlquist, 1990) in which all residues are labelled.

For the purposes of this thesis the most cost effective and straightforward way to produce uniformly labelled proteins is to use a bacterial expression host, most commonly *Escherichia coli* (*E.coli*). These cells are ideal for this purpose as they

grow rapidly, have a high plasmid copy number and are easy to maintain. Furthermore, there are many strains and vectors available, which provide a high level of control over the expression process. Generally *E.coli* will grow in a variety of growth media and rarely require the addition of expensive growth supplements. It is this property that allows the expression of isotopically labelled protein in milligram quantities suitable for NMR and unlabelled protein for other studies.

This chapter details the techniques and work-flow involved in producing both unlabelled and isotopically labelled wild-type medin. The protein prepared using the methods described here was later used to characterise the aggregation properties of medin in subsequent chapters.

## **3.2. Materials and methods**

### **3.2.1. Recombinant expression of medin in *E.coli***

#### **3.2.1.1. Construct design**

The medin gene was purchased from GenScript (U.S.A.) and inserted into a pMALC2X (Invitrogen) vector. The pMALC2X vector enables the expression and purification of the protein of interest fused to maltose binding protein (MBP). The cloned gene is inserted downstream of the *MalE* gene which gives rise to the expression of an MBP-fusion protein. It also encodes for a factor Xa cleavage site that allows for the separation of the target protein from MBP. In addition the pMALC2X vector contains ampicillin resistance gene for selection. This vector was chosen for the high solubility of MBP which would increase the yield of soluble protein and minimise the incorporation of medin into inclusion bodies (Waugh, 2005) A tobacco etch virus (TEV) protease cleavage site was also cloned into the vector as TEV protease was available in-house thus making it more cost effective than purchasing factor Xa (Figure 3.1).

A.

```
5' G↓AA TTC [GAA AAC CTG TAC TTC CAG↓GGT] CGT CTG GAC
AAA CAG GGT AAC TTC AAC GCT TGG GTT GCT GGT TCC TAC
GGT AAC GAC CAG TGG CTG CAG GTT GAC CTG GGT TCC TCC
AAA GAA GTT ACC GGT ATC ATC ACC CAG GGT GCT CGT AAC
TTC GGT TCC GTT CAG TTC GTT GCT (TAG) GC A↓AGCT T 3'
```

B.



**Figure 3.1 (A) Deoxyribose nucleic acid (DNA) sequence of the inserted gene containing the TEV protease recognition sequence (red) followed by medin sequence (black) and the stop codon in brackets. (B) Schematic of the resultant fusion protein. Red arrows indicate the protease cleavage site, black arrows indicate restriction sites.**

### 3.2.1.2. Preparation of chemically-competent cells

Competence is the term given to cells that are capable of incorporating extracellular deoxyribose nucleic acid (DNA) from the environment into their own genome. Some cells are naturally competent, but most cells need to be treated to make them transiently permeable to exogenous DNA. There are two principal methods for this process, electroshock and chemical competency; this work employed the use of chemically competent cells.

The protocol used is a variant of the Hanahan method (Hanahan et al., 1991) and is based on the use of calcium chloride. Prior to starting the protocol all glass and plastic ware was free of detergent and chilled. A single colony was selected from a streaked agar plate containing the plasmid and used to inoculate 2 ml of super optimal broth (SOB) media (Appendix 1) and grown overnight at 25 °C. Sterile glycerol was added to 1 ml of this starter culture at a final concentration of 15 % (v/v) and transferred to a cryovial for long term storage at -80 °C. The remaining 1



ml of culture was used to inoculate 250 ml of SOB media and grown at 20 °C until an optical density (O D) of 0.3 at 600 nm ( $OD_{600} = 0.3$ ) was obtained. Cells were harvested by centrifugation at 3000 g for 10 minutes and gently resuspended in 80 ml of ice cold CMB80 buffer (Appendix 1). The resuspended cells were then incubated on ice for 20 minutes and centrifuged a second time at 4 °C prior to resuspension in 10 ml of ice cold CMB80 buffer. The OD at 600 nm of the suspension was tested by adding 20  $\mu$ L of the cell suspension to 200  $\mu$ L of super optimal media with catabolite repression (SOC) media. The OD was adjusted to between 1 and 1.5 by the addition of CMB80 buffer. Aliquots of 50  $\mu$ L were generated and stored at -80 °C indefinitely.

#### **3.2.1.3. Transformation of vector into chemically competent cell lines.**

Transformation is the process whereby an engineered vector is incorporated into the genome of competent cells. A standard method was followed for all vectors and cell lines (Sambrook, 1989). In brief 50  $\mu$ L aliquots of competent cells were thawed on ice prior to the addition of 2  $\mu$ L of target vector. Following incubation on ice for 30 minutes the cells were heat-shocked at 42 °C for 45 seconds and then returned to ice. 950  $\mu$ L of SOB media was then added and the cells incubated at 37 °C for 30 minutes prior to plating out on agar plates containing antibiotic. The agar plates were incubated at 37 °C overnight and then stored at 4 °C.

Two *E.coli* strains were used in the process of this work, BL21 (DE3) cells for expression and XL1 blue cells for cloning, both supplied by NEB.

#### **3.2.1.4. Expression tests**

In order to identify the optimal conditions for growth and expression of the target protein, trials of temperature and expression time were carried out. The vector was

transformed into *E.coli* BL21 (DE3) competent cells. A colony was then selected and grown in 50 ml of Luria broth (LB) (Appendix 1) containing antibiotic overnight at 37 °C. This culture was then spun down at 3000 x g and used to inoculate 2 L of LB. The culture was incubated at 37 °C until an OD<sub>600</sub> = 0.8 was reached. The culture was then divided equally into 3 flasks. Then they were equilibrated at three different temperatures 20 °C, 30 °C and 37 °C. Expression of the fusion protein was induced by the addition of isopropyl β-D-thiogalactoside (IPTG) to a final concentration of 1 mM. Cell growth was then allowed to continue and samples were taken from all three flasks after 3 and 5 hours incubation. The remaining cultures at 20 °C and 30 °C were allowed to continue growing overnight and harvested the following morning. All the samples were then lysed by sonication and then centrifuged at 19000 x g, the supernatant and pellet were collected for analysis by sodium-dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE).

The resuspended cells were subjected to a single freeze-thaw cycle prior to 2 French press cycles in the presence of DNase at a final concentration of 20 ng/ml and a cComplete™ EDTA-free protease inhibitor cocktail tablet (Roche applied sciences). The lysed cells were then centrifuged at 19000 x g to remove the cell debris and the supernatant filtered through a 0.22 μm Acrodisc™.

#### **3.2.1.5. Expression of medin in the pMALC2X vector**

The transformation, overnight culture, inoculation of the media and initial growth were carried out as described in section 3.2.1.4, using terrific broth (TB) (Appendix 1) instead of LB as it resulted in a higher cell density and thus a higher yield of expressed protein. As detailed above, the fusion protein expression was induced by



Cloning was carried out as described in Berrow *et. al.*(Berrow *et al.*, 2009). In brief the medin sequence was amplified using the polymerase chain reaction (PCR) with specific primers (Sigma Aldrich) designed for insertion into pOPINS, red areas indicate the vector specific sequence and black is template specific:

**Forward primer**

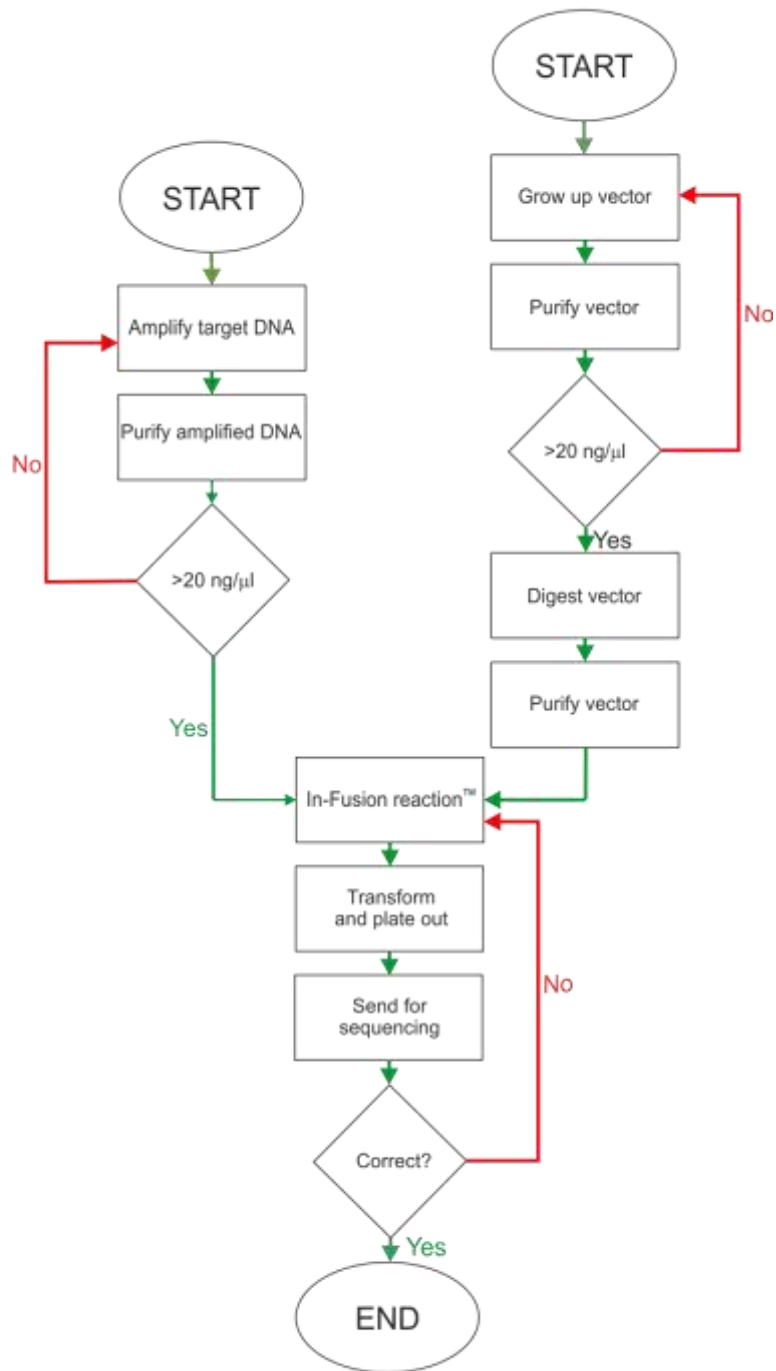
GC GAA CAG ATC GGT GGT CGT CTG GAC AAA CAG GGT AAC

**Reverse Primer**

ATG GTC TAG AAA GCT TTA AGC AAC GAA CTG AAC GGA AC

pOPINS was grown up in XL1 blue cells in LB at 37 °C overnight and then isolated from the cells using a midi-plasmid prep kit (Sigma Aldrich). The purified plasmid was then digested for 2 hrs at 37 °C with HindIII and Kpn1 (NEB). The amplified medin sequence and the digested vector were then added to the dried down In-Fusion™ mixture (Clontech) and incubated at 42 °C for 30 minutes. 2 µL of the mixture was then used to transform an aliquot of XL1 blue cells and plated out onto an agar plate containing 0.1 mM IPTG, 40 µg/mL XGal and 35 µg/mL of kanamycin. Several white colonies were selected, grown in 5 mL of TB and the plasmid isolated using a midi-plasmid prep kit (Sigma Aldrich). The purified plasmid was then sequenced (GATC, London U.K.).

A.



B.



**Figure 3.3 (A.) Schematic of the sub-cloning work-flow. (B.) Schematic of the fusion protein generated from the pOPINS-medin construct, arrow indicates cleavage site.**

### **3.2.1.7. Agarose gel electrophoresis**

A 1 % agarose gel was run to confirm that the PCR reaction and digestion were successful. To construct the gel, 1 g of agarose was dissolved in 100 ml of tris-acetate-EDTA (TAE) buffer (Appendix 1) by heating, poured into the gel cast and 0.5 µg/ml ethidium bromide added and mixed. Once set the gel was transferred to a gel tank and covered in TAE buffer. 5 µl of the PCR reaction or digested vector was mixed with 1 µl 5 x gel loading buffer (NEB) and loaded. The gel was run at 100 V for 60 minutes prior to visualisation under UV light.

### **3.2.1.8. Expression of pOPINS construct**

Initially expression tests of the new construct were carried out as described in section 3.2.1.4 and the optimal conditions were shown to be overnight incubation at 20 °C. The overall expression and lysis protocol was carried out as described above. The only exception being that kanamycin was used as opposed to ampicillin for selection, at a final concentration of 35 µg/ml.

For the preparation of <sup>13</sup>C and <sup>15</sup>N labelled medin, BL21 (DE3) cells were initially grown overnight in LB and then transferred to minimal media for expression with <sup>13</sup>C<sub>6</sub> glucose as the only carbon source and <sup>15</sup>NH<sub>4</sub>Cl as the only nitrogen source (See Appendix 1 for full details). All other variables remained the same as for those described for rich media.

## **3.2.2. Purification of MBP- medin fusion protein**

### **3.2.2.1. Amylose affinity purification**

The fusion protein can be isolated from crude extract by the use of affinity chromatography. The MBP-tagged protein will bind to amylose resin, whereas other

protein will flow through the column. The bound target protein can be eluted by the addition of maltose to the buffer which out competes MBP for amylose binding.

A 10 ml self-packed amylose column (GE Healthcare, NEB) was equilibrated with 5 column volumes (CV) of amylose binding buffer (Appendix 1), 5 CV of maltose buffer (Appendix 1) and a further 5 CV of amylose binding buffer. The filtered supernatant was then loaded onto the column at a flow rate of 2 ml/minute using a peristaltic pump. The column was washed with 10 CV of amylose binding buffer to remove any non-specific binding. The fusion protein was eluted with 3 CV of maltose buffer in 3 ml fractions. Fractions were tested for the presence and quantity of fusion protein using Bradford reagent (Sigma Aldrich). 10  $\mu$ L of the fraction was added to 990  $\mu$ L of Bradford reagent and the absorbance at 590 nm recorded and compared to a standard curve (Marion M, 1976). The fractions containing the fusion protein were then pooled prior to cleavage.

#### **3.2.2.2. Cleavage**

TEV protease was expressed and purified in-house as described in Appendix 2. TEV protease was added in a ratio of 1:20 with the fusion protein and incubated for 4 hours at 37 °C.

#### **3.2.2.3. Isolating medin**

Due to the large difference in molecular weight between medin and MBP gel filtration was used to isolate medin. Prior to gel filtration the sample was concentrated to a volume of 2 ml using an Amicon® Ultra-15 centrifugal filter with a 30 kDa cut off (Millipore) to optimise separation. A Superdex 75 26/60 gel filtration column (GE Healthcare) was equilibrated overnight (10 CV) in amylose binding buffer at a flow rate of 2 ml/min. The concentrated sample was then loaded

and run at 2 ml/min on an ÄKTA protein purifier (GE Healthcare). Elution was monitored by absorbance at 280 nm and 2 ml fractions were collected.

#### **3.2.2.4. Purification of SUMO-Medin fusion protein**

The SUMO-medin fusion protein can also be isolated from crude extract with the use of affinity chromatography. The presence of the His<sub>6</sub>-tag enables the use of immobilised metal affinity chromatography (IMAC), using resin functionalised with Ni<sup>2+</sup> ions. The histidines bind to the nickel ions whilst other proteins wash through the column. The bound protein is then eluted with imidazole, which competes with the Ni<sup>2+</sup> to chelate the histidines.

As far as possible the purification was carried out on ice. A 5 ml His-Trap FF™ column (GE Healthcare) was equilibrated with 5 CV of His-trap buffer A and 5 CV of his-trap buffer B followed by a further 5 CV of His-trap buffer A at 4 ml/min. Following equilibration, the filtered supernatant was loaded onto the column using an ÄKTA purification system at 1 ml/min (GE Healthcare).

A step-wise gradient was employed to remove non-specific contaminant proteins. This was followed by a linear gradient (Figure 3.9), to elute the target protein. Fractions of 3 ml were collected and those containing the fusion protein were pooled and desalted to remove the imidazole in preparation for the reverse His-trap purification step after cleavage. A Hi-Prep desalting column 26/10 (GE Healthcare) was equilibrated with 2 CV of His-trap buffer A prior to sample loading via the peristaltic pump. 2 ml fractions of the desalted protein were pooled in preparation for cleavage.



### **3.2.2.5. Cleavage and removal of the His<sub>6</sub>-SUMO tag.**

The His<sub>6</sub>-SUMO tag was removed by incubation with SUMO-protease (produced in-house), at a final concentration of 2.5 μM, at 25 °C for 3 hours. In order to remove the tag, the mixture was passed back down an equilibrated 5 ml His-Trap FF™ column (GE Healthcare) and the flow through containing the medin peptide was collected and snap frozen in liquid nitrogen for later use.

### **3.2.3. Characterisation**

#### **3.2.3.1. SDS PAGE gel analysis and staining methods**

The expression and purification of medin was analysed by SDS-PAGE using the Laemmli method (Laemmli, 1970). Analysis was carried out using 15% polyacrylamide gels in a Bio-Rad gel electrophoresis system. 5 μl protein samples were added to 5 μl of 2 x sample buffer, boiled and loaded onto the gel. All gels were run at 200 V for 60 minutes and then stained with Coomassie Brilliant Blue G-250 stain and de-stained with destain solution (Appendix 1).

Medin was difficult to visualise with the Coomassie Brilliant Blue G-250 with normal SDS-PAGE gels so other gel compositions and visualisation techniques were considered.

Tris-Tricine gels were constructed and run as described by Schagger and von Jagow omitting the spacer gel as it is usually only necessary for separating peptides with a molecular weight below 3 kDa (Schägger and von Jagow, 1987). See Appendix 1 for further details of buffers and gel construction.

Silver staining was carried out as published in Cell Biology: a laboratory handbook (Celis, 2006). In brief, the gel was fixed for a minimum of 2 hours and then washed in 20 % EtOH, The gel was rinsed in sensitising solution, and then washed twice

with dH<sub>2</sub>O. Chilled silver stain was added to the washed gel and incubated for 20 minutes. The stain was removed and the gel washed a further two times with dH<sub>2</sub>O prior to the addition of developing solution.

Glutaraldehyde cross-linking was carried out in an attempt to reduce protein leaching from the gel (Hopwood, 1972). Cross linking was performed with consecutive washes of methanol, acetic acid and water in different ratios (Appendix 1). The gel was then incubated with 1 % glutaraldehyde in H<sub>2</sub>O for 30 minutes, followed by further washes and then stained as normal with Coomassie Brilliant Blue G-250.

Trichloroacetic acid (TCA) can be used to precipitate proteins within a polyacrylamide gel (Link and LaBaer, 2011). The gel was first incubated with 10 % TCA for 30 minutes and then washed with 50 % methanol and 10 % acetic acid mixture in H<sub>2</sub>O prior to staining with Coomassie Brilliant Blue G-250.

#### **3.2.3.2. Western blot analysis**

Polyclonal anti-bodies against two epitopes of medin, corresponding to residues 1-10 and 19-31, were raised in rabbit and purified by GenicBio (Hong Kong). Anti-rabbit IgG peroxidase conjugate (A6154) was purchased from Sigma Aldrich.

SDS-PAGE 15% gels were run as described above and then transferred to polyvinylidene fluoride (PVDF) membrane in ice-cold transfer buffer (Appendix 1) for 1 hr at 100 V. The membrane was blocked overnight at room temperature with PBS with 5 mg/ml BSA to prevent non-specific binding. The primary antibody was added at a concentration of 1:10000 in PBS containing 5 % Tween. Following a 1 hr incubation at room temperature, the membrane was washed and then incubated with anti-rabbit IgG peroxidase conjugate (A6154) as per the manufacturers' instructions. The immunoreactive bands were visualised by the addition of 0.5 mg/ml

3,3'-diaminobenzidine tetrahydrochloride (DAB) (Sigma Aldrich) with 10  $\mu$ L hydrogen peroxidase in 50 ml of phosphate buffered saline (PBS) (Appendix 1).

### **3.2.3.3. Tryptic digest MS-MS.**

Purified peptide (2  $\mu$ g) was resuspended in 50 mM ammonium bicarbonate pH 7.8 containing 2 M urea and digested overnight with 0.1  $\mu$ g trypsin. The resulting peptides were cleaned up using C18 Zip Tips (Millipore) in readiness for analysis by either MALDI-MS or ES-MS. MALDI-MS analysis was performed using a MALDI-TOF instrument (Waters-Micromass) on samples dried on to the target plate using a saturated solution of alpha-cyano-4 hydroxycinnamic acid in 50 % acetonitrile/0.1 % trifluoroacetic acid. For ES-MS analysis, the sample was infused into the nano electrospray source of the mass spectrometer (Waters Q-ToF micro), at a flow rate of 150 $\mu$ l/hour, via a gas tight syringe. The positive ion mass spectrum of the sample was recorded.

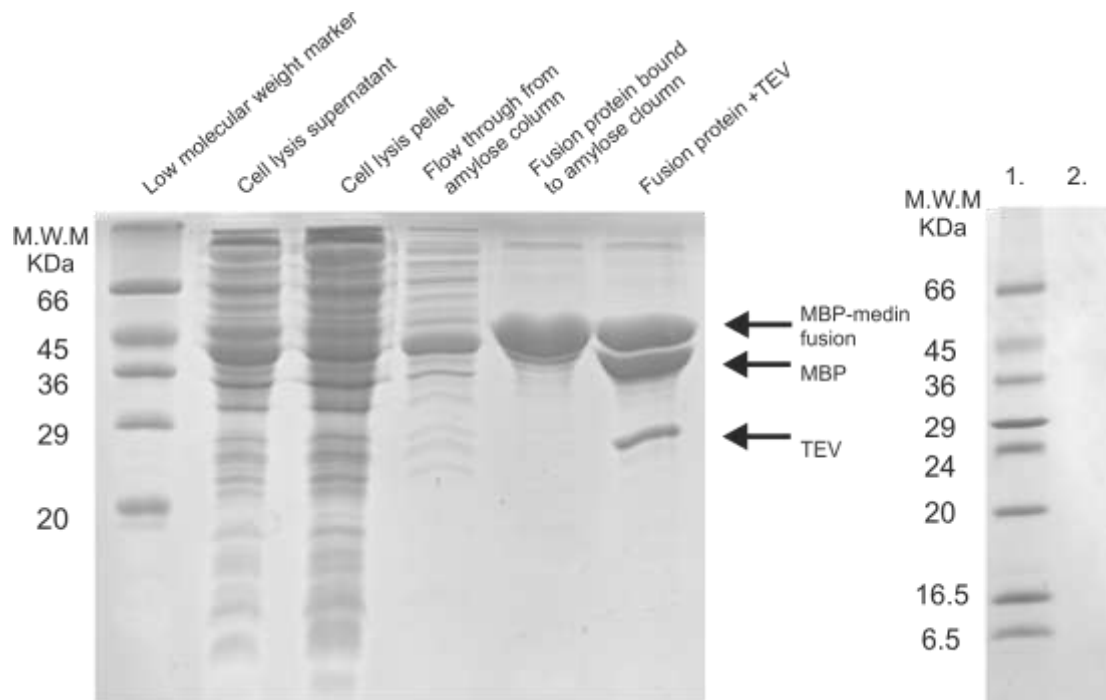
## **3.3. Results and Discussion**

### **3.3.1. Cloning, expression and purification of medin**

#### **3.3.1.1. pMALC2X construct**

The pMALC2X construct was transformed into BL21 (DE3) competent cells for expression of the fusion protein. BL21 (DE3) cells are a widely used expressing strain of *E.coli*. as they have reduced protease activity which often results in high protein yields. Expression tests are a quick and efficient way of identifying the optimal expression conditions for a construct and for pMALC2X these conditions were shown to be 3 hours incubation at 30 °C. The yield of fusion protein was further improved by the use of TB media instead of LB. Despite the use of the MBP-tag to improve solubility, approximately half of the expression fusion protein was

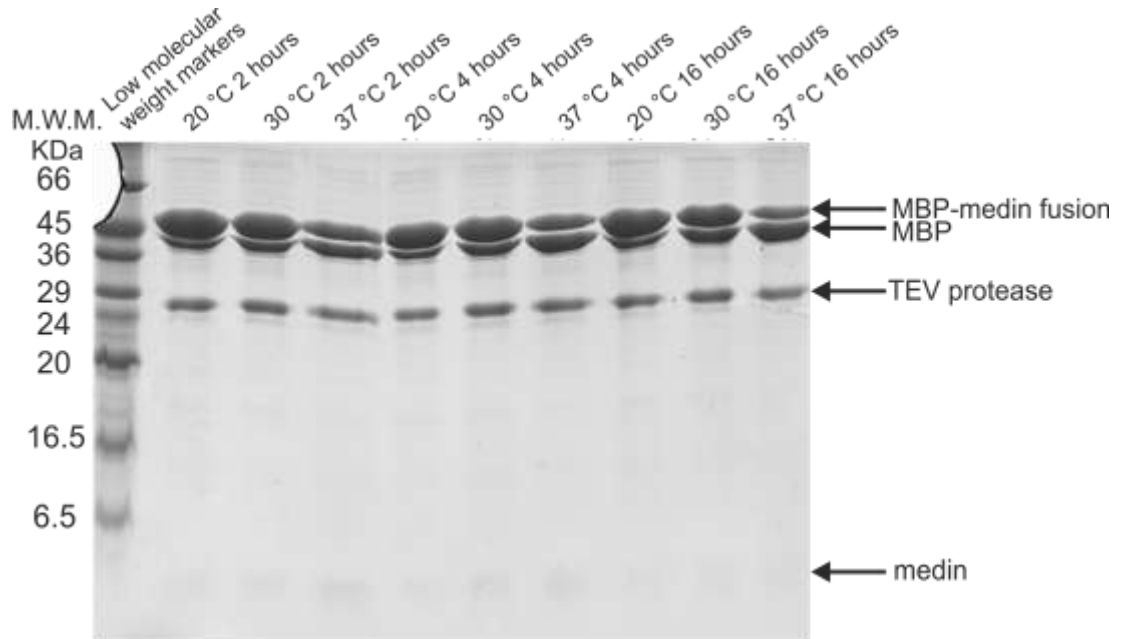
found in the pellet; nevertheless, the yield of fusion-protein was sufficient to continue with only the soluble portion (Figure 3.4, lanes 2 and 3).



**Figure 3.4 15 % SDS-PAGE of the expression and purification of medin in the pMALC2X vector. Gel on the right demonstrates the improved performance of the increased amylose column. Lane 1. Low-molecular weight markers, lane 2. flow-through from the 10 ml amylose column showing full fusion protein capture. Both gels were stained with Coomassie Brilliant blue stain.**

The affinity purification using amylose resin to bind the MBP tag was successful (Figure 3.4, lane 5) with few contaminants remaining after this initial stage of purification. However the yield was so high that it exceeded the binding capacity of the amylose column and appeared in the flow through (Figure 3.4, lane 4). This problem was overcome by increasing the bed volume of the column from 5 ml to 10 ml of amylose resin, thus increasing the binding capacity, which then resulted in 100 % capture (right-hand gel, lane 2, Figure 3.4). There was no fusion protein observed in the flow through.

Cleavage of the fusion protein by TEV protease was inefficient under normal conditions (overnight incubation at 4°C) and resulted in only approximately 50 % (Figure 3.4, lane 6). As such a cleavage trial of different temperatures and incubation times was carried out in an attempt to improve the yield (Figure 3.5).

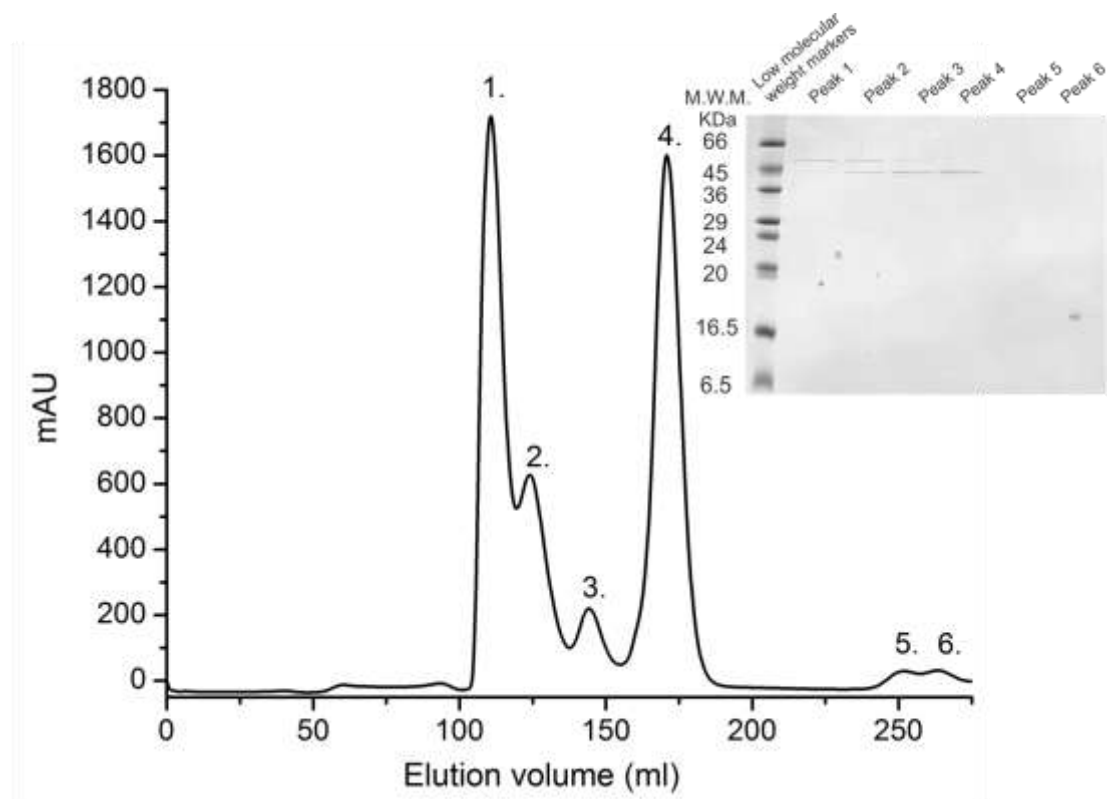


**Figure 3.5 15 % SDS-PAGE of the MBP-fusion protein cleavage trial results, each lane represents a different condition. Gel is visualised using Coomassie Brilliant blue stain.**

The best cleavage was observed at 37 °C after 16 hrs of incubation (Figure 3.5, lane 10.) There was also a large amount of precipitate under these conditions therefore it was decided that cleavage at 37 °C for 4 hrs was the best compromise between cleavage and the amount of precipitate. Normally TEV protease demonstrates very efficient cleavage and it is unknown why it is poor in this case (Kapust and Waugh, 2000). There are a few possible explanations for this poor cleavage. Medin, like many amyloid proteins, is highly hydrophobic and it may be wrapping around and

'sticking' to the MBP and therefore occluding the cleavage site. A second possibility is that fusion protein is aggregating, again blocking the cleavage site. There have been reports of MBP aggregation (Zanier et al., 2007) and as expected, medin is prone to self-assembly.

Size exclusion chromatography was used to isolate medin from the cleavage mixture. This was considered to be a suitable purification method as there is a large difference in molecular weight between the proteins, 45 kDa and 5.5 kDa for MBP and medin respectively. Initially the protein mixture was concentrated down to 2 ml in an Amicon® Ultra-15 centrifugal filter with a 30 kDa cut off (Millipore) to ensure good resolution the column. It was loaded onto a Superdex 75 26/60 gel filtration column (GE Healthcare) and all the peak fractions were kept for SDS-PAGE analysis. As can be seen in Figure 3.6 (peak 1), a large amount of protein elutes in the void volume suggesting that the species present is larger than 75 kDa. Theoretically the largest species present would be any intact fusion protein with a molecular weight of 48 kDa which should not elute in the void volume. Furthermore the size of the peak suggests that it is unlikely to be a contaminant. SDS-PAGE analysis confirms that this peak contains un-cleaved fusion protein. It is likely that the fusion protein has aggregated and is therefore unable to enter the column matrix but dissociates under the denaturing conditions of SDS-PAGE. This would also be consistent with occlusion of the cleavage site.



**Figure 3.6 Gel filtration chromatogram of cleavage mixture. Concentrated cleavage mixture was run on a Superdex 75 26/60 gel filtration column (GE Healthcare) in 20 mM Tris, 0.2 M NaCl, pH 7.4 at 2 ml/min. 15 % SDS-PAGE analysis of peaks**

Peak 4 contains MBP alone (peak 4, inset, Figure 3.6). The two smaller peaks (2 and 3) at ~120 and ~140 ml are shown to include both un-cleaved and isolated MBP (Figure 3.6). It is probable that they represent different sized aggregates or conformers. Medin should elute at approximately 250 ml, a small peak can be seen at the correct elution volume, yet the intensity is not what would be expected and there was no protein band visible following SDS-PAGE analysis of the peak. This may again be a result of aggregation; however it is more likely that hydrophobic nature of medin has caused it to adhere to the filter in the concentration step or prior to loading. Due to the numerous problems with the purification, it was decided to sub-

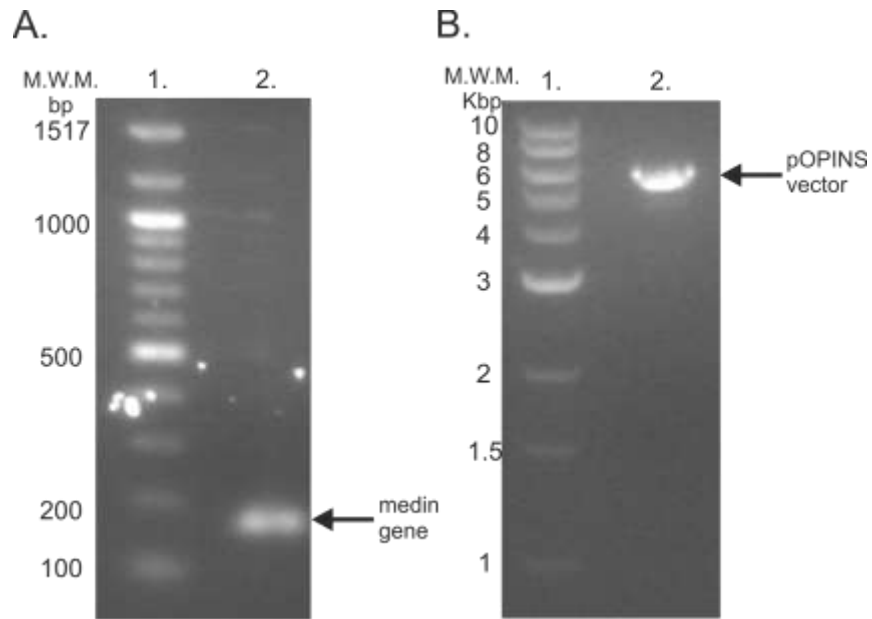
clone medin into an alternative vector that may improve or eliminate the problems encountered with the pMALC2X vector.

#### **3.3.1.2. pOPINS construct**

The pOPINS vector was chosen because the fusion partner His<sub>6</sub>-SUMO is highly soluble and a more comparable size to medin. Furthermore the SUMO protease recognises tertiary structure rather than sequence and therefore may overcome some of the cleavage problems that occurred with the pMALC2X construct. Moreover the His<sub>6</sub> element provides a simple method of fusion protein purification and subsequent removal of the tag (Figure 3.2).

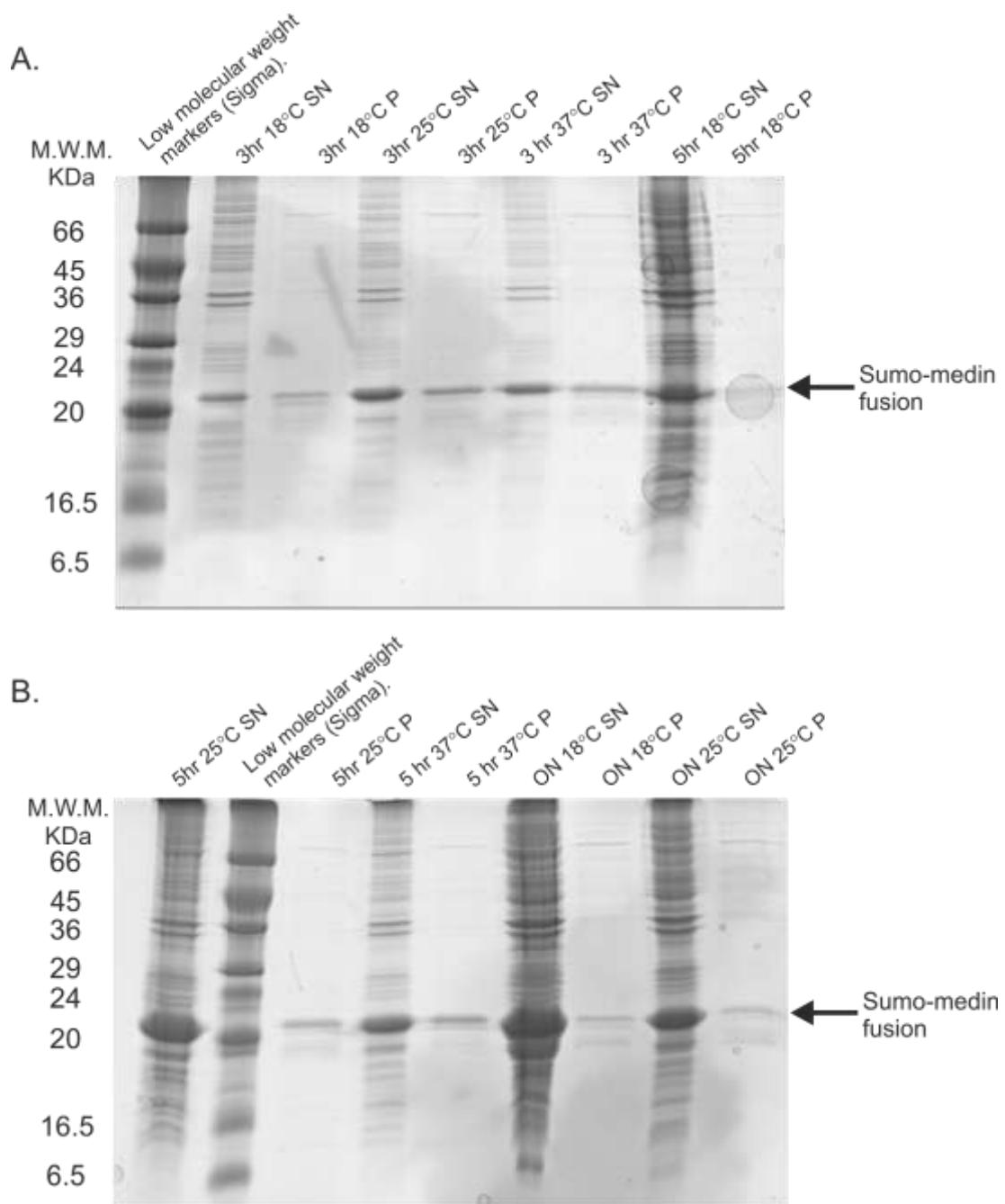
The medin sequence was amplified using the vector specific primers successfully as confirmed by agarose gel electrophoresis, (Figure 3.7, panel A). The pOPINS vector was grown, purified and linearised as described in (Berrow et al., 2007). This was confirmed by visualisation on an agarose gel with a strong band at ~ 5.5 Kbp (Figure 3.7). The In-Fusion™ reaction was successfully carried out and the resultant mixture transformed into XL1 Blue cells and plated onto an LB-Agar plate containing kanamycin, XGal and IPTG. Following an overnight incubation, a mixture of blue and white colonies was observed. Several white colonies were selected, grown and the plasmids purified and sent for sequencing. The sequencing data confirmed the insertion of the medin gene into the pOPINS vector (Appendix 2).





**Figure 3.7 Agarose gel of sub-cloning elements. Amplified medin gene generated using the vector specific primers (A). Purified linearised pOPINS vector (B).**

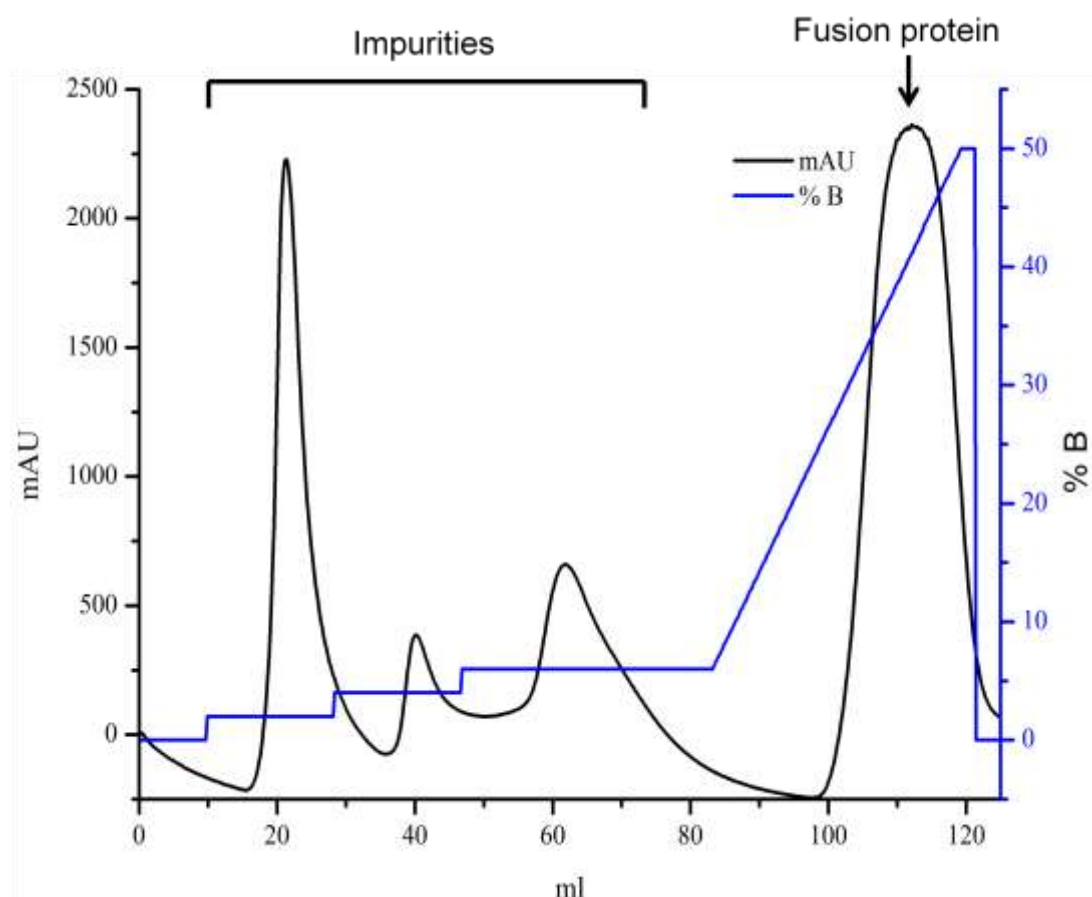
The pOPINS vector containing the medin gene was transformed into BL21 (DE3) cells for expression. Initially expression tests were carried out to identify the optimal conditions of growth and expression.



**Figure 3.8 Expression trials of medin-pOPINS vector in BL21 (DE3) cells. Lanes 1 and 11. Low molecular weight markers (Sigma Aldrich). All other lanes represent different growth conditions (SN = supernatant, P = pellet).**

These tests indicated that the optimal conditions for yield and solubility were overnight incubation at 18 °C (Figure 3.8, panel B, ON 18°C SN). These conditions were also optimal for expression in isotopically labelled media. The cells were harvested and lysed as described for the pMALC2X vector. The initial his-trap™

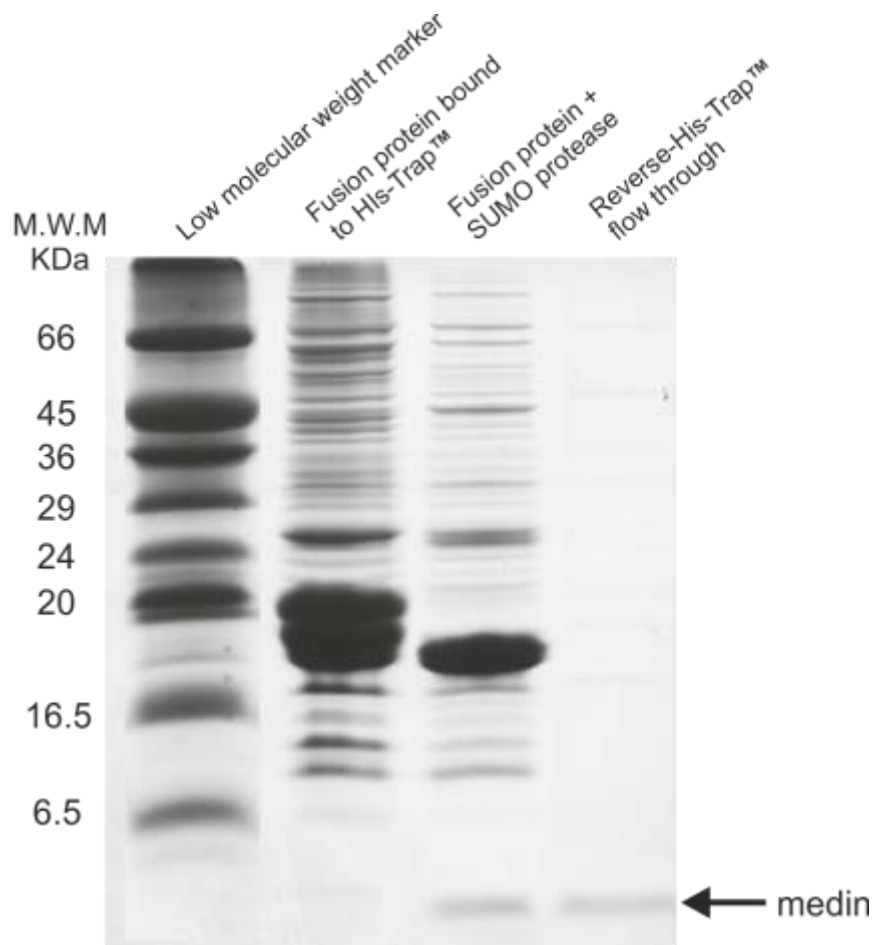
purification was very effective at isolating the fusion protein. The small step gradients eliminate any non-specific binding and then a large amount of protein elutes at the upper end of the imidazole gradient consistent with the fusion protein (Figure 3.9).



**Figure 3.9** Affinity purification chromatogram of the fusion protein isolation. Sample was run on a 5 ml his-trap FF column (GE Healthcare) in his trap buffer A with a gradient of his-trap buffer B (Appendix 1).

SDS-PAGE analysis of the large peak indicated the presence of three large protein bands approximately 20 kDa in size that bound to the his-trap™ column (Figure 3.10, lane 2). It may be that these represent three different conformations of fusion protein. The pooled fractions containing the fusion protein were desalted back into his-trap buffer A prior to cleavage. Cleavage was carried out at 25 °C for 2 hours,

approximately 60 % of the fusion protein was cleaved within this time (Figure 3.10, lanes 2 and 3). It appears that the upper two protein bands are cleaved successfully, however the lower protein band remains un-cleaved. As it is has run the furthest through the gel matrix, it is possible that in this third conformation the medin peptide is wrapped tightly around the SUMO fusion partner making it more compact and thus enabling it to travel further than its counterparts. As a consequence the cleavage site may be inaccessible to the SUMO-protease.



**Figure 3.10 Tris-Tricine SDS-PAGE analysis of the purification and cleavage of medin within the pOPINS construct**

### **3.3.2. Characterisation of medin**

#### **3.3.2.1. Visualisation**

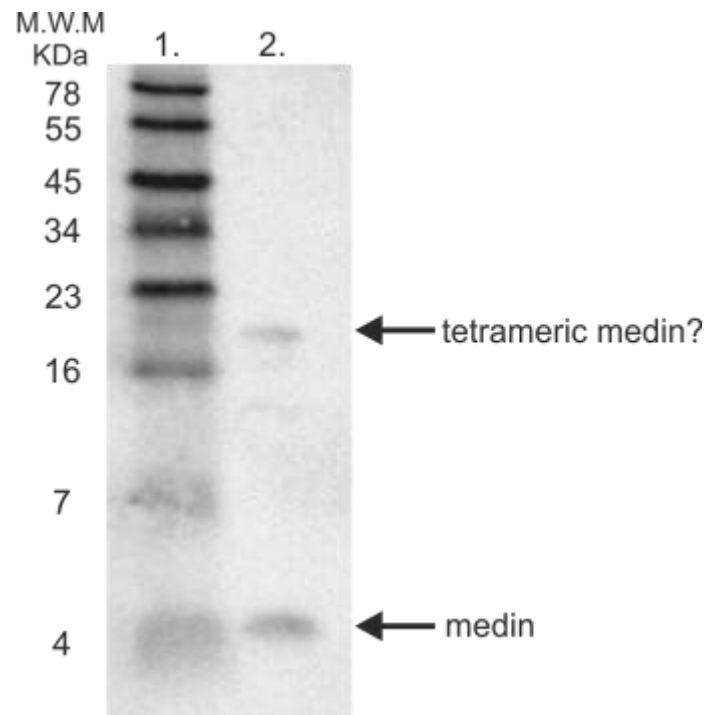
Throughout the purification, it has been difficult to observe medin upon SDS-PAGE analysis with normal Coomassie staining. Initially it was thought that the concentration might be too low. Therefore, for the cleavage trail, the lanes were overloaded to identify if this was the problem (Figure 3.5). Faint bands can be seen in Figure 3.5 at the appropriate molecular weight (5.5 kDa), however, they are at a much lower intensity than expected for the amount of protein loaded. The next line of investigation was to explore alternative stains to eliminate the unlikely possibility that medin does not bind Coomassie stain. This is often seen for small amyloid peptides including A $\beta$  (M. Pandya, personal communication). Ponceau red and silver staining was tried with limited success. The third possibility was that medin, as a small protein, was washing out of the gel during staining and destaining. Several methods were attempted to investigate this possibility; TCA precipitation, gluteraldehyde cross-linking and membrane transfer. The aim of all of these methods was to immobilise the protein in some way to prevent leaching during wash steps. Unfortunately, all of these methods proved to be unsuccessful, in that they were no more effective than the initial overloading.

The best method for visualising medin is to overload lanes of a tris-trycine gel as can be seen in Figure 3.10. The denser gel helps to prevent lateral diffusion within the gel and concentrates the sample. This can then be stained with Coomassie stain and destained as described previously. Although this method did provide positive results it was highly dependent on concentration and only very concentrated samples were visible. It was not possible to concentrate medin to high levels as the process

accelerated aggregation and resulted in high order aggregates that did not enter the stacking gel.

### 3.3.2.2. Western blot analysis

Further verification can be obtained by Western blotting with specific medin antibodies, thus providing confirmation that the purified product is indeed medin. Figure 3.11 shows the presence of two immunoreactive bands present in the flow through following reverse purification.



**Figure 3.11 Western blot analysis of medin. Lane 1. SeeBlue prestained markers (Invitrogen). Lane 2. Flow through from reverse-his purification indicating two immunoreactive bands at approximately 5 kDa and 20 kDa.**

The band at approximately 5 kDa corresponds to monomeric medin. The second fainter band at approximately 20 kDa could be tetrameric medin or a small amount of uncut fusion protein.

### 3.3.2.3. Tryptic digest and mass spectrometry

Additional confirmation of the sequence was confirmed by tryptic digest and subsequent mass spectrometry. Results from the digest show the presence of peptides with masses that match those predicted for the medin sequence (Table 3.1). This data suggest that the purified peptide is indeed medin. The only expected peptide missing was the C-terminal fragment with the sequence NH<sup>2</sup>-NFGSVQFVA-COOH. This peptide has previously been identified as the most amyloidogenic fragment of medin (Larsson et al., 2007), and it is possible that upon cleavage it aggregates very rapidly and therefore is no longer amenable to mass spectrometry.

**Table 3.1 Tryptic digest fragments, depicting the sequence, expected mass from digest (calculated using MASCOT server) and actual masses obtained from MALDI and ES-MS. Fragments, R and LDK are too small to reliably obtain masses using MALDI and ES-MS.**

Fragment sequence	Expected mass	MALDI	ES-MS
R	-	-	-
LDK	-	-	-
QGNFNAWVAGSYGNDQWLQVDLGSSK	2841.3	2841.2	
EVTGIITQGAR	1144.6	1144.5	1144.5
NFGSVQFVA	968.4		

## 3.4. Discussion

Aortic medial amyloid is the most common form of localised amyloid yet there remains no published method for the large scale expression and purification of medin. Chemically synthesised proteins are expensive to purchase and often difficult

to obtain therefore it is preferable to use an expression host which is more cost effective and provides a replenishable source of peptide. Furthermore, there is evidence that synthetically produced and recombinant proteins behave differently *in vitro* (Finder et al., 2010). The lack of a reliable expression protocol has no doubt limited research on medin and AMA. The objective of this chapter was to establish expression and purification protocols for the production of recombinant medin for future characterisation studies. Moreover, the use of an *E.coli* expression host enables the relatively facile incorporation of isotope labels required for NMR structural analysis.

Medin was ultimately cloned into the pOPINS vector and expressed in BL21 (DE3) *E.coli* cells. Small proteins are often susceptible to proteolytic degradation; the use of a fusion protein can help protect the peptide from degradation. The use of affinity or solubility tags can also assist in the purification. The pOPINS vector contains both a His<sub>6</sub> affinity tag and a SUMO solubility tag and these tags improved the expression and purification of medin. Overall, this system produced approximately 4 mg of purified medin per litre of culture.

One of the biggest challenges of this work was confirmation and visualisation of the purified product. It seems likely that the difficulties in visualising medin by SDS-PAGE are likely a result of two effects, low concentration and a degree of washing out during staining. Similar difficulties have been observed for other amyloid peptides (personal communication, M.,Pandya,). This is also reflected in the numerous alternative staining techniques described in the literature (Walsh et al., 2009). Attempts to overcome this by concentrating medin to high levels were hampered by aggregation. Despite these issues, it was possible to visualise a protein band that correlated with the correct molecular weight for medin in some tris-trycine



gels (Figure 3.10). It was also possible to confirm the sequence by a combination of Western blot analysis (Figure 3.11) and tryptic digests coupled with mass spectrometry (Table 3.1).

This work also demonstrates for the first time the expression of uniform  $^{13}\text{C}/^{15}\text{N}$  isotopically labelled medin suitable for NMR structural studies. Expression in minimal media produced 3 mg of uniform  $^{13}\text{C}/^{15}\text{N}$  isotopically labelled medin per litre of culture.

Many amyloid related diseases are caused by mutated variants of amyloid proteins (Connors et al., 2000). These variants can have different aggregation kinetics and structures (Tycko et al., 2009). This recombinant expression system has now been adapted to express mutants of medin (see Chapter 5). This could help identify key residues within the medin sequence that are important for aggregation and could ultimately become therapeutic targets.

These protocols will now enable in depth studies into the biophysical and structural properties of medin and could lead to a greater understanding of how it affects the physiology of the aorta and other blood vessels.

## 4. Biophysical characterisation of medin aggregation

### 4.1. Introduction

Despite the ubiquitous nature of aortic medial amyloid, remarkably little is known about how medin aggregates into amyloid fibrils. Preliminary information on the kinetics and morphology of synthetic medin was presented in chapter 2. The aim of this chapter is to characterise the aggregation of medin in greater detail using a range of biophysical techniques and recombinantly prepared peptide.

Previous biophysical studies carried out on medin have identified that the C-terminus with the sequence NH<sub>2</sub>-NFGSVQFV-COOH, is highly amyloidogenic in isolation and shares similarity with fragments of the islet amyloid polypeptide (IAPP) (Haggqvist et al., 1999, Reches and Gazit, 2004). The region NH<sub>2</sub>-NFGSV-COOH was identified as the minimum motif for amyloid formation. Additionally, the role of the phenylalanine residues in the aggregation process was investigated and when substituted for non-aromatic residues, resulted in a reduction of aggregation potential. This suggests that the two phenylalanine residues could have an important role in the aggregation process, possibly through  $\pi$ -stacking interactions (Reches and Gazit, 2004).

Studies by Larsson *et al.* (2007) used a variety of synthetic and recombinant fragments and full length medin and tested them for their ability to bind Congo Red, Thioflavin T and to form fibrillar structures visible by electron microscopy. They concluded that the C-terminal 18-19 residues comprised the amyloid promoting region and that full length medin aggregated after approximately 30 hours (Larsson et al., 2007). They further suggested that the two phenylalanine residues discussed by

Reches and Gazit (2004) are not essential for fibril formation of the full length protein. This work closely examined the sequence requirements for fibril formation but did not investigate the kinetic or structural aspects of medin aggregation.

Additional work by Madine and Middleton demonstrated that full length synthetic medin aggregated over a period of 8 days and the fluorescence intensity, size of aggregates and toxicity of the resultant species could all be modulated by the addition of inhibitor compounds or glycosaminoglycans (Madine and Middleton, 2010). Similar to the work conducted by Larsson et al. (2007) the studies by Madine and Middleton (2009) did not aim to characterise the aggregation of medin in detail, rather, investigate the modulation of medin aggregation. Both studies have used ThT as the primary technique to investigate medin aggregation. Furthermore, these studies have principally been done on synthetically produced medin often in the presence of organic solvents which do not accurately represent the physiological environment.

Here, a range of biophysical techniques are used to study the aggregation of recombinant medin. They can be broadly divided into techniques that provide information on the mechanism and kinetics of aggregation and the structure and morphology of aggregates.

Arguably the primary technique used to study amyloid kinetics is the thioflavin T (ThT) assay. ThT is a cationic benzothiazole dye that shows an increase in fluorescence upon amyloid binding (Vassar and Culling, 1959). In the free form, ThT has excitation and emission values of 385 nm and 445 nm respectively. However, binding to amyloid gives rise to a new excitation maximum at 450 nm, and enhanced emission at 482 nm (Levine, 1993). The specificity of ThT for amyloid has led it to be considered as a standard technique for the study of amyloid aggregation growth

kinetics (Naiki et al., 1989). Although medin has been studied using the ThT assay previously the conditions under which it was examined are highly variable and further work is warranted to establish reproducible data under physiologically relevant conditions.

Dynamic light scattering (DLS) can be used to monitor the size of particles over the aggregation period (Gast and Modler, 2005, Goldberg, 1999). The combination of ThT and DLS has successfully been used to gain information about the kinetics and mechanism of poly Q peptides (Streets et al., 2013).

Intrinsic fluorescence (IF) is a technique that exploits the fluorescent properties of certain amino acids: namely Trp, Tyr and Phe. The maximum emission fluorescence of these amino acids is sensitive to the local environment and can provide information about structural changes during aggregation. IF is often used to probe protein folding and dynamics (Huiyuan, 2009) and has been used to study the aggregation process in many different amyloid systems (Maji et al., 2005, Munishkina and Fink, 2007, Amaro et al., 2011). IF is used here to monitor changes in the tryptophan environment over the aggregation time course.

Circular dichroism is a spectroscopic technique that can be used to study the secondary structure of proteins and peptides (Huiyuan, 2009). CD is an averaging technique and provides information about the conformational state of all the molecules within a given sample (Whitmore and Wallace, 2008). Quantitative analysis of the secondary structure elements can be achieved by deconvolution of experimental spectra using libraries of known structures (Whitmore and Wallace, 2008). CD can be used to monitor the changes in secondary structure associated with aggregation and amyloid formation. For example, CD analysis identified the presence of  $\alpha$ -helical intermediates present during A $\beta$  amyloidogenesis (Kirkitadze

et al., 2001). Furthermore CD can be used to identify structural changes associated with pathogenic mutations (Adachi et al., 2013). Here, CD is used to monitor amyloid formation over time and thus provides both structural and kinetic information.

The morphology of amyloid aggregates can be studied by transmission electron microscopy (Gras et al., 2011). This technique is commonly used to study amyloid fibrils and intermediate species in the amyloid pathway and can provide information about structural features of different amyloid assemblies (Petkova et al., 2005a, Fändrich, 2012).

The dynamic process of amyloid fibril formation and the non-crystalline nature of fibrillar protein aggregates make these proteins difficult to study and characterise. It is therefore necessary to use a range of relatively low-resolution techniques in combination to obtain information about the kinetic and structural properties of amyloid proteins which can assist in the development of effective diagnostics and therapeutics.

The aim of this work was to use a range of techniques to examine the aggregation kinetics and structural characteristics of full-length recombinant medin under physiological conditions. Moreover, this work aimed to use different experimental conditions to learn more about the inhibitory compounds described by Madine and Middleton (2009) and in doing so investigate their effect on recombinant medin in aqueous conditions.

## 4.2. Materials and methods

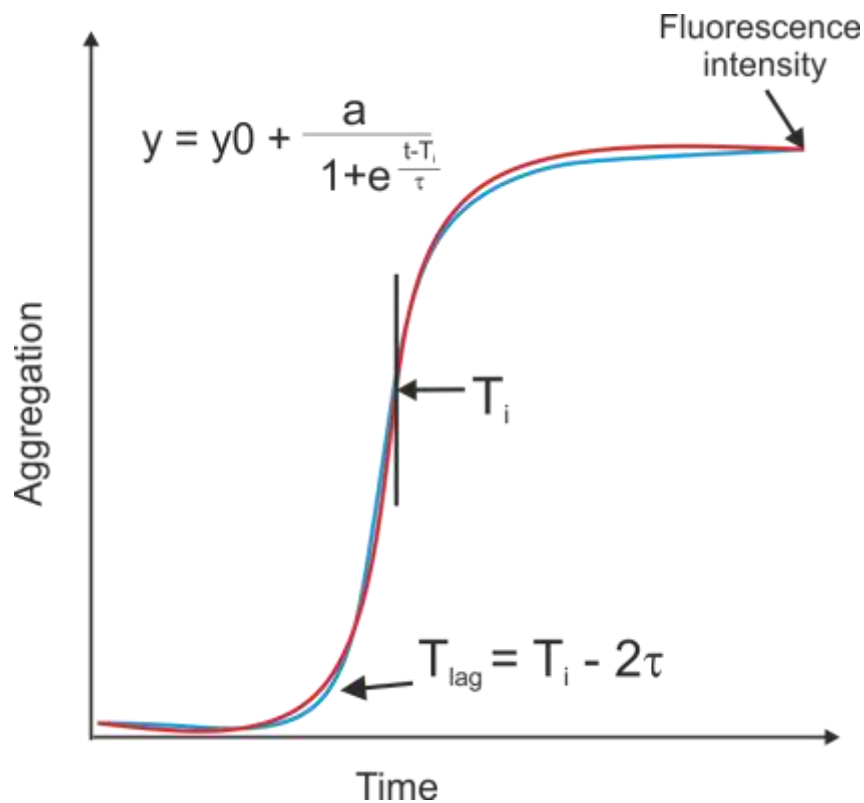
### 4.2.1. Aggregation kinetics studied by Thioflavin T

ThT assays were carried out on a Flexstation 3 microplate reader (Molecular Devices Ltd.). Experiments were carried out in 96-well blackwalled, clear bottomed microplates (Nunc) with data recorded using bottom read mode. The use of a plate reader during this assay enables many more points to be sampled compared to the method described in section 2.2.4.1. ThT solution was injected into the protein samples (20  $\mu$ M peptide concentration in aggregation buffer) at the start of the read, to a final ThT concentration of 20  $\mu$ M in aggregation buffer (Appendix 1). The wells were mixed with two cycles of trituration prior to recording data. Fluorescent measurements, representing the average of 6 flashes, were recorded every 5 minutes throughout the duration of the experiments, with the photomultiplier tube sensitivity set to medium. The dye was excited at 450 nm and the emission signal recorded at 485 nm with a cutoff of 475 nm. Unless otherwise stated, the assay was conducted at 30 °C with no agitation. The resulting data was then analysed and the lag time and fluorescence intensity were recorded. Kinetics were analysed from the data using the equation 4.1 (Alvarez-Martinez et al., 2011):

$$y = y_0 + \frac{a}{1 + e^{-\frac{t - T_i}{\tau}}} \quad 4.1$$

where,  $y$  is the polymerisation,  $y_0$  is the starting fluorescence,  $a$  is the difference between  $y_0$  and the fluorescence maximum,  $t$  is time,  $T_i$  is the inflection point of the sigmoidal, and the slope  $1/\tau$  is the rate of polymerisation (Figure 4.1). The lag time can be calculated using equation 4.2:

$$T_{lag} = T_i - 2\tau \quad 4.2$$



**Figure 4.1** Schematic of an amyloid growth curve, illustrating the different kinetic points. Kinetic parameters were obtained by fitting equation  $y = y_0 + \frac{a}{1 + e^{\frac{t-T_i}{\tau}}}$  and the lag time was calculated using the equation  $T_{lag} = T_i - 2\tau$ .

#### 4.2.2. Aggregation kinetics by intrinsic fluorescence

The vast majority of proteins contain amino acids that are intrinsically fluorescent: namely phenylalanine, tyrosine and tryptophan. The intrinsic fluorescence (IF) of these amino acids, particularly tryptophan, are heavily influenced by their local environment. This property can be exploited to investigate the tertiary or quaternary structure of proteins (Vivian and Callis, 2001, Andrews and Forster, 1972, Lakowicz, 2006). Solvent-exposed tryptophan residues usually give a fluorescence spectrum with a maximum at 355 nm, whereas tryptophan residues buried in the hydrophobic core have blue shifted emission spectra in the range of 309-335 nm. In this study,

intrinsic fluorescence of the tryptophan residues at positions 11 and 21 was used to monitor aggregation by observing the shift in the spectrum from solvent-exposed, to buried in the aggregated form. Intrinsic fluorescence measurements were carried out on a Cary Eclipse Varian fluorescence spectrometer operating on a 20  $\mu$ M medin solution in aggregation buffer (Appendix 1) at 30 °C. Tryptophan residues were excited at 279 nm and the emission spectra recorded between 300 nm and 400 nm with a band pass of 5 nm. The fluorescence maxima at 355 nm and 335 nm were recorded, normalised and plotted as a function of time.

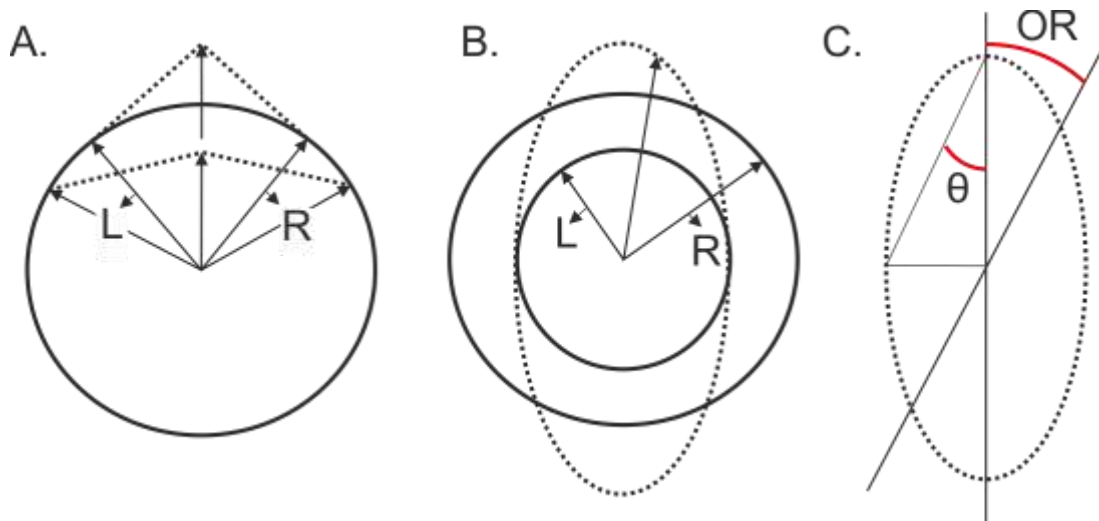
#### **4.2.3. Determination of aggregate size distribution by dynamic light scattering**

DLS measurements were carried out on a Zetasizer nano instrument (Malvern instruments) on both freshly prepared and aggregated 20  $\mu$ M medin samples in aggregation buffer. Each experiment was carried out in a 3 mm quartz cuvette at 25 °C and results are presented as average values.

#### **4.2.4. Secondary structure analysis by circular dichroism**

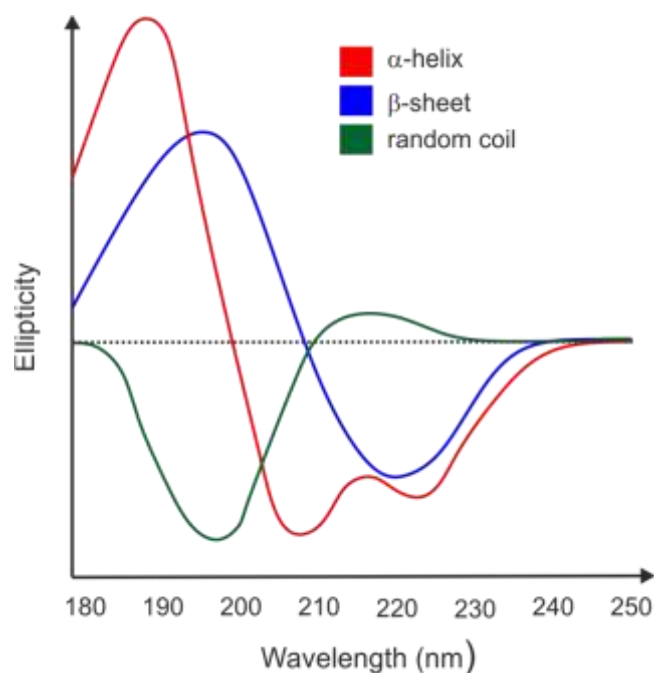
Circular dichroism (CD) spectroscopy is a technique that can be used to study the secondary structure of proteins and peptides. Circularly polarised light is generated when two linearly polarised states become out of phase with one another and rotates along an ellipsoidal path (Panels A and B, Figure 4.2). Circularly polarised light can be either left- (LCP) or right-handed (RCP); CD measures the difference in absorption of LCP and RCP. CD is measured in ellipticity where the ratio of the minor to major axis of the ellipse is the tangent of the angle ( $\theta$ ) and OR is the optical rotation (Panel C, Figure 4.2).





**Figure 4.2 Schematic of linearly polarised light (A), circularly polarised light (B). The origin of ellipticity values, where the ratio of the minor to major axis of the ellipse is the tangent of the angle ( $\theta$ ) and OR is the optical rotation (C.)**

CD is used here to examine the secondary structure of peptides and their stability over time. Each pure secondary structure element gives a signature spectrum in far – UV CD (Figure 4.3). The secondary structure content of spectra can be analysed using various fitting algorithms such as CONTILL and CDSSTR (Van Stokkum, 1990, Sreerama, 2000, Manavalan, 1987) and using a basis set of spectra of appropriate proteins.



**Figure 4.3 Typical CD spectra of pure secondary structure elements.**

Changes in secondary structure of medin during aggregation were monitored using far-UV synchrotron radiation circular dichroism (SRCDD). Experiments were carried out at beam line B23 at Diamond light source (Oxford, U.K.). Freshly prepared medin was incubated at a concentration of 200  $\mu\text{M}$ , at 30  $^{\circ}\text{C}$ , in the cuvette (121.000-QS Hellma, Southend-on-Sea, U.K.) for the duration of the time course. Single scans were recorded every two hours between 180 and 190 nm, using a slit width of 0.5 mm, 0.5 cm path length and a scan rate of 1 nm/sec, over 44 hours. The baseline was subtracted prior to secondary structure analysis. Analysis was carried out using Olis<sup>®</sup> GolbalWorks software. Data at each time point was subject to two fitting methods, CONTILL and CDSSTR, using either basis sets 8 or 11 (Van Stokkum, 1990, Sreerama, 2000, Manavalan, 1987). The best fit, as determined by the normalised spectral fit standard deviation, was selected and the percentage  $\alpha$ -helix,  $\beta$ -sheet, turn and random coil content recorded.

#### **4.2.5. Analysis of morphology by transmission electron microscopy**

Morphologies of the insoluble medin aggregates were visualised using transmission electron microscopy (TEM). Peptide suspensions (10  $\mu$ L) were loaded onto carbon coated copper grids held between tweezers. Excess solution was removed by placing the grid perpendicular to blotting paper. The grid was then inverted onto a droplet of dH<sub>2</sub>O to wash of excess peptide aggregates. The grid was blotted again before being inverted onto a droplet of 4 % uranyl acetate to negatively stain the aggregates. Excess stain was removed by blotting before allow the grid to air dry prior to visualisation. The grids were observed using a Technai 10 electron microscope operating at 100 kV.

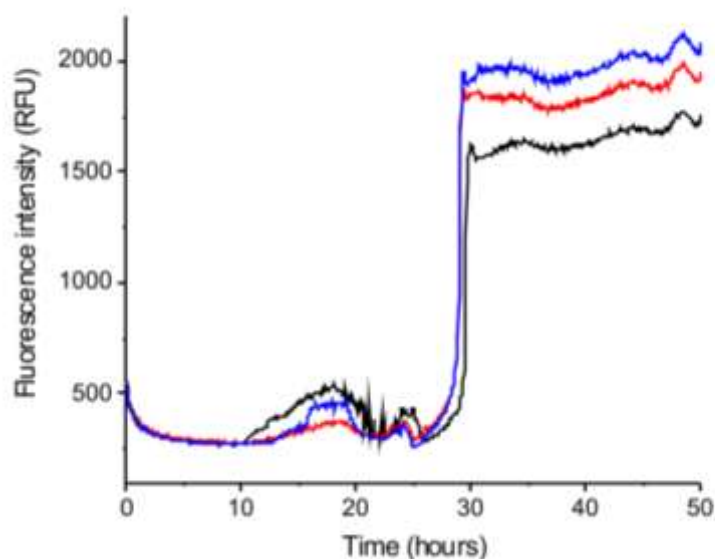
## 4.3. Results

### 4.3.1. Aggregation kinetics of medin

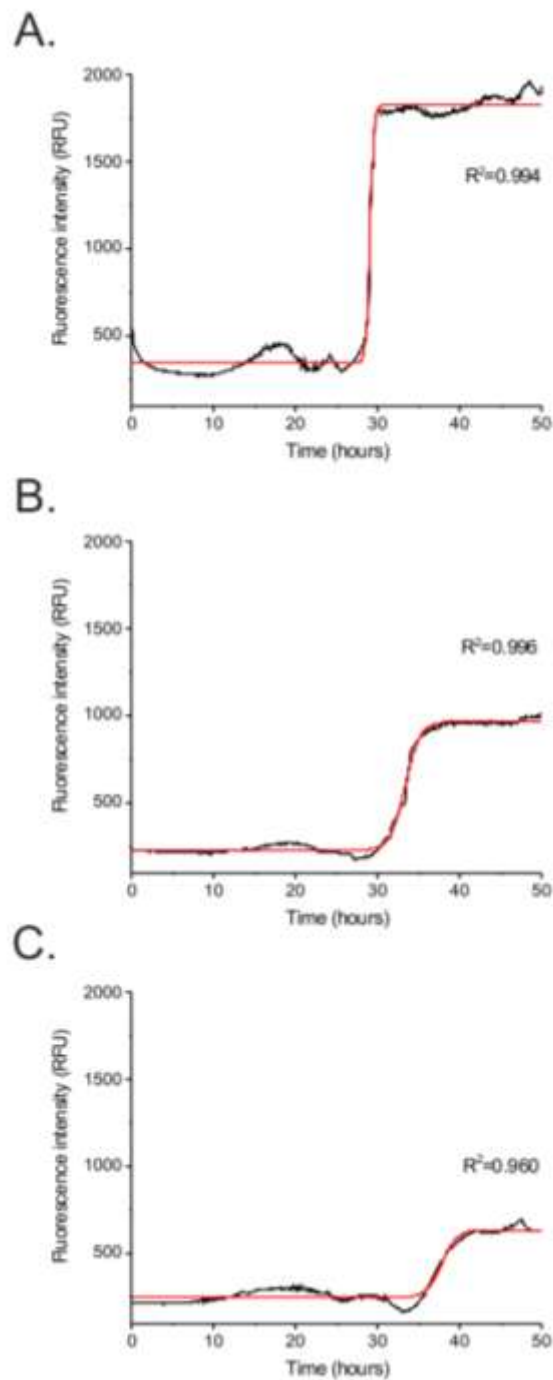
The primary aim of this section of work was to establish the aggregation characteristics of recombinant wild type medin, to enhance understanding of this ubiquitous protein and provide reproducible conditions for future testing of inhibitors of medin aggregation.

#### 4.3.1.1. Investigating the concentration dependence of medin aggregation

Initially the ThT binding assay was used to see if our recombinant medin was capable of binding ThT and to monitor the aggregation over time. These measurements were carried out in triplicate an example of three replicate data sets is shown in Figure 4.4. Each curve was fitted using the equation described in section 4.2.1. and information about the lag time, rate constant and overall fluorescence intensity collated.



**Figure 4.4** Three replicate ThT data sets for medin incubated at 200  $\mu\text{M}$  in aggregation buffer. Data were recorded on a fluoromax spectrometer at 30  $^{\circ}\text{C}$  under quiescent conditions.

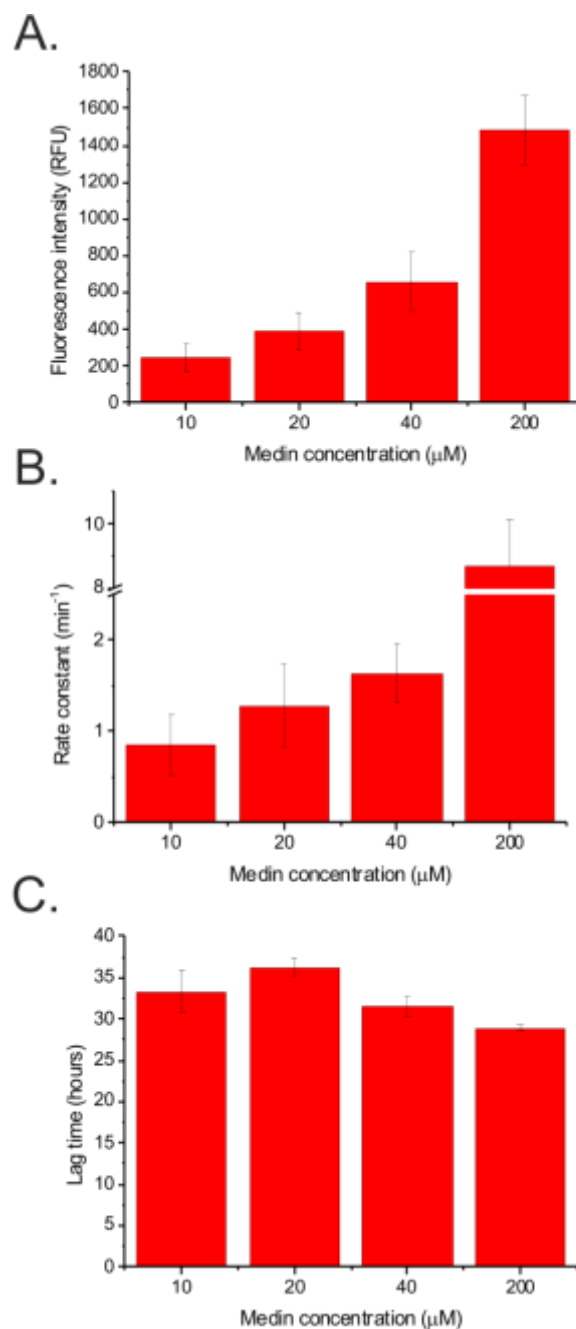


**Figure 4.5** The effect of protein concentration upon medin aggregation measured by thioflavin T fluorescence. Data were recorded on a fluoromax spectrometer at 30 °C under quiescent conditions at the following medin concentrations: 200, 40 and 20  $\mu$ M (A, B and C respectively). The red line represent the fit using the following equation:  $y = y_0 + \frac{a}{1 + e^{-\frac{t-T_i}{\tau}}}$ . The adjusted  $R^2$  values indicating the goodness of fit for each curve are reported on the right hand side. Data represent mean data for 3 replicates.

Different concentrations of medin, (10, 20, 40, and 200  $\mu\text{M}$ ) were incubated with ThT and monitored over time. All conditions gave rise to a characteristic rise in fluorescence intensity over time (Figure 4.5). Data were then processed as described in section 4.2.1, with the lag time, rate constant and fluorescence intensity calculated and plotted against concentration (Figure 4.6). Fluorescence intensity appears to correlate with concentration as expected (Panel A, Figure 4.6). The rate constant values also correlate with concentration (Panel B, Figure 4.6). These data indicate that wild type medin aggregates after approximately 33 hours under the conditions used here (Panel C, Figure 4.6). This is consistent with previous studies on recombinant medin which suggested a lag time of approximately 30 hours (Larsson et al., 2007). It is also consistent with previous data carried out on synthetically produced medin. Madine *et al.* showed that synthetic medin incubated at 0.25 mM concentration with agitation, demonstrated a dramatic increase in fluorescence intensity after three days. The ThT assay was carried out at discrete time points, (0, 3 and 8 days) so we are unable to determine if the nucleation point is comparable (Madine and Middleton, 2010). They also showed that there was a further smaller increase in fluorescence intensity after 8 days of incubation; however it was not possible to assess recombinant medin over such a long time course in the plate reader. Overall, the lag time does not significantly change with concentration which is surprising, it may be due to the specific conditions of the assay and further work would be required to investigate this.

These experiments were also used to identify the optimal concentration of medin for future experiments; 20  $\mu\text{M}$  was selected as it gave a good compromise between the amount of material required and the observed fluorescence intensity under these

conditions. Furthermore, it is comparable with conditions used to study other amyloid proteins using this technique (Schubert et al., 1995, Cribbs et al., 2000).



**Figure 4.6 Analysis of the effect of protein concentration on fluorescence intensity (A), lag time (B) and rate constant (C). Values were extracted from fitting exponential curves to the data and recording the value at the base of the exponential phase, the maximum fluorescence intensity and the calculated rate constants for each condition. Data represent mean and standard error values for 3 replicates.**

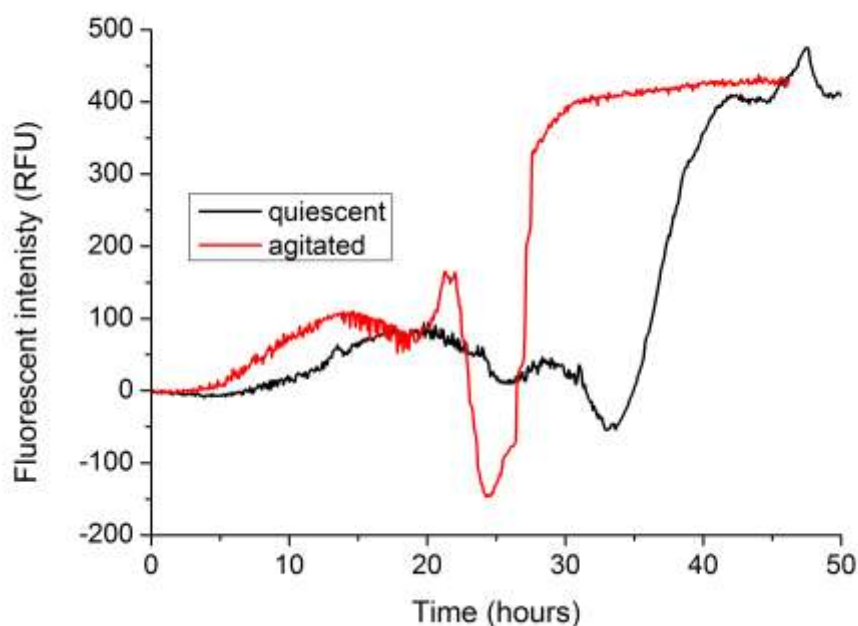
Interestingly, there are also more subtle differences that occur within the lag time. At 200  $\mu\text{M}$  concentration there is a small hump followed by a dip that precedes the exponential increase and they are reproducible. This is also seen at lower concentrations to a lesser degree. It is not clear what these changes represent; it is possible that they indicate the formation of oligomers or other intermediates but further investigation is required to confirm this.

Furthermore there is variation in the data after the exponential phase, and in the steady state, these changes varied between individual repeats unlike the humps preceding the exponential phase. In particular there are small humps visible at 20 and 200  $\mu\text{M}$  after approximately 50 hours. Again, it is not clear what is causing these differences and further investigation is required.

#### **4.3.1.2. Investigating the effect of agitation on medin aggregation**

It has previously been shown that agitation during fibril formation can significantly reduced lag time by expediting nucleation (Xue et al., 2008, Nielsen et al., 2001, Petkova et al., 2005b). In order to investigate whether medin also exhibited the same phenomenon a 20  $\mu\text{M}$  solution was monitored over time. Conditions were identical except for the addition of orbital shaking, for 250 seconds between reads.

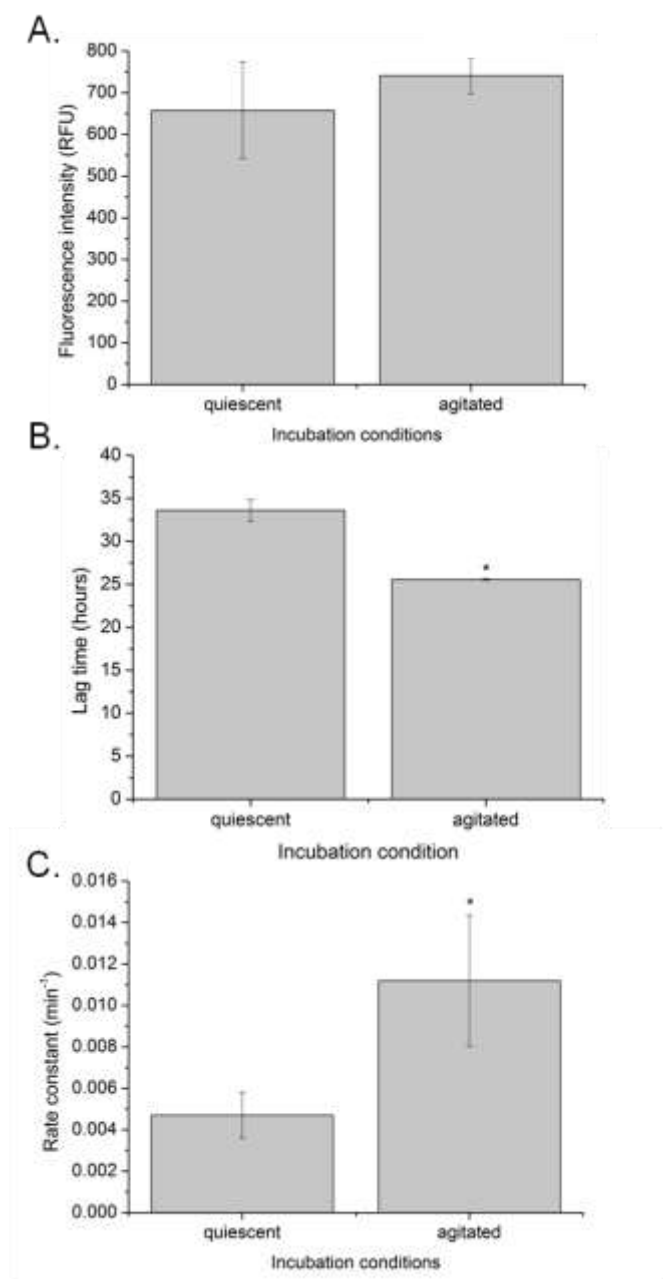




**Figure 4.7** The effect of agitation upon medin aggregation measured by thioflavin T fluorescence. Data were recorded on a Fluoromax spectrometer at 30 °C on a 20  $\mu$ M medin solution in aggregation buffer. The agitated sample was subjected to orbital shaking for 250 s between reads. Data are representative means of three replicates.

This data showed that aggregation kinetics of medin could be changed by the addition of agitation (Figure 4.7). The lag time was reduced from 33 hours to 25 hours suggesting that the agitation did increase the rate of nucleation (Figure 4.8, panel B). Additionally, agitation significantly affected rate but not the fluorescence intensity (Figure 4.8 panels A and C). This conforms to the hypothesis of agitation expediting nucleation.

Moreover, the addition of agitation also greatly exaggerated the dip observed before the exponential increase in fluorescence but the reason behind this remains unknown and demands further scrutiny.

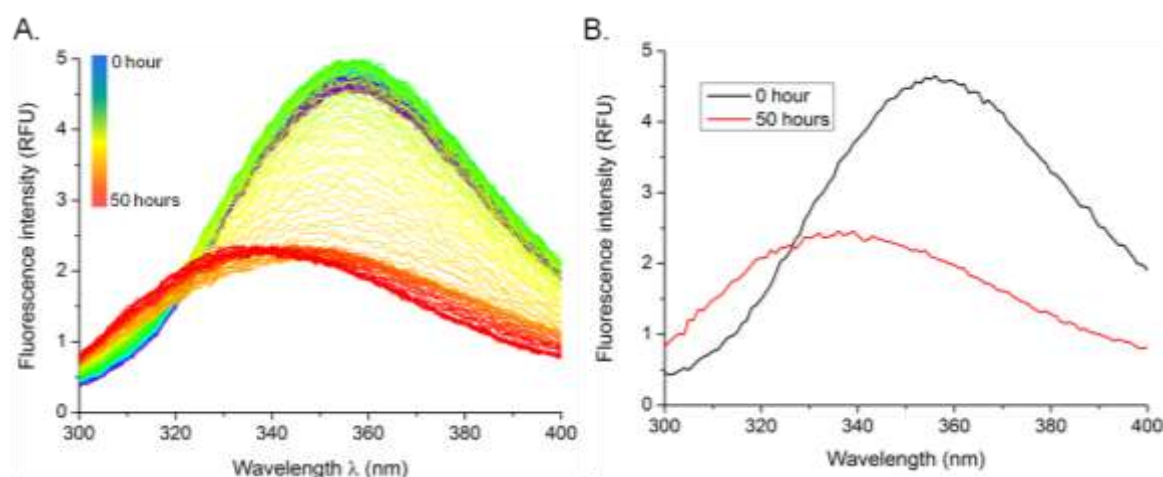


**Figure 4.8** Analysis of the effect of agitation on fluorescence intensity (A), lag time (B) and rate constant (C). Values were extracted from fitting exponential curves to the data and recording the value at the base of the exponential phase, the maximum fluorescence intensity and the calculated rate constants for each condition. Data represent mean and standard error values of 3 replicates.

#### 4.3.1.3. Kinetic analysis by intrinsic tryptophan fluorescence

Intrinsic fluorescence measurements were carried out to investigate the effect of aggregation on the aromatic residues within medin, particularly tryptophan residues.

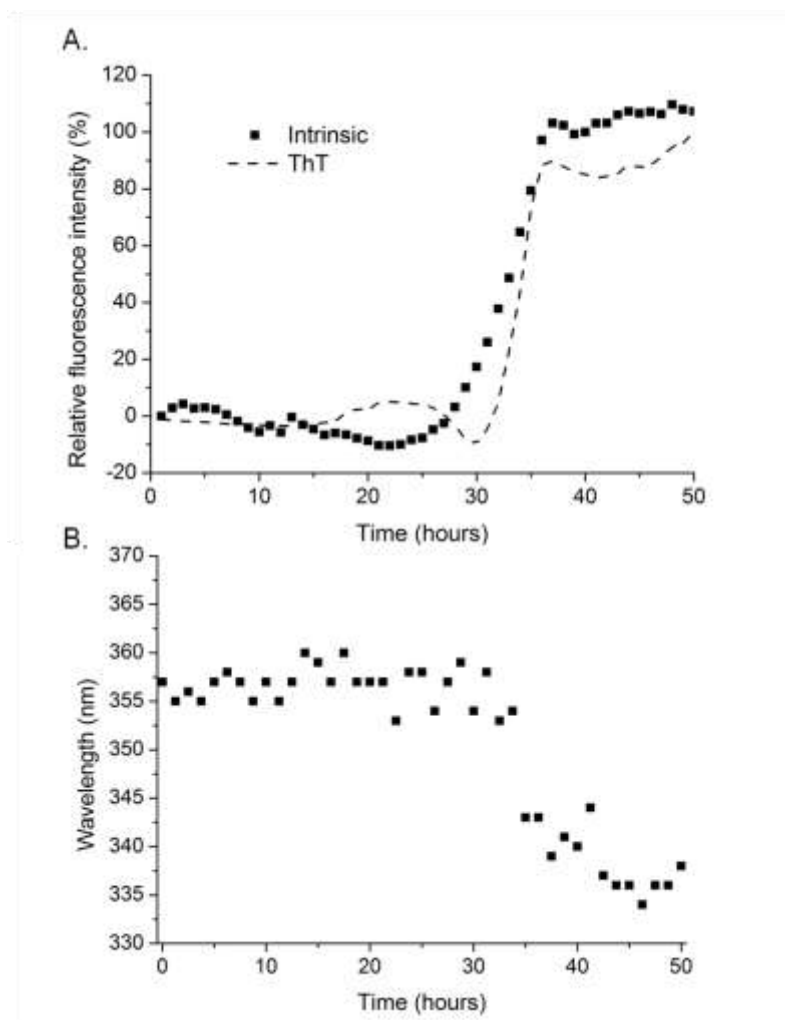
It also acts as an independent method to assess the aggregation kinetics of medin under different experimental conditions. These data show that the tryptophan residues within medin are initially solvent exposed with a maximum fluorescence at 355 nm. However, over time, the maximum shifts to 335 nm which is consistent with the tryptophan residues becoming buried and less solvent exposed (Figure 4.9).



**Figure 4.9** Changes in intrinsic tryptophan fluorescence over time at 30 °C, at a protein concentration of 20  $\mu$ M. Tryptophan residues were excited at 279 nm and emission spectra were recorded from 300-400 nm. Data were collected every 15 minutes for 50 hours, raw data shown in (A) and initial and final spectra shown in (B).

The data were then further analysed to establish the point at which the tryptophan residues change from solvent exposed to a buried environment, by plotting the change in fluorescence intensities at 355 nm and the wavelength at which the maximum fluorescence was observed at each time point. Figure 4.10, panel A shows the percentage decrease in fluorescence at 355 nm compared with the increase in fluorescence intensity observed during the ThT measurements, over time. The intrinsic fluorescence measurements are in good agreement with the ThT findings under quiescent conditions, suggesting that the change observed in the tryptophan residue environments occurs at the same time as the ThT fluorescence increases,

approximately 33 hours after the start of incubation (Figure 4.10 panel A). Panel B shows the change in maximum wavelength over time and illustrates the conversion from 355 to 335 nm that coincides with the increase in fluorescence intensity and ThT fluorescence (Figure 4.10).

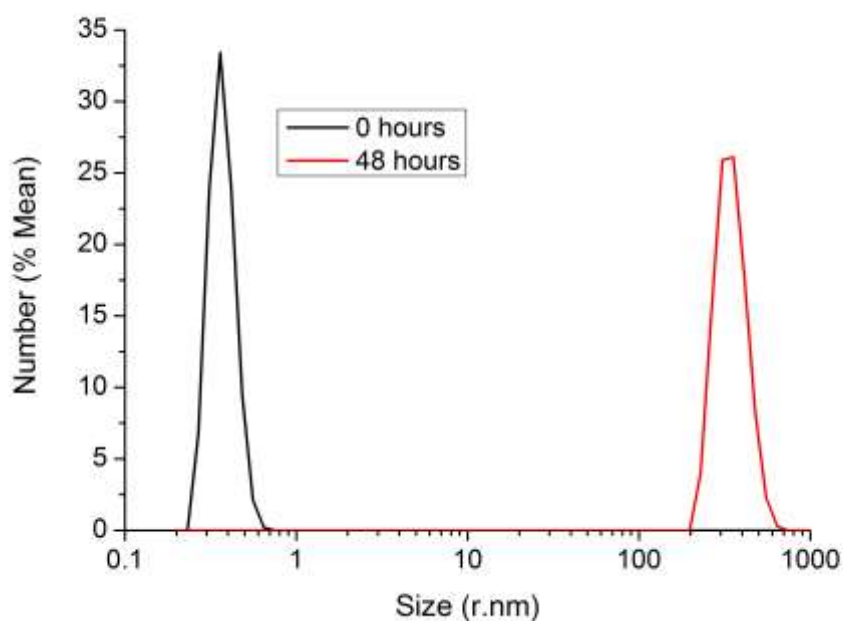


**Figure 4.10 Percentage decrease in tryptophan fluorescence at 355 nm as a function of time (squares) compared to percentage increase in thioflavin T fluorescence over time (dashed line) (A). Change in maximum wavelength as a function of time (B.)**

#### 4.3.1.4. Dynamic light scattering

Dynamic light scattering is a technique that provides information about the size distribution of particles in solution. DLS was used here to monitor changes in particle size pre- and post aggregation. As can be seen in Figure 4.11, at 0 hours the

particle size is consistent with monomeric peptide species in a compact conformation, whereas after 48 hours much larger higher aggregates have formed. These data indicate, like the ThT and IF measurements, that aggregation occurs within 48 hours under quiescent conditions in physiological buffer.



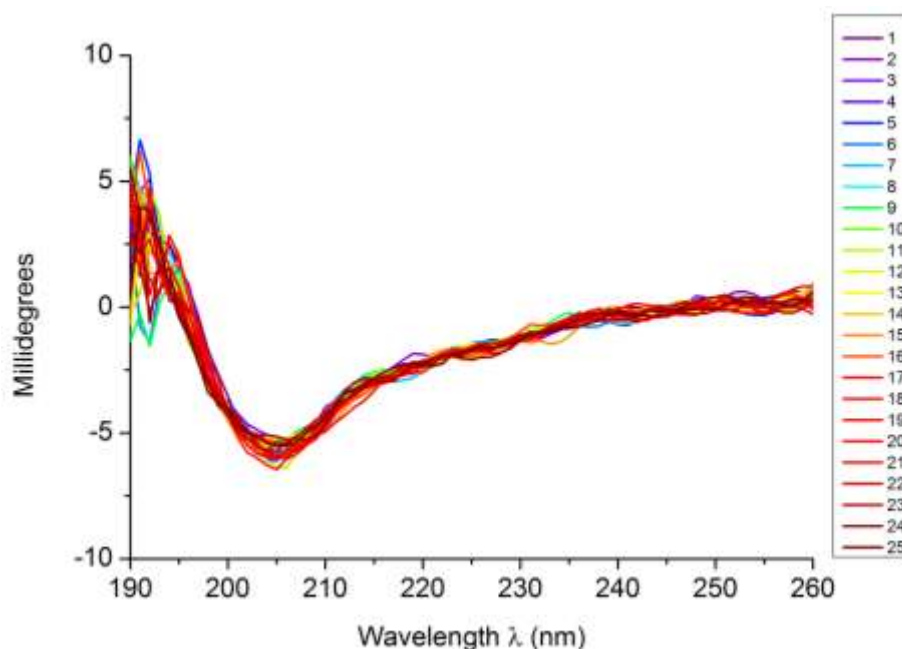
**Figure 4.11 Dynamic light scattering data for medin (20  $\mu$ M in aggregation buffer) at 0 and 48 hours. Data recorded using a Zetasizer nano instrument operating at 25  $^{\circ}$ C; data shows an increase in particle size over time.**

### 4.3.2. Structure and morphology of medin aggregates

#### 4.3.2.1. Circular dichroism

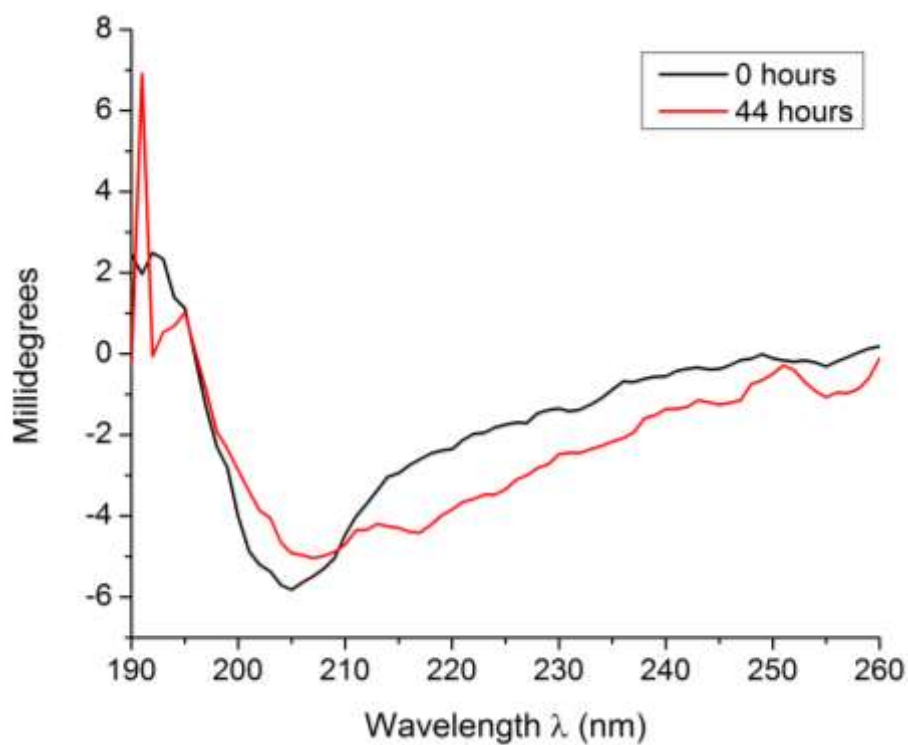
The stability of medin, at the early stages of aggregation, was analysed by CD. A medin sample was repeatedly exposed to single scans of UV radiation over a prolonged period (Figure 4.12). This experiment both acted as a control for the CD analysis shown above and as a means to test the stability of medin. The data show no signs of degradation over 25 scans, suggesting that medin is stable to radiation and that the results shown above in Figure 4.15 are a result of aggregation not

degradation. Thus, medin could be scanned repeatedly over a period of time and any changes in the spectrum could be attributed to time-dependent structural transition not radiation damage.

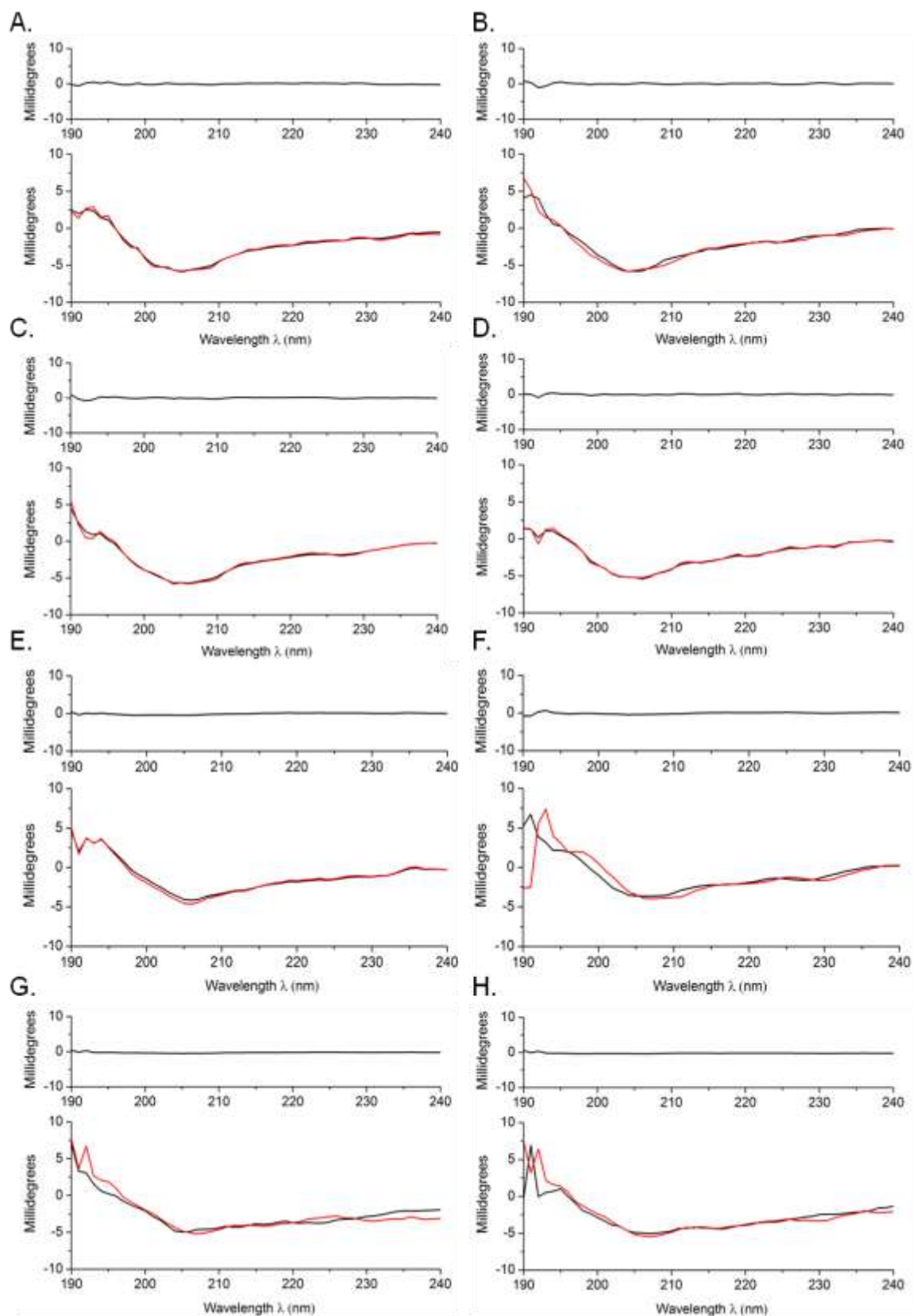


**Figure 4.12 Analysis of medin stability in response to UV radiation. Data were recorded over 25 scans at 30 °C on medin (200 μM, in CD buffer) using far-UV synchrotron radiation circular dichroism at beam line B23 at Diamond light source. Numbers represent number of repeated exposures to UV.**

Far UV - CD analysis of medin was carried out over time to investigate any changes in secondary structure that accompany aggregation. Experiments were carried out using far-UV synchrotron radiation circular dichroism (SRCD). The sample was incubated in the cuvette at 30 °C and scanned every two hours to monitor changes over the aggregation time course identified from ThT measurements. Spectra obtained at 0 hours and 44 hours were compared and show a shift towards more  $\beta$ -sheet (Figure 4.13).



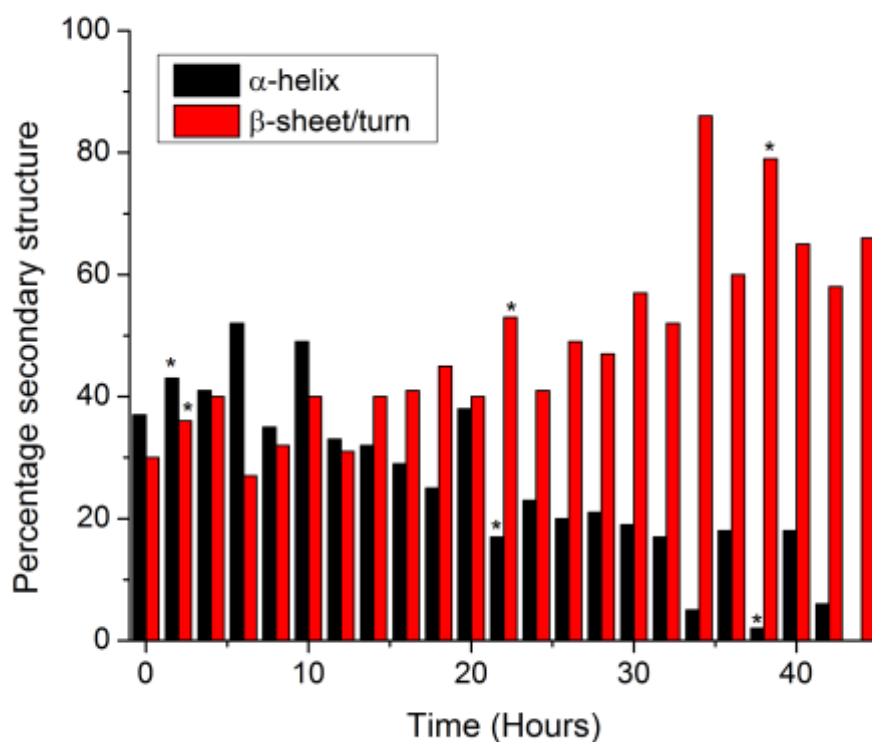
**Figure 4.13** CD spectra for medin comparing start and end points (0 and 44 hours). Data were recorded at 30 °C on medin (200  $\mu$ M, in CD buffer) using far-UV synchrotron radiation circular dichroism at beam line B23 at Diamond light source.



**Figure 4.14** CD spectra for medin (black) with fits (red) and residuals (above) at 0,6,12,18,24,30,36 and 42 hours, A,B,C,D,E,F,G and H respectively.



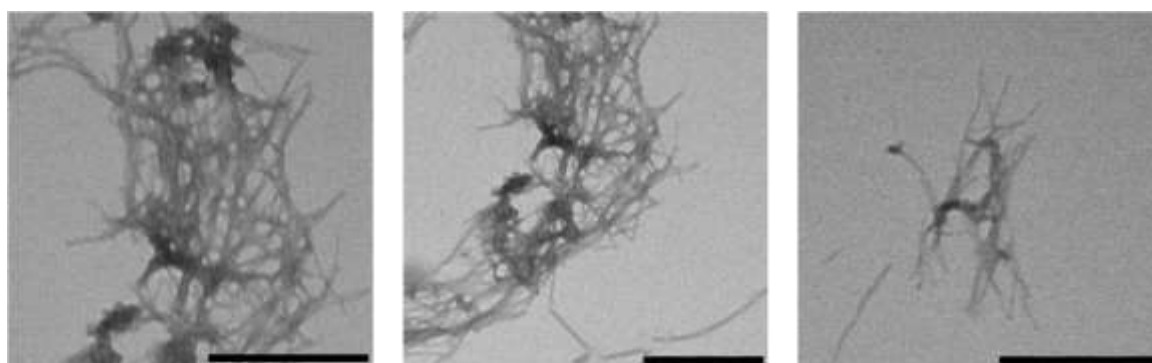
The secondary structure content present at each scan was calculated using CONTIN or CDSSTR secondary structure fitting software and percentage values extracted and plotted over time (Figure 4.15). Throughout the time course there is a reduction in the percentage of  $\alpha$ -helix and an increase in the percentage of  $\beta$ -sheet and  $\beta$ -turn (Figure 4.15) in line with previous studies on amyloid formation (Gross, 2000). At 50 hours, there is approximately 65 %  $\beta$ -sheet/turn present and virtually no  $\alpha$ -helix remaining (Figure 4.15). The amount of random coil present remained relatively constant throughout.



**Figure 4.15** Secondary structure analysis of medin over 44 hour time course. CD data were analysed using, CONTIN or CDSSTR secondary structure fitting software, percentage values recorded, and plotted over time. Asterisk indicate standard deviation values for the fits of over 0.1.

#### 4.3.2.2. Electron microscopy

Negative stain electron microscopy was used to examine the morphology of medin aggregates formed at 20  $\mu\text{M}$  concentration in aggregation buffer (Appendix 1) under quiescent conditions. As can be seen in Figure 4.16, medin forms dense networks of long, straight fibrils under these conditions. They are generally un-branched and approximately 500 nm in length.



**Figure 4.16** Negatively stained electron microscopy images of medin fibrils formed under quiescent conditions, at 30 °C, at a medin concentration of 20  $\mu\text{M}$  in aggregation buffer. Scale bars represent 500 nm.

#### 4.3.3. Conditions for future characterisation

In order to test the modulation of medin aggregation by inhibitory compounds or co-localising factors it was necessary to select a set of conditions for future assays. A peptide concentration of 20  $\mu\text{M}$  was chosen as it gave reproducible ThT data, without using too much material. Also, it was decided to conduct ThT assays without agitation, as these data gave more reproducible results and had a longer lag time in which to observe any potential affects.

#### 4.3.4. Affect of inhibitors on medin aggregation

It is yet unknown what affect medin has on the physiology of the arteries in which it is found or whether it is pathogenic as suspected for other amyloid proteins. It has been suggested that oligomeric medin may cause damage to the elastic structures of

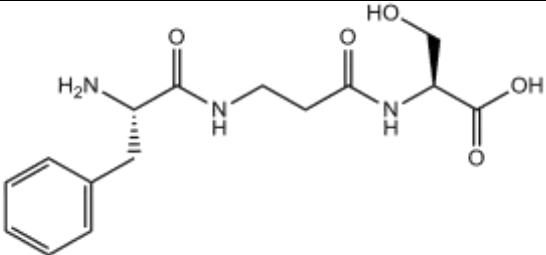
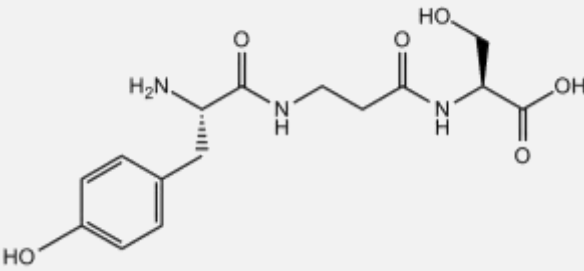
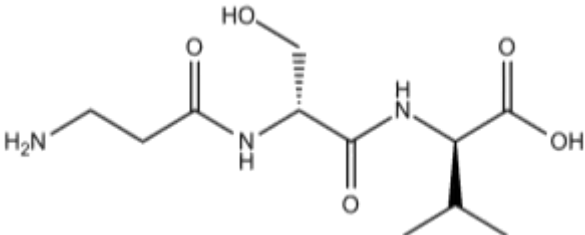
the arteries and render them more prone to rupture (Peng et al., 2007). If this is found to be the case it may be possible to target medin therapeutically. One strategy for inhibition of medin aggregation is to use modified peptides designed to bind to the aggregation prone regions of the target peptide, also known as the self recognition element (SRE). This approach has been successfully used to inhibit other amyloid proteins such as  $\alpha$ -synuclein, amyloid- $\beta$  (A $\beta$ ) and amylin (Madine et al., 2008b, Madine et al., 2009b). This work employs the use of a  $\beta$ -sheet breaking residue,  $\beta$ -alanine or 3-aminopropanoic acid ( $\beta$ A), to disrupt the hydrogen bonding pattern that stabilises the cross- $\beta$  structure within aggregates, and prevent further association of peptides.

Previous work by Madine *et al.* (2009) designed three modified peptides to bind to a previously identified SRE at the C-terminus of medin, with the sequence NH<sub>2</sub>-NFGSVQFVA-COOH in particular the NH<sub>2</sub>-FGSV-COOH motif (Madine and Middleton, 2010, Larsson et al., 2007). The aim of this work was to test two of these inhibitors F( $\beta$ A)S and ( $\beta$ A)SV and a third novel peptide Y( $\beta$ A)S using the plate reader ThT assay developed throughout this work (Table 4.1). Only one of the two modified peptides, designed by Madine and Middleton (2010), with the sequence F( $\beta$ A)S, was capable of preventing the formation of large aggregates. The other, with the sequences ( $\beta$ A)SV, did not prevent the formation of large aggregates.

Previously the peptides were analysed by ThT fluorescence on incubated samples that were frozen at discrete time points and defrosted prior to analysis (Madine and Middleton, 2010). Given the improved capabilities of ThT assays carried on the plate reader, the aim was to investigate the effect the peptides had on the three parameters discussed above, lag time, fluorescence intensity and rate.

It is thought that  $\pi$ -stacking interactions may play a role in the aggregation process (Gazit, 2002). The inclusion of aromatic residues within the peptide sequences may therefore improve binding to the SRE and ameliorate the efficacy of inhibition. The third new peptide was designed based on the success of F( $\beta$ A)S by swapping the phenylalanine residue for a second aromatic residue, tyrosine to give (Y $\beta$ AS).

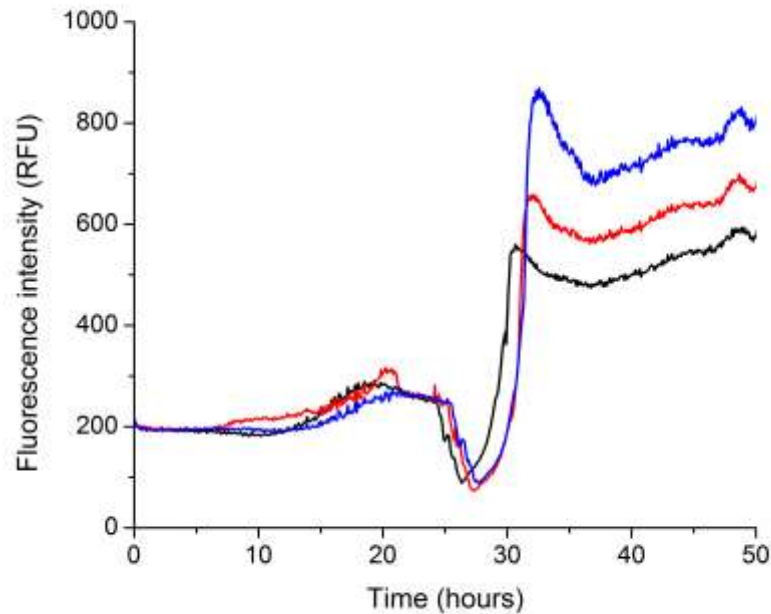
**Table 4.1 Summary of peptide inhibitors designed against the motif, FGS at the C-terminus of medin using  $\beta$ -alanine to extend the backbone and prevent hydrogen bonding.**

Amino acid composition	Structure	Mr (Da)
F( $\beta$ A)S		323.34
Y( $\beta$ A)S		339.34
( $\beta$ A)SV		275.30

#### 4.3.4.1. $\beta$ -sheet breaking inhibitors

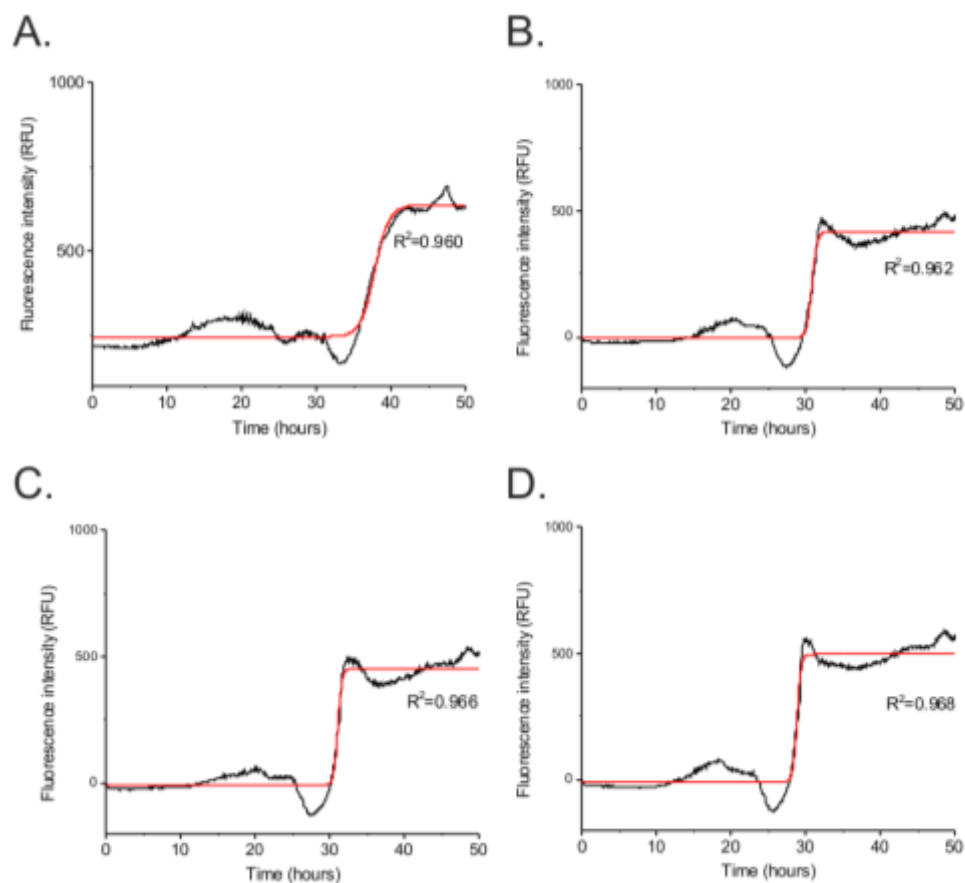
Each of these modified peptides (Table 4.1) was incubated with a 20  $\mu$ M solution of medin and analysed by ThT to see if they have the same effect on the aggregation kinetics of medin as previously described. Moreover, to investigate whether the

exchange of a phenylalanine residue to a tyrosine is as effective as F( $\beta$ A)S. The inhibitors were added two concentrations, both equimolar and in five times excess.



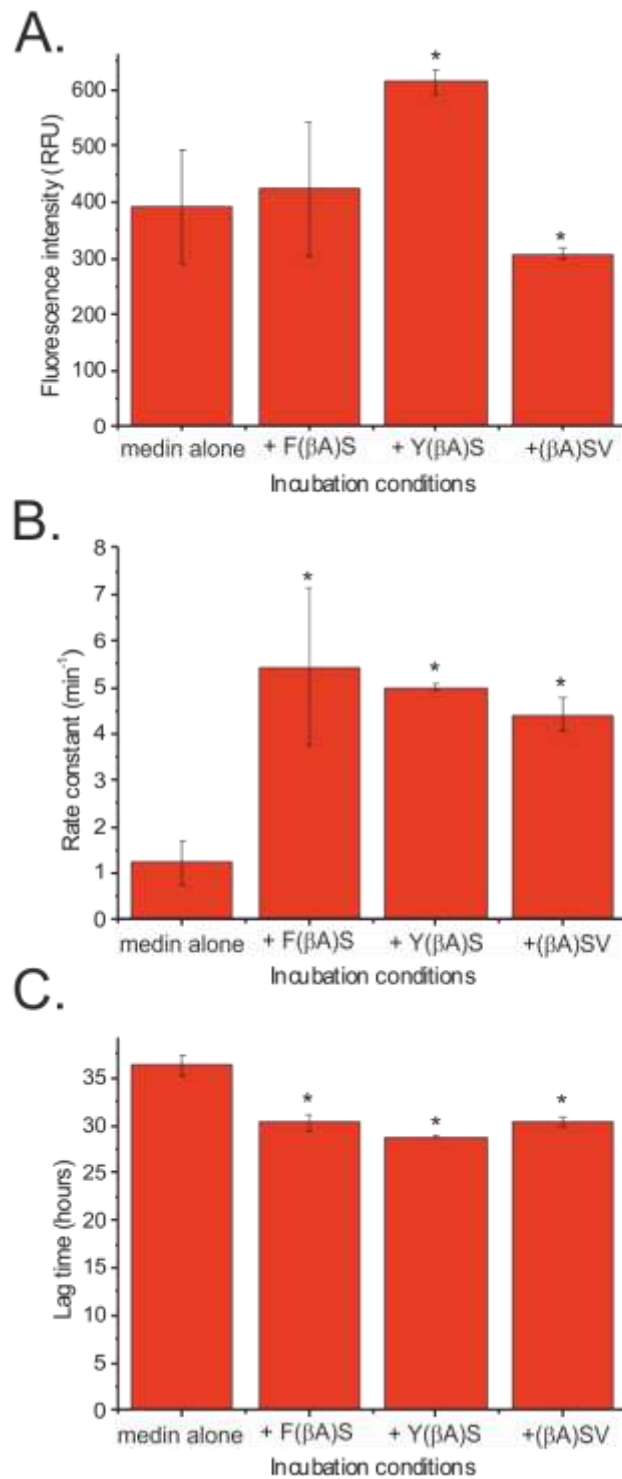
**Figure 4.17 Three replicate ThT data sets for medin +F( $\beta$ A)S (equimolar) incubated at 20  $\mu$ M in aggregation buffer. Data were recorded on a fluoromax spectrometer at 30  $^{\circ}$ C under quiescent conditions.**

Data shown in Figure 4.17 represent three replicate data sets for medin in the presence of equimolar F( $\beta$ A)S. The lag time and rate in particular are quite consistent but there is more variability in the maximum fluorescence intensity.



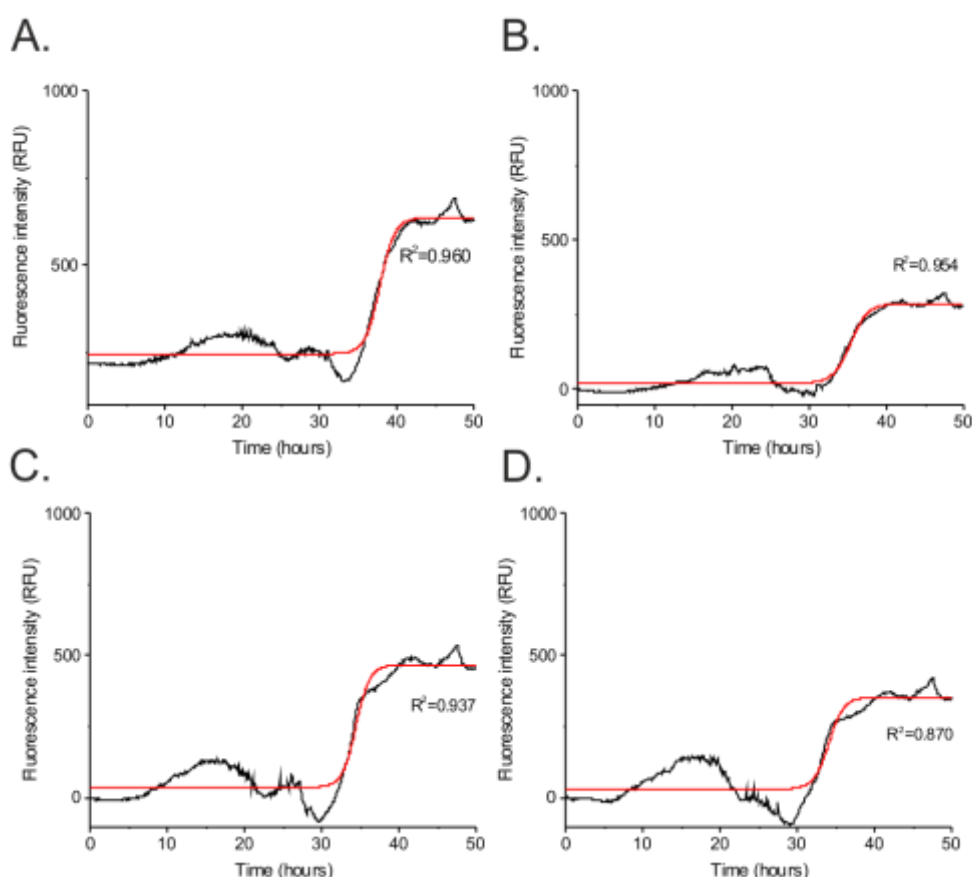
**Figure 4.18** The effect of inhibitor peptides at equimolar concentration upon medin aggregation measured using thioflavin T assay. Medin alone (A), F( $\beta$ A)S (B), ( $\beta$ A)SV (C) and Y( $\beta$ A)S (D). Data were recorded on a fluoromax spectrometer at 30 °C on a 20  $\mu$ M medin solution in aggregation buffer. The red line represents the fit using the following equation:  $y = y_0 + \frac{a}{1 + e^{-\frac{t-T_l}{\tau}}}$ . The adjusted  $R^2$  values indicating the goodness of fit for each curve are reported on the right hand side. Data represent mean data for 3 replicates.

Data initially suggest that all three inhibitors have an effect on the aggregation kinetics of medin by altering the fluorescence intensity, reducing the lag time and increasing the rate (Figure 4.18). Individual analysis of the lag time, fluorescence intensity and rate constants however, show that when compared to medin alone, the increased rate and reduced lag time are significant at the  $p=0.05$  level (Figure 4.19). These inhibitors appear to actually accelerate medin aggregation by reducing the lag time however the presence of inhibitors reduced the rate.



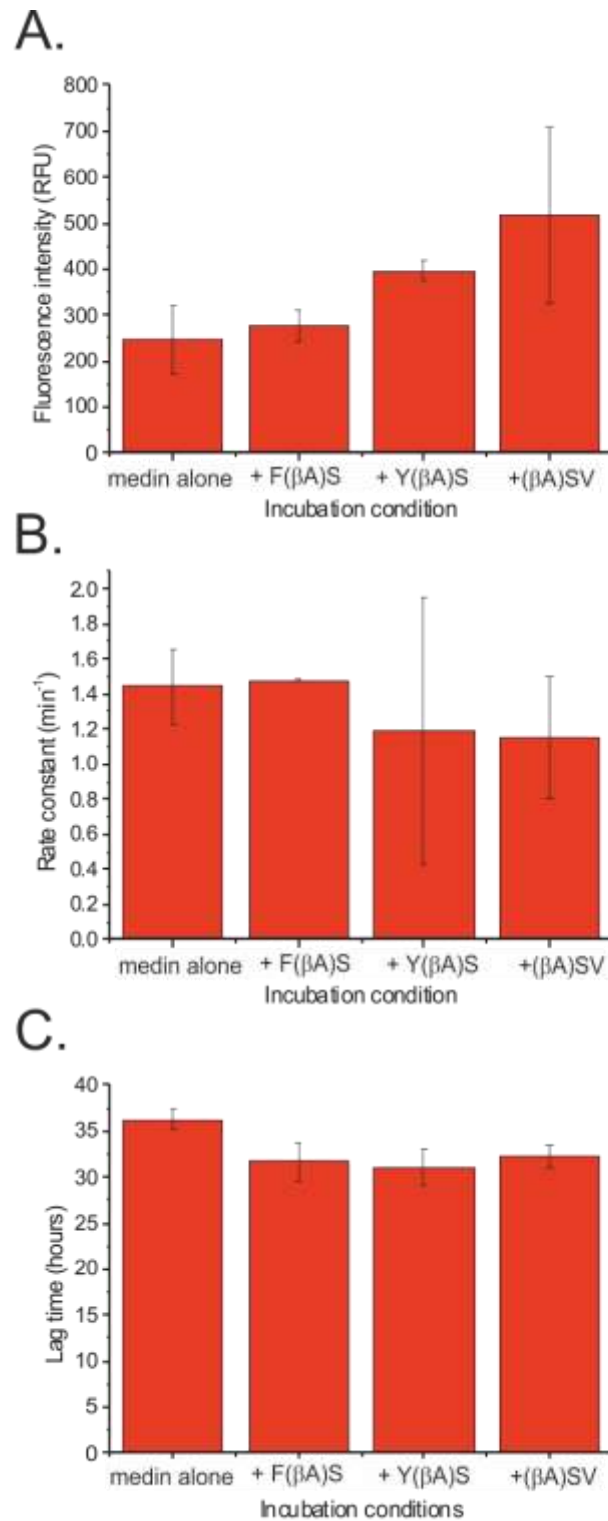
**Figure 4.19** The effect of inhibitor peptides at equimolar concentration upon medin aggregation measured using thioflavin T assay. Data comparing fluorescence intensity (A) lag time (B) and rate of elongation (C). Asterisk indicate statistical significance at the  $p=0.05$  level.

Further experiments were carried out to investigate the effect of incubating medin with a five times excess concentration of the inhibitor peptides. Unlike the ThT results at equimolar concentration, medin incubated with inhibitors at a five times excess concentration showed no significant changes to any of the parameters measured (Figure 4.20 and 4.21).



**Figure 4.20** The effect of inhibitor peptides at five times excess (100  $\mu\text{M}$ ) concentration upon medin aggregation measured using thioflavin T assay. Medin alone (A), F( $\beta\text{A}$ )S (B), ( $\beta\text{A}$ )SV (C) and Y( $\beta\text{A}$ )S (D). Data was recorded on a fluoromax spectrometer at 30  $^{\circ}\text{C}$  under quiescent conditions with a medin concentration of 20  $\mu\text{M}$ . The red line represents the fit using the following equation:  $y = y_0 + \frac{a}{1 + e^{-\frac{t-T_i}{\tau}}}$ . The adjusted  $R^2$  values indicating the goodness of fit for each curve are reported on the right hand side. Data represent mean data for 3 replicates.





**Figure 4.21** The effect of inhibitor peptides at five times excess ( $100 \mu\text{M}$ ) concentration upon medin aggregation measured using thioflavin T assay. Data comparing fluorescence intensity (A) lag time (B) and rate of elongation (C).

#### **4.4. Discussion**

There is currently very little known about the aggregation kinetics, structure and morphology of medin. Studies of the properties of medin are highly justified given its prevalence and the reported links with aortic aneurysm and dissection (Peng et al., 2007). The aim of this work was to increase understanding of the aggregation process of full length recombinant medin under physiologically relevant conditions. This information can then be used to investigate the effect of mutations, co-localising factors and small molecules on the aggregation kinetics of full length medin.

The thioflavin T fluorescence assay (Levine, 1999) is often used to monitor the aggregation of amyloid proteins and it is used here to obtain aggregation profiles for full length medin under a variety of conditions. Thioflavin T analysis of the effect of increasing peptide concentration upon medin aggregation was consistent with previous theories about nucleation-dependant amyloid growth (Cabriolu et al., 2010). Increasing concentration is directly proportional to observed fluorescence intensity and to a lesser degree the rate constant (Figure 4.6). There was also an observed reduction in lag time, however this was less pronounced than expected. This may represent the conditions under which the assay was conducted and further work needs to be carried out to accurately determine the relationship between lag time and concentration. The affect of agitation upon medin aggregation was also investigated. Agitation reduced the lag time by approximately a third (Figure 4.8) is consistent with previous studies on the effect of agitation. These studies on  $\beta$ 2-microglobulin demonstrated that agitation reduced the lag time preceding the exponential phase (Sasahara et al., 2008). It is likely that this reduction is a result of increased fragmentation and thus increased nucleation rate. It may be that agitation is required

in order to observe greater concentration dependence within lag times. It is also interesting to consider the environment in which medin is located *in vivo*. Aortic medial amyloid is most commonly found in close association with the elastic structures of the arteries, a highly dynamic environment. The pressure and forces at this location may have a role in the mechanism of aggregation. The preliminary results shown here agree with previous theories about the role of agitation and given the location of medin deposition, warrant further investigation.

Following kinetic analysis using the classical ThT assay, different aspects of medin aggregation were investigated using a range of biophysical techniques. Moreover, these different techniques can collectively build a reliable picture of medin aggregation.

Intrinsic fluorescence data was collected to investigate the role of tryptophan residues during the aggregation of medin. This data shows a clear switch in tryptophan environment from solvent exposed to buried over the time course of aggregation (Figure 4.9) This switch also coincides with the increase in fluorescence observed in the ThT measurements (Panel A, Figure 4.10). Results are consistent with a change in tryptophan environment associated with aggregation. It may be that the tryptophan residues form stabilising  $\pi$ -stacking interactions although, given the time course, it is unlikely that contributions from these residues are the drivers behind medin aggregation. With such a driver one would expect the switch in environments to precede the increase in ThT fluorescence. The exchange of tryptophan environment from solvent exposed to buried is in agreement with the models proposed by Sawaya *et al.* (2007) of closely packed inter-digitated steric-zippers. It is not possible, from these experiments, to specify the class of steric zipper but they are consistent with a dry steric-zipper interface between  $\beta$ -sheets.

Dynamic light scattering measurements indicated that medin is present in a compact monomeric form at the onset of the time course, and after 48 hours only much larger, presumably aggregated medin remains (Figure 4.11).

Taken together these data strongly suggest that at a 20  $\mu\text{M}$  concentration medin aggregates after approximately 33 hours. Not only do these quite different methods all suggest a lag time of 33 hours, this is also in agreement with data published by Larsson *et al* (2006) on recombinant medin. However, these findings are not consistent with data published by Madine and Middleton (2009) who observed that full length synthetic medin required 8 days to aggregate. It is likely that this is a result of the different conditions used to aggregate medin. It has previously been shown that synthetically produced amyloid proteins aggregate slower than recombinant peptides, probably as a result of residual contaminants from the synthesis (Finder *et al.*, 2010). Furthermore, in the previous study, medin was aggregated at 25  $^{\circ}\text{C}$  whereas these data were collected on medin samples incubated at 30  $^{\circ}\text{C}$ .

Secondary structure analysis of medin by circular dichroism suggested a rearrangement in secondary structure over the observed time course from  $\alpha$ -helix to  $\beta$ -sheet (Figure 4.15) suggestive of amyloid formation (Tomaselli *et al.*, 2006). The transition from  $\alpha$ -helical conformations to  $\beta$ -sheet rich structures is thought to play a role in the infectivity of prion proteins (Pan *et al.*, 1993). It may be that secondary structure transitions trigger the aggregation of amyloid proteins via a similar mechanism. This is supported by the comparison of CD data to the ThT data. The conversion from  $\alpha$ -helix to  $\beta$ -sheet coincides with the increase in ThT fluorescence associated with the dye binding to the cross- $\beta$  structure of the aggregates.

Electron microscopy of the aggregated samples showed the presence on dense fibrillar networks of similar dimensions and morphology to other amyloid forming proteins (Figure 4.16). Although these images clearly show the presence of fibrillar aggregates, it is not possible to detect any specific detail about the structure of the fibrils as has been shown with other amyloid proteins such as A $\beta$  (Paravastu et al., 2008).

Although this work has gone a long way to improve our understanding of how medin aggregates, it should be stressed that these data relate only to the assay conditions used here and further work must be carried out to fully understand process under a variety of conditions.

Initial work on tripeptides, hypothesised to inhibit medin aggregation, was also reported. The efficacy of two previously designed peptides and a third novel inhibitor were tested on medin under the conditions established here. Unfortunately the effects of the two previously designed inhibitors were not accurately mirrored here. This is likely to be consequence of several factors, namely differences in the assay, such as: time course, temperature, peptide concentration and peptide preparation method.

On the contrary at equimolar concentration the tripeptides appear to accelerate medin aggregation (Figure 4.18 and 4.19).

Further work needs to be done to fully characterise the effects of these short peptides upon recombinant medin under more physiological conditions. Furthermore, it also very important to ensure that the peptides do not contribute to the formation of medin aggregates. Cell based toxicity assays must also be included in further work to investigate the toxicity of medin aggregates and the effect of tripeptide addition. Previous cytotoxicity assays conducted on synthetic medin demonstrated that medin incubated for 0, 1 and 3 days had a cytotoxic effect on bovine endothelial cells

however following 7 days incubation medin aggregates were no longer toxic (Madine and Middleton, 2010).

Overall, the results described here have begun to characterise the aggregation process of medin and provide a benchmark for future work detailed in later chapters.

## 5. Toward a structural model for medin fibrils

### 5.1. Introduction

As established in previous chapters, in order to design effective modulators of amyloid formation it is useful to identify the aggregation prone regions of an amyloidogenic peptide. There are several computational methods available to assist in this identification and the role of predicted regions can then be tested experimentally through a variety of molecular biology and biophysical techniques. It is important to identify key interactions such as hydrogen bond networks, salt bridges and aromatic interactions within this region that can be subsequently targeted therapeutically.

#### 5.1.1. Stabilising interactions within amyloid fibrils

There are many different stabilising interactions that are thought to contribute to the formation and stability of amyloid fibrils in general. As discussed in chapter 1, hydrogen bonds formed from the backbone amine groups are the primary stabilising interaction common to all amyloid fibrils. In addition  $\pi$ -stacking of aromatic residues and salt bridges have also been shown to be important (Tu and Raleigh, 2013, Petkova et al., 2002).

Several studies have reported on the role of  $\pi$ -stacking interactions within amyloid structures with different conclusions.  $\pi$ -stacking is described as an attractive non-covalent interaction between aromatic groups, such as those found in phenylalanine, tyrosine and tryptophan. Some indicate that aromatic residues are very important for amyloid formation, (Gazit, 2002, Jones et al., 2003, Tu and Raleigh, 2013) although there is some debate as to whether this is merely due to the hydrophobic nature of these residues rather than  $\pi$ -stacking interactions themselves (Bemporad et al., 2006, Larsson et al., 2007, Lakshmanan et al., 2013). Several studies have discussed the

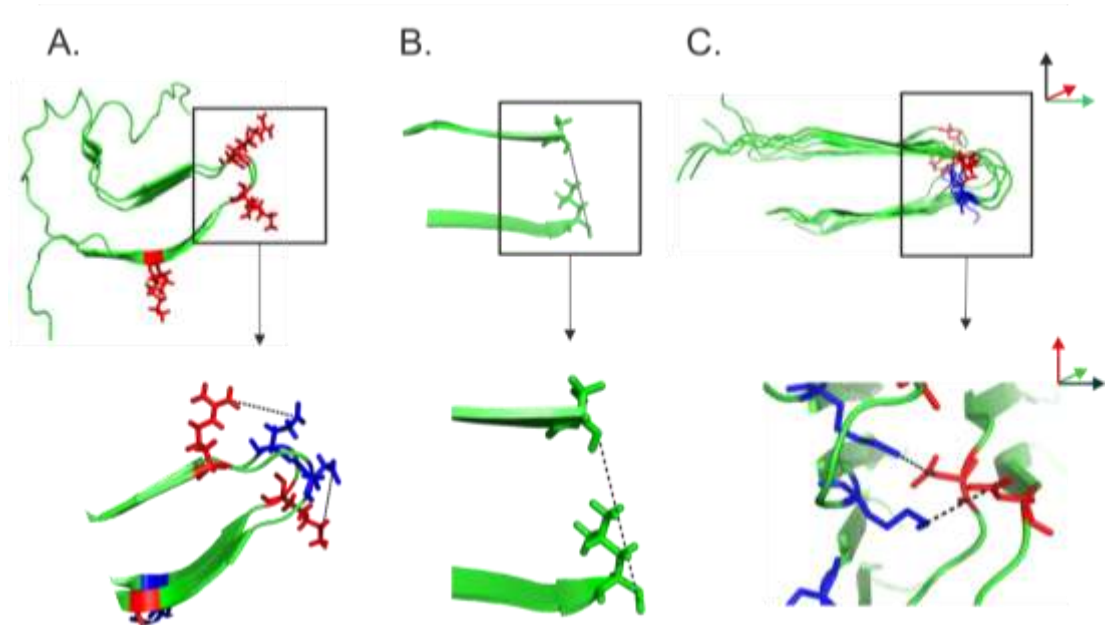
role of  $\pi$ -stacking interactions within medin specifically. Early investigations on the C-terminal region of medin by Gazit *et al.*, (2002) suggested that phenylalanines at positions 43 and 48 are important for amyloid formation, however later studies by Larsson *et al.* (2007) showed that substitution of these residues for alanines did not disrupt fibril formation.

There is also accumulating evidence that salt bridges, between the side-chains of acidic and basic residues, may play an important role in amyloidogenesis or in the stabilisation of the resulting pre-fibrillar and fibrillar aggregates. The salt bridges are non-covalent interactions that contribute to protein stability. They are most commonly formed between the anionic carboxylate ( $\text{RCOO}^-$ ) of either aspartic acid or glutamic acid and the cationic ammonium ( $\text{RNH}_3^+$ ) from lysine or the guanidinium ( $\text{RNHC}(\text{NH}_2)_2^+$ ) of arginine.

Salt bridges have been shown to be present in several amyloid-like assemblies such as  $\text{A}\beta_{1-40}$  (Petkova *et al.*, 2002), Het-S (Wasmer *et al.*, 2008, Gendoo and Harrison, 2011) and prion fibrils (Yau and Sharpe, 2012). These ionic interactions can be inter- and/or intra-molecular and can form between side chain groups or the terminal groups. Examples of the variety of salt bridge interactions within fibrils are shown in Figure 5.1. Panel A, Figure 5.1 shows a single molecule of the HET-s protein from *Podospora anserine* forming a left-handed  $\beta$ -solenoid structure, stabilised by three intra-molecular salt bridges (Wasmer *et al.*, 2008). Fibrils formed from a fragment of the mammalian prion protein (PrP), with face-to-face packing of opposite sheets in an anti-parallel arrangement that is stabilised by a salt bridge between the N- and C-termini (Panel B, Figure 5.1) (Yau and Sharpe, 2012).  $\text{A}\beta_{1-40}$  fibrils, when grown



under agitated conditions, are stabilised by inter-molecular salt bridges (Panel C, Figure 5.1).



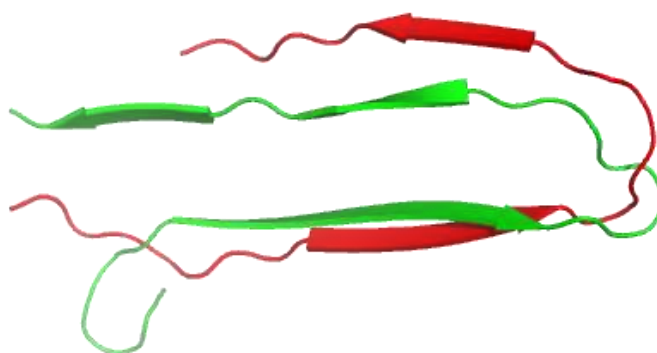
**Figure 5.1** Examples of different salt bridge interactions stabilising fibril structures: intra-molecular salt bridges in Het-S (PDB ID: 2RNM) (A), a salt bridge formed between the termini of a fragment of PrP (B) (Yau and Sharpe, 2012) and inter-molecular salt bridges formed in A $\beta$ <sub>1-40</sub> fibrils (PDB ID: 2LMO) (C). All viewed down the fibril long axis (top) and expanded regions shown below.

It is not known what stabilising interactions are present within medin fibrils. There is some information available about the role of aromatic residues within medin assemblies (Reches and Gazit, 2004, Larsson et al., 2007, Davies et al., 2012, Madine et al., 2009a) but it is not known if salt bridges play a role in stabilising medin fibrils. The latter interactions could occur in principle because medin has two lysine and two arginine residues and one glutamate and three aspartate residues.

### 5.1.2. Similarities between medin and other amyloid proteins

There is a large amount of data on the kinetic, morphological and structural characteristics of many different amyloid proteins. This information can be used to

guide and inform studies on less well characterised amyloid-forming proteins such as medin. There is also a large amount of interest in identifying any common amyloidogenic motifs or ‘self recognition elements’ that could be targeted therapeutically. One strategy to assist in structural studies is to model the unknown fibril structure on a known amyloid structure and then develop and refine the model with experimentally-derived restraints. The best models are generated from templates that have highest sequence identity and are of a similar size. Two well characterised amyloid proteins, A $\beta$  and IAPP, were considered to be possible templates. They are a relatively similar in size to medin, (4.5 kDa and 3.9 kDa respectively) and structural models for the fibrils of both proteins are available. Interestingly however they have a broadly similar fibril structure comprising a single hairpin, despite the low sequence identity (23 %) between them (Figure 5.2).



**Figure 5.2 Aligned single subunits of A $\beta$  (red), extracted from PDB entry 2LMO and amylin (green) fibrils (PDB files from R.Tycko, Bethesda). Alignment was carried out in PyMol V 0.9.**

In order to select the optimal template, sequence alignment of medin with both A $\beta$ <sub>1-40</sub> and IAPP was performed. As expected, there was low overall sequence identity, 13 % between medin and IAPP and 16 % between medin and A $\beta$ <sub>1-40</sub>. Local sequence alignment between the three proteins revealed only 16 % identity between medin and

amylin over a 19 amino acid overlap whereas there was 38 % sequence identity between medin and A $\beta$ <sub>1-40</sub> over a 16 amino acid overlap.

```

Medin  RLDKQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQGARNFGSVQFVA 50
A $\beta$  1-40 --DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGV-----V----- 40
          * : . . . : : : : * : * * . * : * : . * . *

```

**Figure 5.3 Sequence alignment between medin and A $\beta$ . Asterisks indicate conserved residues, colons indicate strongly similar residues and full stops indicate weakly similar residues. Residues highlighted in red indicate amyloidogenic regions predicted by Zyggregator (Tartaglia and Vendruscolo, 2008). Blue box highlights an area of particular similarity.**

A $\beta$ <sub>1-40</sub> was selected as a template for modelling medin fibrils due to the higher sequence identity between medin and A $\beta$ <sub>1-40</sub>. Furthermore, there is a particular area of conservation between residues 25 and 30 of medin with the sequence NH<sub>2</sub>-DLGSSK-COOH. In A $\beta$ , residues D23 and K28 (corresponding to D25 and K30 in medin) lie within the turn and form an inter-molecular salt bridge (Paravastu et al., 2008). The remarkable similarity between the sequences between residues 25-30 led to the hypothesis that medin may also contain a salt bridge at that location that may drive or stabilise aggregation.

#### 5.1.2.1. Salt bridge in A $\beta$ <sub>1-40</sub>: D23 and K28

A combination of limited proteolysis and solution state NMR have shown that residues 21-30 in both A $\beta$ <sub>1-40</sub> and A $\beta$ <sub>1-42</sub> are protease resistant and likely comprise the amyloidogenic region of A $\beta$  (Lazo et al., 2005). Molecular dynamics techniques coupled with solution-state NMR experiments suggest that this fragment is stabilised by a salt bridge between the side chain groups of D23 and K28 (Fawzi et al., 2008). Solid-state NMR measurements on full length A $\beta$ <sub>1-40</sub> fibrils proved that, in some morphologies, fibrils were indeed stabilised by an inter-molecular salt bridge

between residues D23 and K28 (Figure 5.1, panel C), (Petkova et al., 2005b, Petkova et al., 2002). Further studies have confirmed that the region encompassing the D23/K28 salt bridge is critical to the kinetic and structural elements of A $\beta$  aggregation. Kinetic studies on A $\beta$ <sub>1-40</sub> with a  $\beta$ -lactam bridge between D23 and K28 indicated that cross-linking at the turn region i.e. simulating the salt-bridge interaction, eliminated the lag phase and showed a reduced critical concentration for fibril formation, measured by thioflavin T fluorescence. Furthermore these cross-linked peptides were capable of acting as seeds for wild type A $\beta$ <sub>1-40</sub>, suggesting structural similarity (Sciarretta et al., 2005). In addition to cross-linking, mutation experiments promoted turn formation within  $\beta$ -sheets through systematic replacement of amino-acid pairs within the turn region with <sub>D</sub>-PG, an effective turn nucleator. These mutations led to significantly accelerated fibrillogenesis when measured by thioflavin T fluorescence (Doran et al., 2012a). Together this data suggests that turn formation is a rate limiting step in A $\beta$  fibril formation.

Various molecular dynamics experiments have also emphasised the importance of D23 and K28 to the geometry and mechanical properties of A $\beta$  fibrils (Cruz et al., 2012, Berhanu and Hansmann, 2012b).

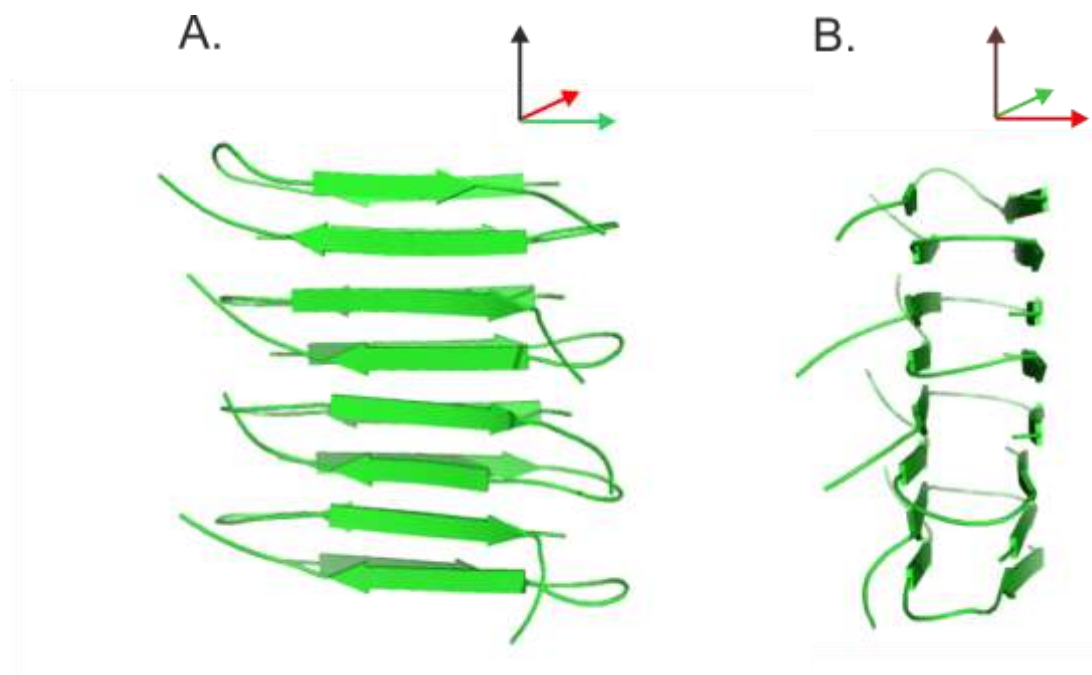
### **5.1.3. A $\beta$ Iowa mutant**

In addition to the structural importance of D23, it has been shown that a mis-sense mutation of this residue to an asparagine in amyloid precursor protein gives rise to a severe cerebral amyloid angiopathy (CAA), widespread neurofibrillary tangles and unusually extensive deposition of A $\beta$  plaques. This mutation was first identified in a family in Iowa suffering from familial Alzheimer's disease (Grabowski et al., 2001). Several studies have begun to characterise the kinetic and pathological properties of

D23N. Incubation of D23N with human smooth muscle cells showed increased fibril formation and pathogenicity when compared with wild type A $\beta$  (Van Nostrand et al., 2001). Transgenic mice carrying the D23N mutation displayed enhanced accumulation of A $\beta$  fibrils in the vasculature of the brain, which strongly stained with thioflavin S. There was further evidence of micro-haemorrhage close to the plaques. Transport studies also indicated that D23N was more readily retained in the brain than WT A $\beta$  (Davis et al., 2004).

Structural characterisation of D23N, by EM, XRFD and solid state NMR indicate a heterogeneous mixture of in-register parallel  $\beta$ -sheet structures, similar to WT A $\beta$ <sub>1-40</sub> and, surprisingly, anti-parallel  $\beta$ -sheet structures. To date, all previous amyloid structural models, where the peptide contains more than one hydrophobic region, are composed of in-register parallel  $\beta$ -sheets (Tycko, 2011). Anti-parallel arrangements have been observed, but were restricted to fragments or short peptides (Tycko et al., 2009). These two structures have different kinetic properties; the anti-parallel structures have a higher nucleation rate but lower elongation rate than their parallel counterparts.

These anti-parallel fibrils are thermodynamically meta-stable and over time convert to parallel structures. Both the parallel and anti-parallel arrangements exhibited toxic effects upon incubation with rat embryonic hippocampal neurons (Qiang et al., 2012). It is thought the anti-parallel structures represent a toxic intermediate in the aggregation pathway which may contribute to the observed enhanced pathology. Seeded growth allowed the isolation and characterisation of the anti-parallel fibril structure (Qiang et al., 2012).



**Figure 5.4 Model of anti-parallel D23N fibrils generated from solid-state NMR restraints, viewed perpendicular to the fibril long axis from the side (A.) and end (B.), highlighting the anti-parallel arrangement of the  $\beta$ -sheets. Loops were smoothed in PyMol to improve clarity.**

Together the clinical, cellular and biophysical data suggest that the aspartic acid at position 23 is critical to the kinetic and structural properties of A $\beta$ , which in turn, manifests in the differences in the clinical pathologies observed.

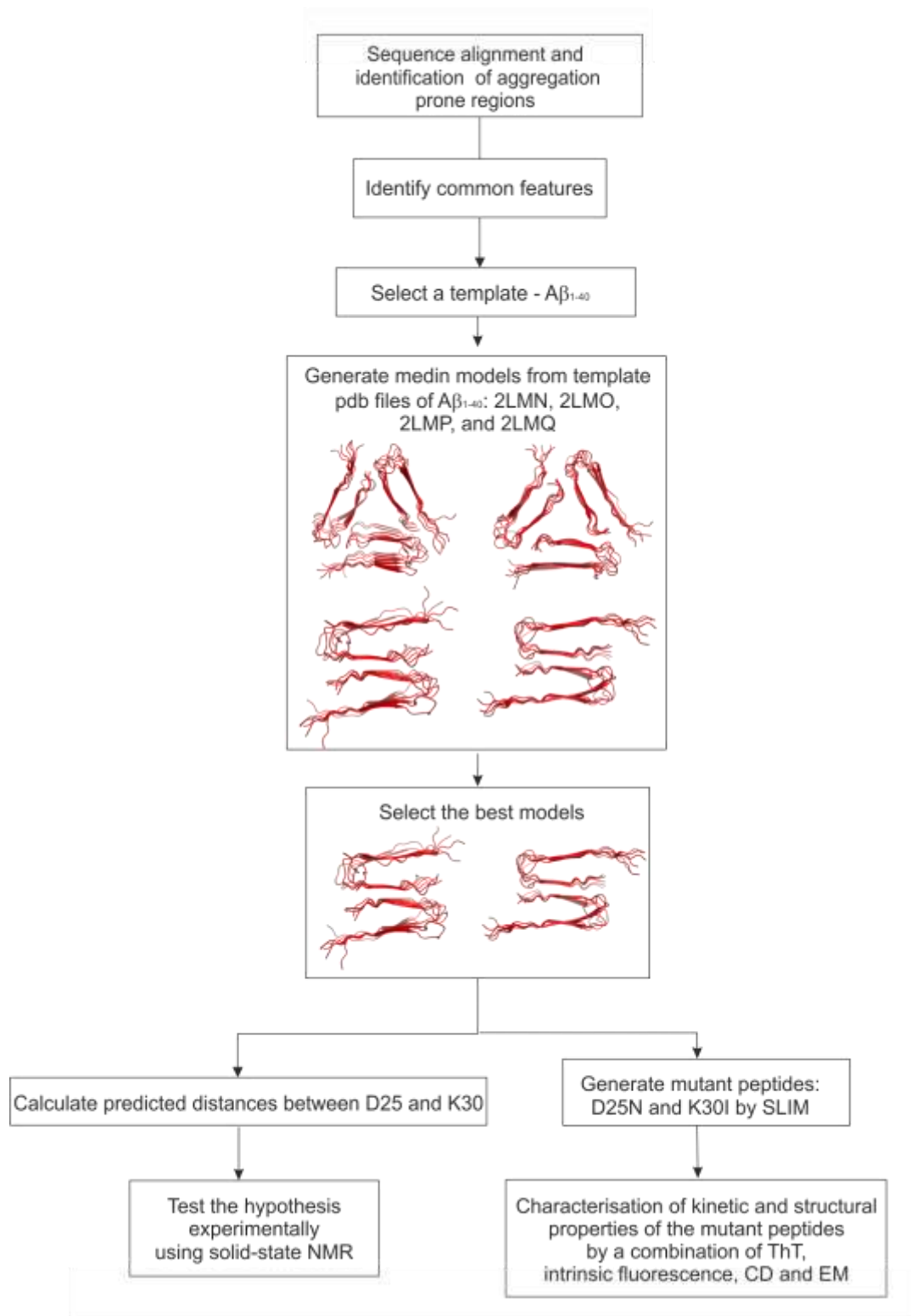
#### **5.1.4. Objectives**

The objective of this chapter is to conduct detailed structural analysis of medin drawing on the sequential similarity with A $\beta$  for which fibril structural models are available. In particular, the focus is on a possible stabilising salt bridge. This hypothesis will be tested by a combination of computational modelling supported by experimental measurements using solid state NMR and site directed mutagenesis.

## **5.2. Materials and methods**

### **5.2.1. Overall strategy**

Computational methods have been developed to highlight key amyloidogenic regions and interactions that can be tested experimentally (Buchete et al., 2005, Nelson and Eisenberg, 2006, Lee et al., 2007). The methods described here were used both to identify similarities between amyloid proteins and find stabilising interactions. These could then be targeted and investigated by both biophysical and structural methods and potentially targeted by inhibitors. Figure 5.5 outlines the overall strategy and the two parallel strands of investigation. Initially, sequential and amyloidogenic similarities between medin and A $\beta$  were identified using bioinformatics tools. This highlighted the region between residues 25 and 30 of medin with the possibility of a stabilising salt bridge between both residues. The kinetic and structural impact of substituting these residues was also investigated using a range of biophysical techniques.



**Figure 5.5 Schematic illustration of the overall experimental strategy detailing the parallel strands of investigation, both structural modelling, direct distance measurement using solid-state NMR and mutant characterisation.**



### **5.2.1.1. Sequence analysis**

As discussed in previous chapters, medin shares sequence similarity with amyloidogenic motifs observed in other amyloid proteins such as IAPP and A $\beta$  (section 2.1). Global and local sequence alignment of human medin against human A $\beta$ <sub>1-40</sub> and IAPP was performed using LAlign server (Huang and Miller, 1991) LAlign was used as it provides both a global sequence identity output and highlights areas of local similarity, which is ideal for identifying possible amyloidogenic motifs that may be missed with global alignments. The alignment was carried out using the BLOSUM50 matrix with a gap open penalty of -12 and a gap extension penalty of -2 with a threshold of 10. Residues 1-9 of A $\beta$  are considered to be unstructured and were omitted from the modelling calculations (Petkova et al., 2002). Similarly, the additional 10 residues at the C-terminus of medin were removed due to the lack of template structure.

### **5.2.1.2. Modelling**

Given the sequence similarities between medin and A $\beta$ , structural models of medin protofilaments were generated using four A $\beta$  structural models deposited within the protein databank (PDB), (PDB IDs: 2LMN, 2LMO, 2LMP and 2LMQ) derived from solid-state NMR measurements. Both single and multi-chain modelling were carried out using Modeller V9.11 on truncated (residues 10-40) medin (Eswar et al., 2007). Input scripts are shown in appendix 4.

### **5.2.1.3. Model validation and distance measurements**

An ensemble of 20 models was generated for each template and the best model of each selected, based on the lowest molecular pdf score (molpdf), which is a standard MODELLER scoring function and is simply a sum of all the restraints (Eswar et al.,

2007). WHAT IF scores were then calculated for each model and compared with those for the template models (Vriend, 1990). WHAT IF calculates an overall RMSD score to indicate how all residues within the peptide are packed. Chimera software is an integrated programme for the analysis and visualisation of molecular structures, here it was used to identify steric clashes within the models (Yang et al., 2012). Visualisation of models and distances between residues were calculated in PyMOL molecular graphics system version 1.5.0.4 (Schrödinger LLC.).

## 5.2.2. Solid state NMR studies

### 5.2.2.1. Preparation of medin for NMR studies.

Uniformly  $^{13}\text{C}$  and  $^{15}\text{N}$  isotopically labelled recombinant medin was incubated at a concentration of 200  $\mu\text{M}$  in aggregation buffer (Appendix 1) for 3 weeks with agitation at 25  $^{\circ}\text{C}$ . The resultant fibrils were harvested by centrifugation at 21,000 g for 1 hour to generate a tightly packed pellet before being transferred to a zirconium 3.2 mm rotor with a Kel-F cap (Bruker, U.K.).

Synthetically produced medin was purchased from Protein peptide research (Cambridge, U.K.) with selective isotope labels. The sequence was designed to include: three uniformly  $^{13}\text{C}$  and  $^{15}\text{N}$  labelled amino acids A13, D25 and K30 (Figure 5.6). A13 was labelled for experiments intended to determine the strand alignment within the fibrils, but time was not available to perform the NMR measurements.



**Figure 5.6 Labelling scheme for full length wild type medin. Asterisks indicate the presence of isotope labels. Residues highlighted in red are uniformly labelled  $^{13}\text{C}$  and  $^{15}\text{N}$ .**

Lyophilised synthetic medin was initially dissolved in hexafluoroisopropanol (HFiP), bath sonicated for 10 minutes and then the HFiP removed under a stream of nitrogen. The samples were then placed on the high vacuum overnight to remove the last remaining traces of HFiP. This process was repeated three times prior to resuspension in 10 % dimethylsulphoxide (DMSO), 90 % H<sub>2</sub>O for aggregation. The synthetic peptide was highly insoluble and required dissolution in 100 % DMSO and subsequent dilution with H<sub>2</sub>O to reduce the percentage of DMSO. Medin was incubated at a concentration of 200 μM with agitation at room temperature for up to 21 days. The peptide suspension was then centrifuged at 21,000 g for 1 hour to generate a tightly packed pellet before being transferred to a zirconium 4 mm rotor with a Kel-F cap (Bruker, U.K.) in preparation for NMR studies.

#### **5.2.2.2. One-dimensional spectra of isotopically labelled recombinant medin**

One-dimensional <sup>13</sup>C and <sup>15</sup>N cross-polarization magic-angle spinning spectra were obtained on a Bruker 850 MHz wide-bore solid-state NMR spectrometer at the national solid-state NMR facility (Warwick Univeristy). The experimental parameters were: 14 kHz sample rotation, 1-ms Hartmann-Hahn contact time, 100 kHz proton decoupling and a 1.5-s recycle delay.

#### **5.2.2.3. Two-dimensional DARR spectra of isotopically labelled recombinant medin.**

DARR NMR experiments were performed using a Bruker 850 wide bore spectrometer operating at a static magnetic field of 20 T and using a Bruker 3.2 mm triple resonance probe head in double resonance mode. Samples were maintained at -23°C with a sample rotation frequency of 14 kHz ± 1 Hz. Experiments utilized a <sup>1</sup>H 90° excitation pulse length of 2.5 μs, Hartmann-Hahn cross polarization over a 1-ms

contact time, 3  $\mu$ s  $^{13}\text{C}$   $90^\circ$  pulses, SPINAL proton decoupling at 100 kHz and a 1.5 s recycle delay. The proton field was reduced to 14 kHz during the DARR mixing time of 10 ms. Phase-sensitive spectra were obtained using time-proportional phase incrementation with 420 points in the indirect dimension. The spectrum at each  $t_1$  increment was the result of accumulating between 300 and 1024 transients.

#### **5.2.2.4. Simulated spectra**

Simulated DARR spectra were used to assess the secondary structure content of the experimental DARR spectra. This process comprised two stages, firstly  $^{13}\text{C}$  resonance frequencies were collated and subsequently these chemical shift values were used to simulate DARR spectra using a custom C program written by Prof. D. A. Middleton.

Chemical shift values were calculated using the online server SHIFTX2 (Han et al., 2011). This server uses pre-calculated empirically derived chemical shift hypersurfaces to predict chemical shift values for user defined molecular model. These hypersurfaces encompass information about dihedral angles, side-chain orientations, secondary structure and nearest neighbour effect and are generated from known protein structures. Chemical shift values for two median models were generated using SHIFTX2, the median fibril model generated in section 5.3.1.2, and a linear median model.

A C program developed by Prof. D. A. Middleton used the derived chemical shift values to generate simulated DARR spectra. Time domain signals were simulated as 512 x 512 matrices and complex Fourier transformation was performed in two dimensions to obtain the frequency domain spectra. For the simulated DARR spectra, only short-range couplings between directly bonded spins were considered, and long-

range couplings were neglected to simplify the simulated spectra. Each FID in the  $t_2$  dimension was modulated therefore by no more than three frequencies in  $t_1$  (i.e., depending on the number of bonded carbon atoms), which restricted the number of cross-peaks.

The simulated spectra for the two models were also compared to simulated spectra for purely random coil and  $\alpha$ -helical conformations. These spectra were generated using  $^{13}\text{C}$  resonance frequencies taken directly from standard chemical shift values for  $\alpha$ -helix and random coil (Ulrich et al., 2008).

#### **5.2.2.5. Heteronuclear distance measurements using frequency selective**

##### **REDOR**

Frequency selective REDOR (FSR) experiments were performed on hydrated fibrils using a Bruker Avance 400 spectrometer operating at a magnetic field of 9.3 Tesla. Samples were packed into a 4-mm zirconium rotor and rotated at the magic angle while maintaining the spinning rate automatically to within  $\pm 1$  Hz. All experiments utilized cross-polarization with an initial 4.0- $\mu\text{s}$   $^1\text{H}$   $90^\circ$  excitation pulse, 1-ms Hartmann–Hahn contact time at a matched  $^1\text{H}$  field of 65 kHz, TPPM proton decoupling (Bennett et al., 1995) at a field of 85 kHz following cross-polarization and a 1-s recycle delay. Frequency-selective  $^{13}\text{C}$ -observed REDOR measurements with  $^{15}\text{N}$  dephasing were conducted using the pulse sequence described by (Jaroniec et al., 2001). The MAS frequency was 7400 Hz. To observe  $^{13}\text{C}$  dephasing a train of 82 or 122 non-selective 4  $\mu\text{s}$   $\pi$  pulses was applied at the  $^{15}\text{N}$  frequency every half rotor cycle, corresponding to total dephasing times of 5.5 ms and 8.2 ms, respectively. Frequency-selective 883 ms Gaussian pulses (defined by 1000 points and truncated at 1% of the maximum amplitude), were applied in the centre of the

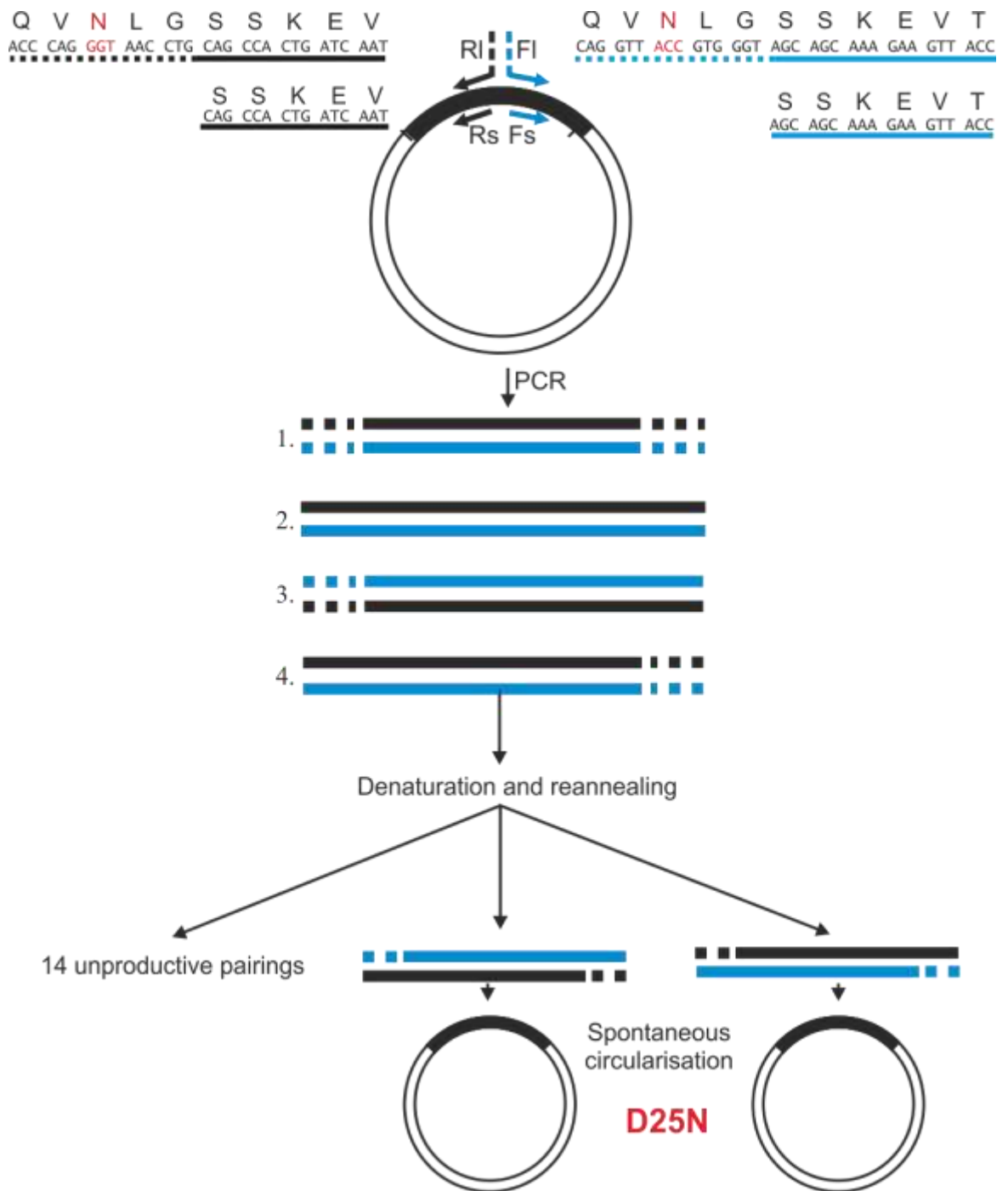
dephasing period. The frequencies of the Gaussian pulses were centred on the amide  $^{13}\text{C}$  and lysine  $\zeta$ - $^{15}\text{N}$  resonance frequencies in order to selectively re-couple the nuclear spins resonating at these frequencies. A second, control measurement was also performed at each dephasing time, omitting all  $^{15}\text{N}$  irradiation so as to observe the loss of  $^{13}\text{C}$  coherence resulting from processes not related to  $^{13}\text{C}$ - $^{15}\text{N}$  re-coupling. Spectra were obtained by averaging 8 blocks, each accumulated from 20,480 transients and alternating experiments with  $^{15}\text{N}$  irradiation and experiments without  $^{15}\text{N}$  irradiation, so as to adjust for any drift in probe tuning during the long acquisition time. The extent of dephasing resulting from  $^{13}\text{C}$ - $^{15}\text{N}$  dipolar coupling ( $S_D$ ) was quantified as the peak intensity observed with  $^{15}\text{N}$  irradiation divided by the intensity observed in the absence of  $^{15}\text{N}$  irradiation.  $S_D$  was translated into a  $^{13}\text{C}$ - $^{15}\text{N}$  distance and using a C program written specifically for that purpose (by Prof. D A Middleton).

### **5.2.3. Mutagenesis**

#### **5.2.3.1. Generation of D25N and K30I mutants by site-directed ligase independent mutagenesis (SLIM).**

In order to investigate the effect of certain residues on the structure and aggregation properties of medin, two mutant proteins were generated using the SLIM method (Chiu et al., 2004). Two residues within wild type medin were selected for mutation; the aspartic acid at position 25 was substituted for an asparagine (D25N) therefore neutralising the charge. The lysine at position 30 was substituted for an isoleucine (K30I) again neutralising the charge and introducing an additional hydrophobic residue (Figure 5.7).





**Figure 5.8 Schematic of SLIM protocol, adapted from (Chiu et al., 2004). Initially an inverse PCR reaction performed with modified primers, generates 4 products. Denaturation and re-annealing of these products results in 16 possible pairings. Two of these pairings are capable of spontaneous circularisation via the tail regions resulting in the mutated plasmid. This example details the primers used to generate D25N medin mutant.**



#### **5.2.3.2. Expression of mutants**

Mutant proteins were expressed under the same conditions as wild type medin detailed in section 3.2.1.8. In brief, plasmids were transformed into BL21 cells, grown overnight and subsequently used to inoculate 2 L of LB. Cells were induced at an O.D. <sub>600</sub> between 0.7 and 1, with IPTG at a final concentration of 1 mM. Cells were harvested by centrifugation and then lysed by sonication. Proteins were purified using a 5 ml HisTrap™ column (G.E.Healthcare), desalted, and then cleaved using SUMO protease, produced in-house. The resultant cleavage mixture was then passed back down the 5 ml HisTrap™ column to remove any un-cleaved protein and the protease.

#### **5.2.4. Characterisation of mutants**

In order to investigate the affect of the substitutions upon medin aggregation, a variety of biophysical techniques were used and the data compared to that collected for wild type medin.

##### **5.2.4.1. Kinetic analysis by thioflavin T assay and intrinsic fluorescence.**

Both assays were carried out on the mutants under the same conditions as the wild type, described in section 4.2.1. In brief, mutant peptides were incubated in aggregation buffer at a concentration of 20 μM in the presence of ThT (20 mM) (Appendix 1). ThT assays were carried out on a Flexstation 3 microplate reader (Molecular Devices Ltd.) using 96-well black-walled, clear bottomed microplates (Nunc) with data recorded using bottom read mode.

#### **5.2.4.2. Structure and morphology analysis by CD and electron microscopy.**

Far-UV CD was used to identify any changes in secondary structure that arose as a result of the substitutions. TEM was carried out to see if there were any visible differences in morphology following 48 hour incubation. Both were done under the same conditions as wild type medin described in section 4.2.4. In brief, the mutant peptides were incubated at a concentration of 200  $\mu\text{M}$ , at 30  $^{\circ}\text{C}$ , in the cuvette (121.000-QS Hellma, Southend-on-Sea, U.K.) for the duration of the time course. Experiments were carried out at beam line B23 at Diamond light source (Oxford, U.K.). Single scans were recorded every two hours between 190 and 240 nm, using a slit width of 0.5 mm, 0.5 cm path length and a scan rate of 1 nm/sec, over 44 hours. Analysis was carried out using Olis<sup>®</sup> GlobalWorks software. Data at each time point was subject to two fitting methods, CONTILL and CDSSTR, using either basis sets 8 or 11 (Van Stokkum, 1990, Sreerama, 2000, Manavalan, 1987). The best fit, as determined by the normalised spectral fit standard deviation, was selected and the percentage  $\alpha$ -helix,  $\beta$ -sheet, turn and random coil content recorded.

### **5.3. Results**

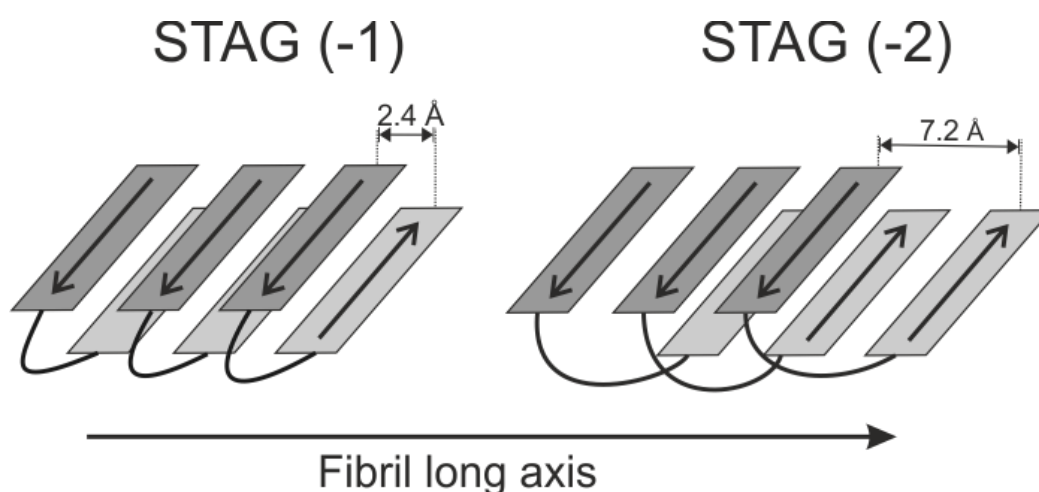
#### **5.3.1. Models**

Molecular modelling based on A $\beta$  fibril models was carried out to visualise any possible turn and to guide elements of the experimental design such as isotope labelling schemes and solid-state NMR pulse sequences.

##### **5.3.1.1. A $\beta$ <sub>1-40</sub> template**

Four A $\beta$  fibril models were used as templates representing two different polymorphs of A $\beta$ <sub>1-40</sub>, with 2- and 3-fold symmetry. Previous structural studies used synthetic isotopically labelled A $\beta$ <sub>1-40</sub>, aggregated under two different conditions, with and

without agitation. Subsequently, multiple rounds of seeding were used to generate two different homogeneous morphologies (Paravastu et al., 2008, Petkova et al., 2005b). The aggregates were then studied, primarily by solid state NMR, and models generated from the resultant restraints (Paravastu et al., 2008, Petkova et al., 2002). Four models have been deposited in the PDB with the following Ids: 2LMO, 2LMN, 2LMP and 2LMQ. Patavastu *et al.* represented each polymorph as two models corresponding to positive (2LMN and 2LMP) and negative stagger (2LMO and 2LMQ). Stagger is used to describe the degree to which each strand is displaced along the fibril long axis from C- to N-terminus (Figure 5.9). Solid state NMR data on A $\beta$  suggest a stagger of  $\pm 2$ , whereby the C terminal  $\beta$ -strand of each molecule is displaced by 7.2 Å, the equivalent of 1.5 times the inter-strand spacing. Paravastu *et al.* (2008) are however unable to distinguish between the two (Petkova et al., 2002). Residues 1-9 of A $\beta$ <sub>1-40</sub> are considered to be unstructured and are therefore not included in the models (Petkova et al., 2002).



**Figure 5.9 Schematic example of stagger along a fibril axis. Images show examples of negative staggers of -1 and -2 respectively.**

A common feature of these models is the tertiary structure of a single A $\beta$  molecule, a single hairpin, which given the sequence similarity at this region, may also represent the tertiary structure of medin.

These models represent two examples of different polymorphs generated from the same primary sequence, and are used here as templates for modelling the tertiary structure of medin. The quaternary packing within the A $\beta$ <sub>1-40</sub> fibrils and the two- and three-fold molecular symmetries (Figure 5.5) of the two morphologies were disregarded in the modelling.

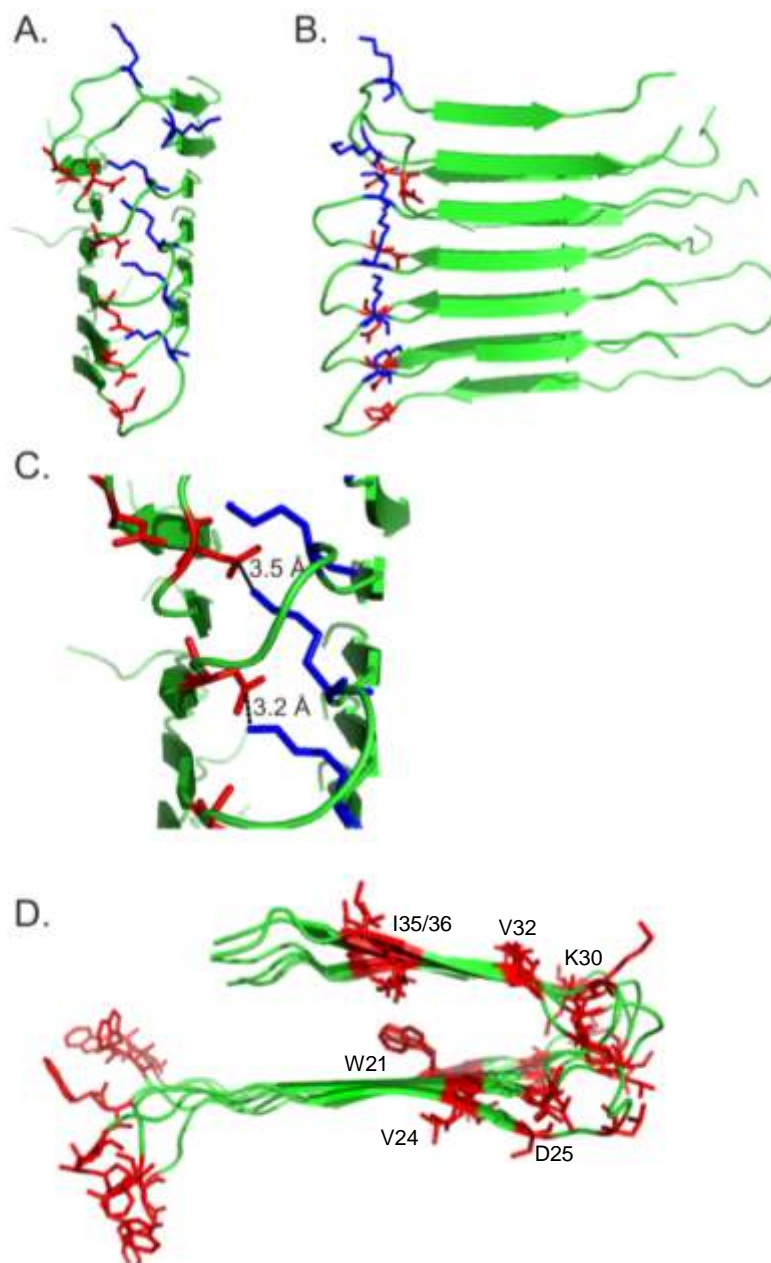
#### **5.3.1.2. Medin models**

Four medin models were generated using Modeller version 9.11 (Eswar et al., 2007) for the truncated medin 10-40 sequence. Truncated medin was used to match the length of the template model. The best model from each ensemble of 20 was selected based on the lowest molpdf score and shown in Figure 5.5. These models gave comparable WHAT IF scores, indicating that the side chain groups are similarly well packed and a similar number of steric clashes calculated by Chimera, to the template models.

The models show several possible stabilising interactions within medin fibrils. There are 9 hydrophobic residues within the proposed amyloid core. It is likely that these residues form hydrophobic interactions that stabilise the structure. In particular W21 and I35 are on opposite sides of the turn and with their side chains oriented inwardly (Panel D Figure 5.10). The phenylalanine residues at positions 8, 43 and 48 are outside the core of the proposed fibril but may form  $\pi$ -stacking interactions that may stabilise the termini. The turn region in the medin model encompasses residues 25-30 and has an appropriate geometry to permit a salt bridge between D25 and K30. Given

the remarkable conservation at these positions, the salt bridge hypothesis was tested experimentally.

The models were visualised in PyMOL (V1.5.0.4) and the distances between C $\gamma$  of D25 and N $\zeta$  of K30 measured. Models based on 2LMO and 2LMN (two-fold symmetry) exhibited many close intermolecular distances between D25 and K30, several less than 4 Å and, therefore, capable of forming ionic interactions (Kumar and Nussinov, 2002) (Figure 5.10). Generally, models based on 2LMP and 2LMQ (three-fold symmetry) gave distances greater than 6 Å (data not shown).



**Figure 5.10 Closer analysis of model medin using 2LMN as a template (A-C). Images show possible intermolecular salt bridge interactions with distances between Lys  $N_{\zeta}$  and Asp  $C_{\gamma}$  shown in black. Lysine residues shown in red, aspartate in blue. Schematic to demonstrate the location of hydrophobic residues (red) within the model (D).**

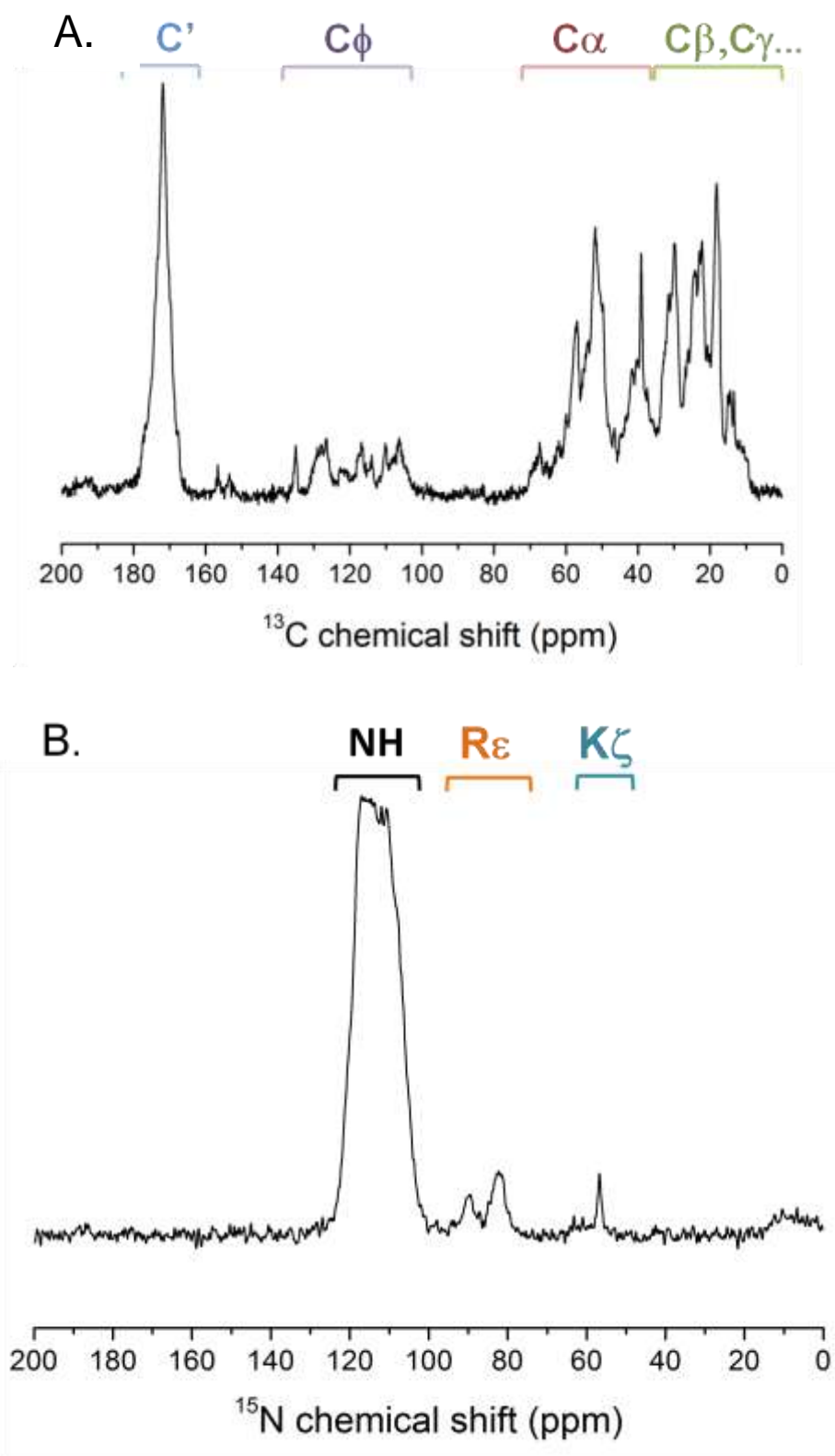
### 5.3.2. Structural studies of wild type recombinant medin by solid state NMR

Chapter 3 described the development of a procedure to produce full-length, uniformly  $^{13}\text{C}$  and  $^{15}\text{N}$  labelled medin. The uniformly labelled protein enabled two-

dimensional solid-state NMR experiments to examine the global conformation of medin in a fibrillar preparation as the first step to test the structural models generated.

#### **5.3.2.1. One-dimensional solid-state NMR**

Proton decoupled  $^{13}\text{C}$  and  $^{15}\text{N}$  spectra were collected for full length recombinant medin. The data show that medin was successfully isotopically labelled with both  $^{13}\text{C}$  and  $^{15}\text{N}$ . Characteristic chemical shifts were annotated according to the Wang and Jardetzky (2002) (Figure 5.11). Quaternary carbon atoms are located between 165-180 ppm, aromatic carbon atoms between 105-135 ppm and the aliphatic carbon atoms between 10-70 ppm. There is a small peak at approximately 160 ppm corresponding to the guanidinium of arginine. (Wang and Jardetzky, 2002).

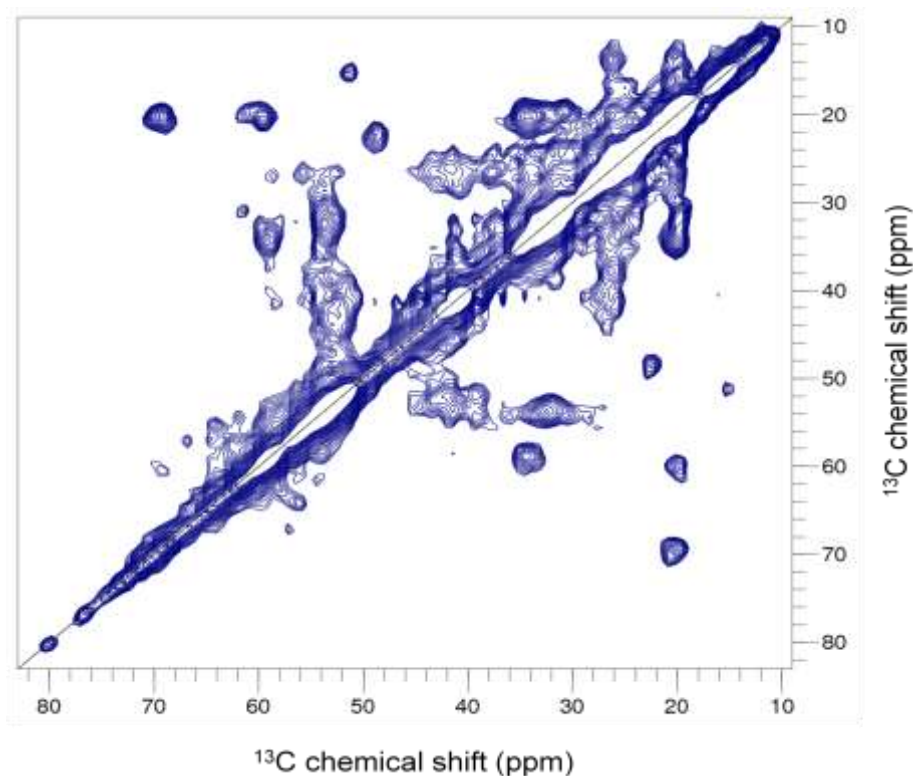


**Figure 5.11** Annotated one dimensional solid-state NMR spectra of uniformly  $^{13}\text{C}$  and  $^{15}\text{N}$  labelled medin fibrils. Data was collected on a Bruker 850 MHz wide-bore spectrometer operating at a static magnetic field of 20 T Proton decoupled  $^{13}\text{C}$  spectrum obtained after 1024 scans (A) Proton decoupled  $^{15}\text{N}$  spectrum obtained after 20480 scans (B).



### 5.3.2.2. DARR spectrum and simulated spectra

Dipolar assisted rotational resonance spectra of full length uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$  labelled medin fibrils were collected on the 850 MHz facility at the University of Warwick. The two-dimensional DARR spectrum correlates  $^{13}\text{C}$  spins through dipolar couplings. As seen in the one-dimensional experiments, the peaks in the DARR spectrum are very broad with a large amount of overlap (Figure 5.12), suggesting that the sample is highly heterogeneous. Some cross-peaks could be assigned using chemical shift data available for other amyloid proteins (Petkova et al., 2005b) but full *de novo* sequential assignment of the spectrum was not possible.

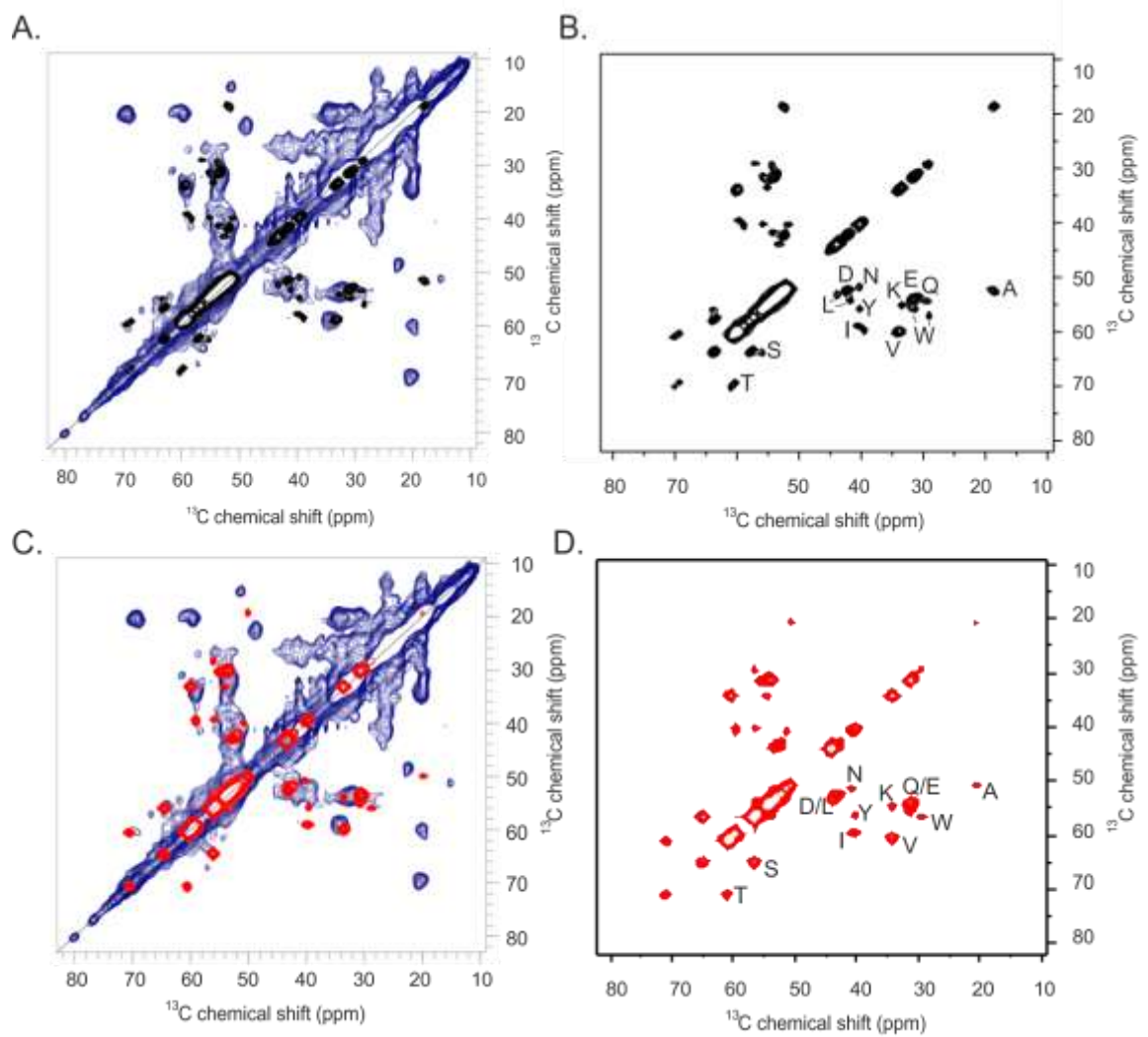


**Figure 5.12 DARR spectrum for uniformly  $^{13}\text{C}/^{15}\text{N}$  labelled recombinant medin following incubation at 25 °C at a concentration of 200  $\mu\text{M}$  in aggregation buffer. Data was collected on a Bruker 850 MHz wide-bore spectrometer operating at a static magnetic field of 20 T with a sample temperature of 23 °C and a spinning rate of 14 KHz.**

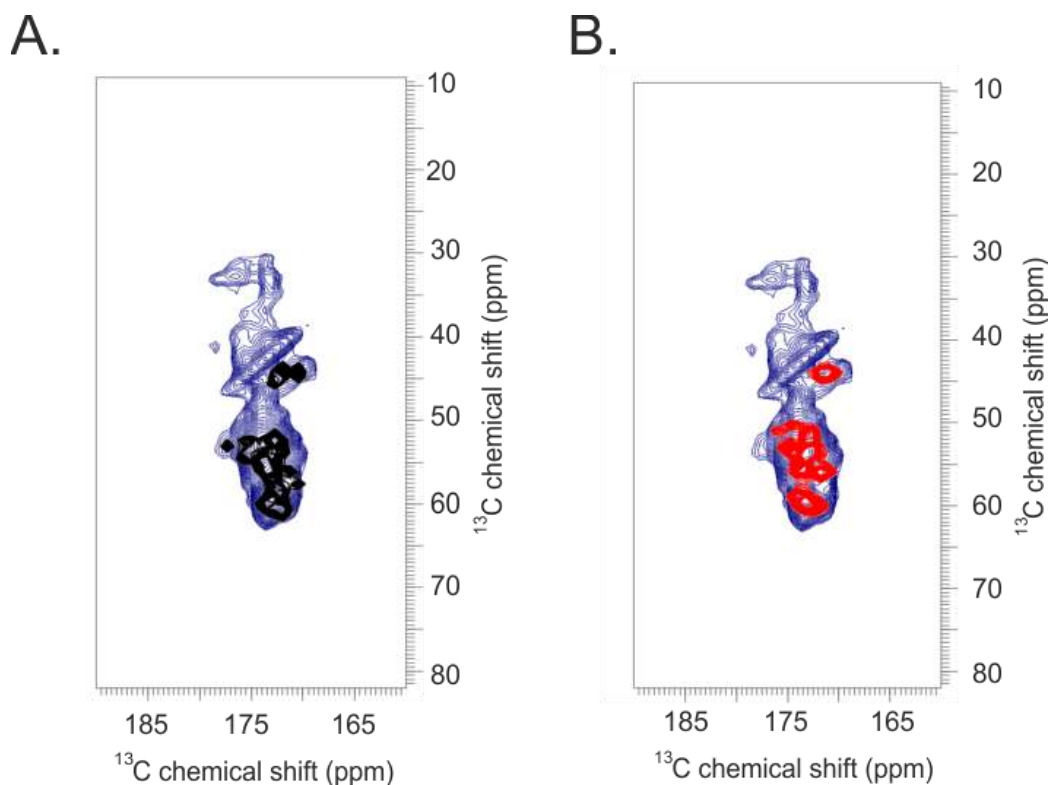
Further multidimensional  $^{13}\text{C}$ - $^{13}\text{C}$  and  $^{13}\text{C}$ - $^{15}\text{N}$  measurements at high field would be necessary to assist the spectral assignment needed for a full structure determination, but these were not possible in the time available. Instead, a different approach was taken in which  $^{13}\text{C}$  chemical shifts were calculated from the medin model shown in Figure 5.10. These shifts were then used to simulate a DARR spectrum to compare with the experimental spectrum.  $^{13}\text{C}$  chemical shifts for  $\text{C}\alpha$ ,  $\text{C}\beta$  and amide carbonyl sites are sensitive to the backbone  $\phi$  and  $\psi$  angles defining secondary structure. Similarity in the experimental and simulated spectra, as judged by the overlap of cross-peaks, would therefore suggest that the model broadly comprised the correct secondary structural elements.

DARR spectra were generated from the hairpin model of medin shown in Figure 5.10, using the online server SHIFTX2 which generates  $^{13}\text{C}$  chemical shifts values from protein structures (Han et al., 2011) and a program to convert the values into a pseudo 2D spectrum. The experimental data (blue) are in good agreement with the simulated spectrum (black, Figure 5.13, 5.14 and 5.15) and overlap in the vast majority of the spectrum. This suggests that medin fibrils have a predominantly  $\beta$ -sheet conformation, and share some structural similarity with the proposed model. A second spectrum was simulated for a model of medin in a continuous  $\beta$ -sheet conformation lacking the hairpin (red, Figure 5.13, 5.14 and 5.15). Although there is again overlap of many cross-peaks, on visual inspection, the match for the linear model appears to be not as close as for the hairpin model. The chemical shift predictions are not accurate enough to distinguish between these two models however.

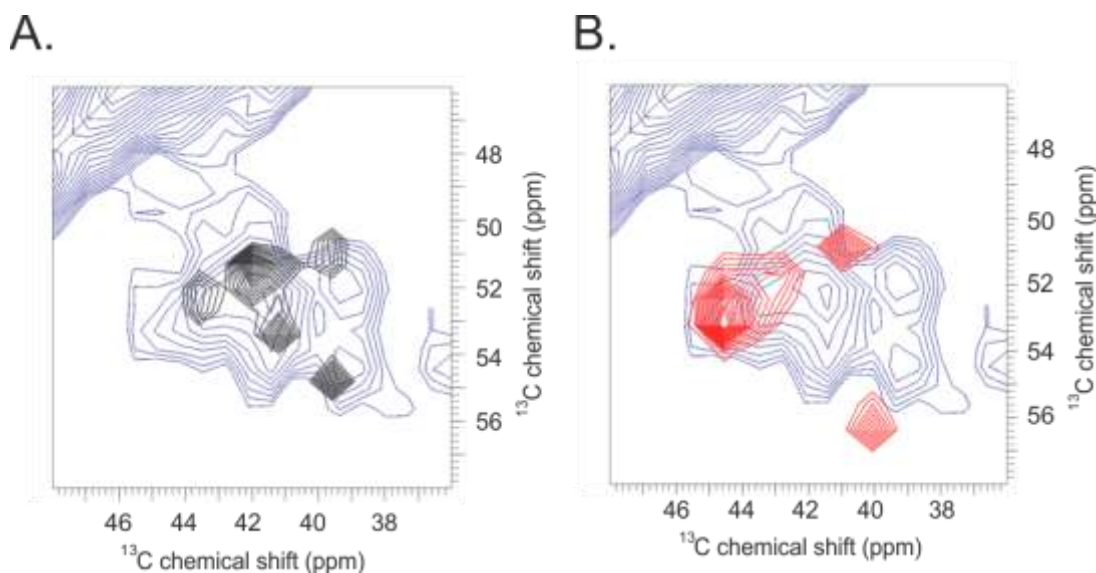
Further simulated spectra were generated from average chemical shift values for medin in random coil and complete  $\alpha$ -helical structures (Appendix 3), taken from the BMRB (Ulrich et al., 2008). In these cases the overlap with the experimental spectra was substantially poorer than for the models above, suggesting that these structural elements do not feature highly in the medin fibrils.



**Figure 5.13 Spectra showing  $\text{C}\alpha$  chemical shifts for recombinant medin (blue) overlaid with corresponding simulated spectra calculated for medin model based on  $\text{A}\beta$  template (black, panel A) and a linear model (red, panel C). The assigned simulated spectra for the medin model based on  $\text{A}\beta$  template and the linear model are shown in panels B and D respectively.**



**Figure 5.14 Carbonyl region of recombinant medin DARR spectrum. Simulated spectra for models based on A $\beta$  (black, panel A) and a linear medin model (red, panel B).**



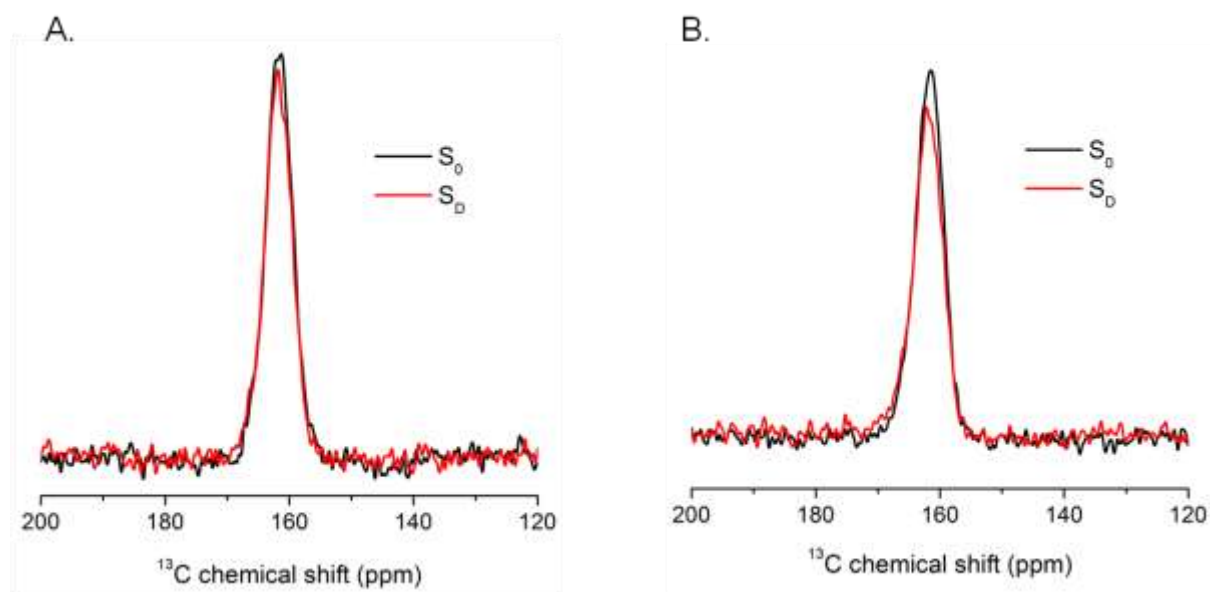
**Figure 5.15 Expanded region of recombinant medin DARR spectrum showing the differences between simulated spectra for models based on A $\beta$  (black, panel A) and a linear medin model (red, panel B).**

### 5.3.3. Testing the salt bridge hypothesis using solid state NMR frequency selective REDOR experiments.

In order to investigate the hypothesis of a salt bridge within medin fibrils, a second synthetic peptide was produced which contained selective isotope labels including  $^{13}\text{C}/^{15}\text{N}$  D25 and K30 (for full details see section 5.2.2.1). This hypothesis was testing using frequency selective REDOR experiments. This technique allows the detection of  $^{13}\text{C}$ - $^{15}\text{N}$  dipole-dipole couplings between a selected spin pair, in this case D25  $\text{C}_\gamma$  and K30  $\text{N}_\zeta$ . Initially a reference spectrum is collected using standard cross-polarisation pulse sequence with a  $90^\circ$  pulse on the observed nucleus (black lines, Figure 5.16). The second spectrum is obtained by the addition of a second set of  $90^\circ$  pulses on the dipolar coupled spin which results in a drop in signal intensity  $\Delta S$  which is related to the internuclear distance between the two spins (red lines, Figure 5.16).

$$\Delta S = \frac{1}{2\pi} \int_0^{\pi/2} \int_0^{2\pi} \cos\left(\pm 2\sqrt{2} \frac{d}{\omega_R} \sin 2\beta \sin \gamma\right) \sin \beta d\beta d\gamma$$

$\Delta S$  can be simulated for different values of the dipolar coupling constant  $d$  until a good fit is achieved. The dipolar coupling constant can then be used to calculate an inter-nuclear distance using equation 2.3. In order to obtain an accurate internuclear distance multiple experiments but be conducted at a variety of dephasing times. Due to the length of the experiments and the loss of intensity at longer dephasing times it was not possible to obtain a precise internuclear distance. However, the limit of  $^{13}\text{C}$ - $^{15}\text{N}$  dipolar coupling is approximately  $5 \text{ \AA}$ . Dephasing observed at the two mixing times could only occur if these is nuclei (D25  $\text{C}_\gamma$  and K30  $\text{N}_\zeta$ ) were within  $5 \text{ \AA}$  and may therefore permit salt bridge formation.



**Figure 5.16** Frequency selective REDOR spectra for full length synthetic medin at two different mixing times, 5.5 ms (A) and 8.2 ms (B). Data was collected on a Bruker Avance 400 spectrometer operating at a magnetic field of 9.3 T. Reference spectra are shown in black and spectra with additional de-phasing are shown in red. The loss in signal intensity is indicative of an internuclear distance of less than 5 Å.

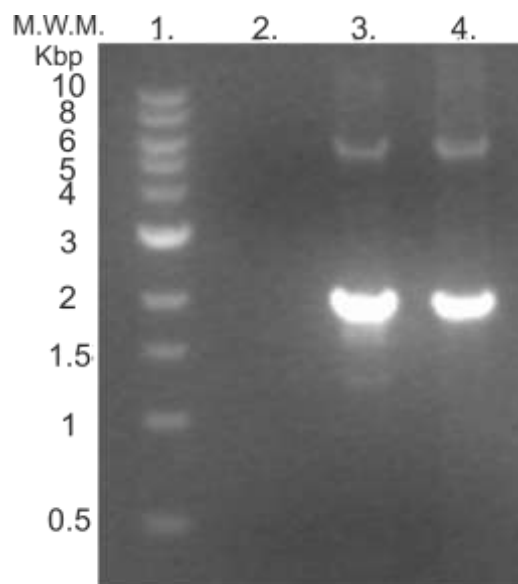
#### 5.3.4. Mutagenesis

Modelling and NMR data suggest that a salt bridge is feasible. A method of testing this experimentally is to create mutants that are prohibitive to salt bridge formation and any effects they may have on the kinetic and morphological properties of medin.

##### 5.3.4.1. Cloning

Two mutants, D25N and K30I were generated using the SLIM method and confirmed by sequencing (Appendix 2). The medin sequence was amplified using two long tailed primers (containing the mutation) and two short primers (Appendix 2). The efficacy of the process can be verified by agarose gel electrophoresis. As anticipated the gel shows two strong bands that correspond to the correct number of base pairs. The control reaction (lane 2, Figure 5.17), containing the template with

no primers present, is blank indicating that the bands seen in lanes 3 and 4 are products of the PCR reaction not the template DNA (Figure 5.17).



**Figure 5.17 Agarose gel electrophoresis of the mutant PCR products. Lane 1. 1 Kbp DNA ladder (NEB). Lane 2. Control PCR reaction with no mutagenesis primers. Lane 3. PCR amplification product from primers containing D25N mutation. Lane 4. PCR amplification product from primers containing K30I mutation.**

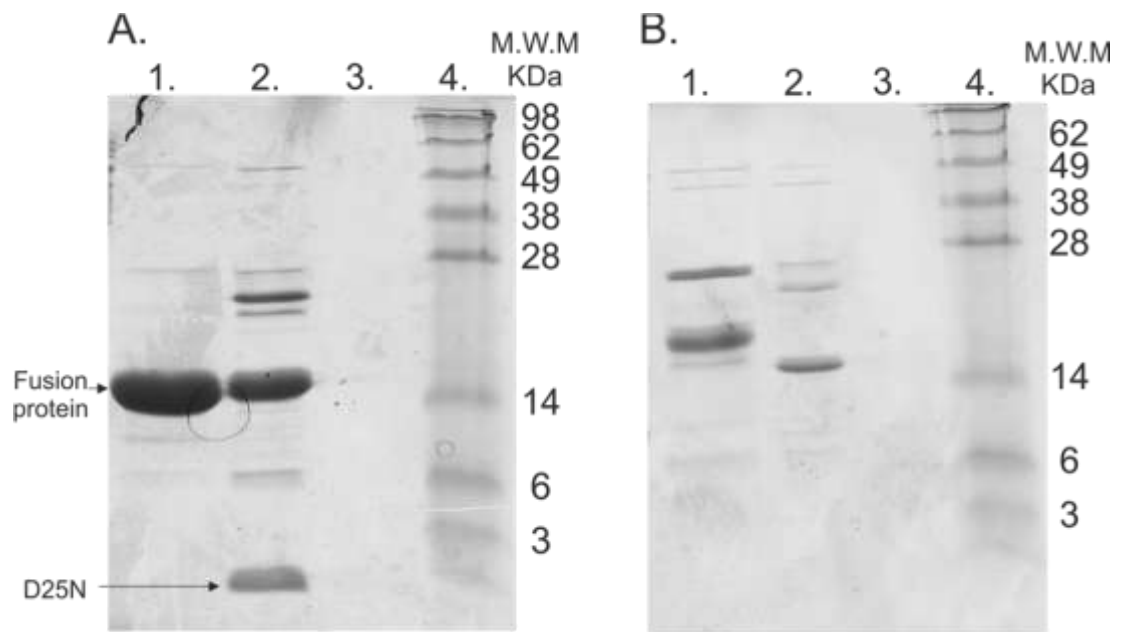
Sequencing of the resultant plasmids showed the presence of the single point mutations, confirming successful cloning (Appendix 2). The mutant proteins were then successfully expressed and purified using the protocol described for medin, in the same vector.

#### **5.3.4.2. Mutant expression**

D25N and K30I were successfully expressed in *E.coli* giving yields of 1.5 and 3 mg of pure protein per litre of culture media. Purification of the mutants followed the protocol established in section 3.2.1.8. The fusion protein was isolated from the crude extract by Ni<sup>2+</sup> affinity chromatography (Lane 1, Figure 5.18). D25N expressed better than K30I giving a higher yield of fusion protein (Panel A, Lane 1



Figure 5.18) The affinity tag was cleaved from both proteins using SUMO protease. K30I was fully cleaved (Panel B, Lane 2 Figure 5.18), it does not seem however that D25N was fully cleaved (Panel A, Lane 2 Figure 5.18). Given the increased yield of D25N it may be that insufficient protease was added. Following cleavage the mutants were isolated from the cleavage mixture using reverse affinity purification.



**Figure 5.18 15 % SDS PAGE gel images of D25N (A.) and K30I (B.) purification and cleavage steps. Lane 1. Eluted from HisTrap™, Lane 2. Cleavage mixture, Lane 3. Reverse HisTrap flow through, Lane 4. SeeBlue® plus 2 protein standard (Invitrogen)**

As described in section 3.2.3.1, WT medin was difficult to visualise using standard electrophoresis and staining methods. The same problems were observed for both mutant peptides. D25N can be seen in lane 2 (Panel A, Figure 5.18) but it is running on the dye front. Furthermore, following reverse affinity purification, despite showing significant absorbance at 280 nm, there is nothing visible on the gel (Lane 3, Panel A, Figure 5.18). The reverse affinity purification dilutes the protein fraction approximately 10-fold. This dilution goes some way to explaining the lack of a

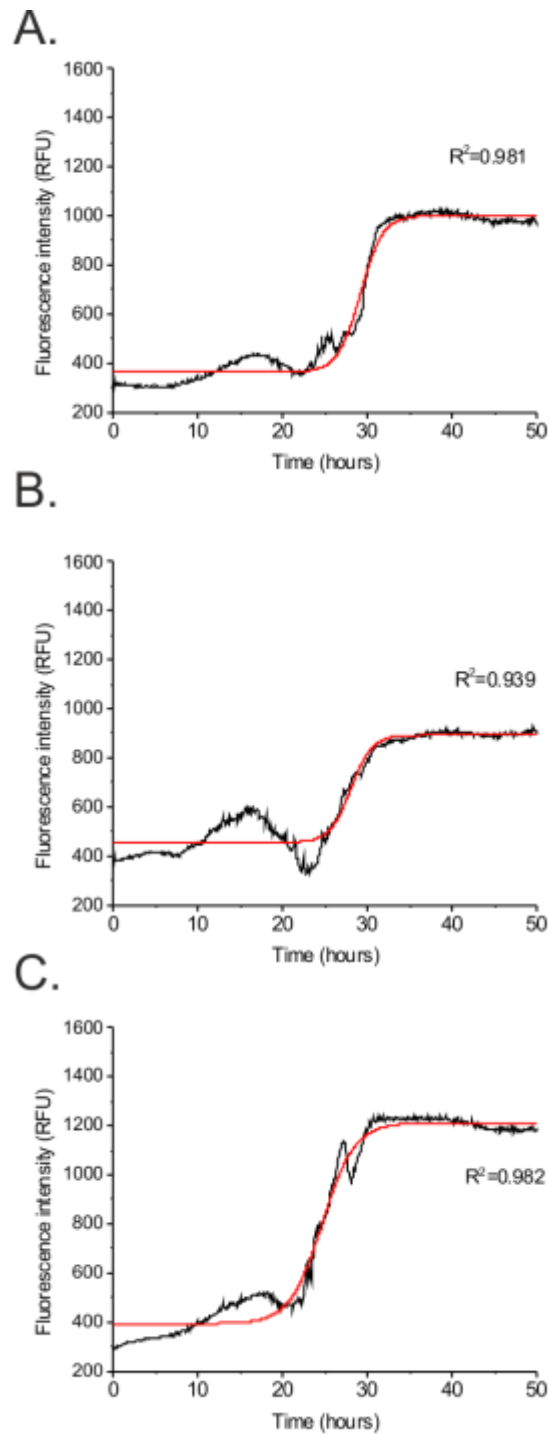
visible protein band in the flow through (Panel A, Figure 5.18). K30I is not visible in either lane (Panel B, Figure 5.19) the lower expression and resultant purification yields may explain why it not visible despite a positive absorbance reading at 280 nm.

The initial expression yield of these mutants was sufficient for the work described here; therefore, the expression, purification and visualisation did not undergo any optimisation.

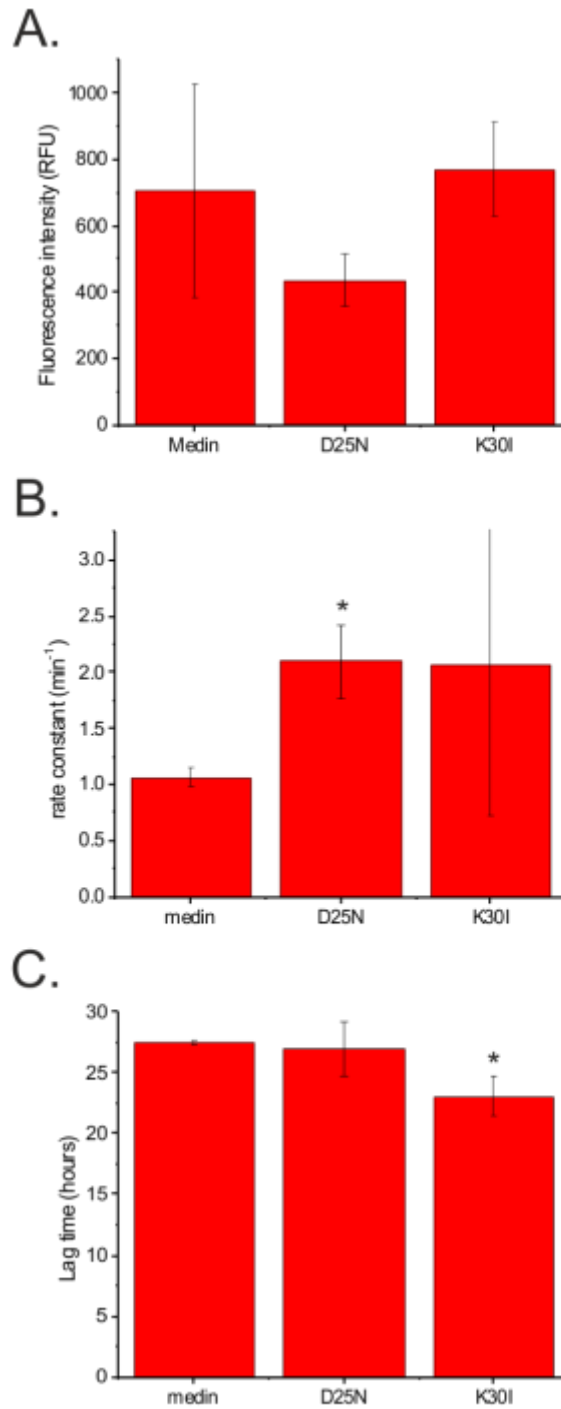
### **5.3.5. Aggregation kinetics of mutant peptides D25N and K30I.**

#### **5.3.5.1. ThT analysis**

ThT fluorescence of the mutant peptides was carried out to assess the effect of the point mutations on the aggregation kinetics of medin. The peptides were analysed under the same conditions as the wild type peptide (section 4.2.1) and three properties compared: the end point fluorescence intensity, lag time and elongation rate constants. Neither of the mutations had a large effect on the aggregation of medin but there are some statistically significant differences. D25N had a faster rate constant than WT medin (Figure 5.19 and 5.20). Conversely, K30I had reduced lag time compared with WT medin (Figure 5.19 and 5.20). There is a relatively small increase in fluorescence intensity after approximately 15 hours, followed by a small decrease. This change precedes the main increase in fluorescence and may represent the formation of oligomers. Interestingly the largest 'hump' is seen for D25N which ultimately results in the lowest overall increase in fluorescence intensity (Panel B, Figure 5.20). This may suggest the formation of off-pathway oligomeric species. This indicates that the mutations do not prevent the formation of amyloid according to the ThT assay.



**Figure 5.19** Aggregation of mutants compared to wild type medin analysed by thioflavin T fluorescence. WT medin, D25N and K30I are shown in panels A-C respectively. Data were recorded on a fluoromax spectrometer at 30 °C on a 20  $\mu$ M protein solution in aggregation buffer. The red line represent the fit using the following equation:  $y = y_0 + \frac{a}{1 + e^{-\frac{t-T_i}{\tau}}}$ . The adjusted  $R^2$  values indicating the goodness of fit for each curve are reported on the right hand side. Data represent mean data for 3 replicates.

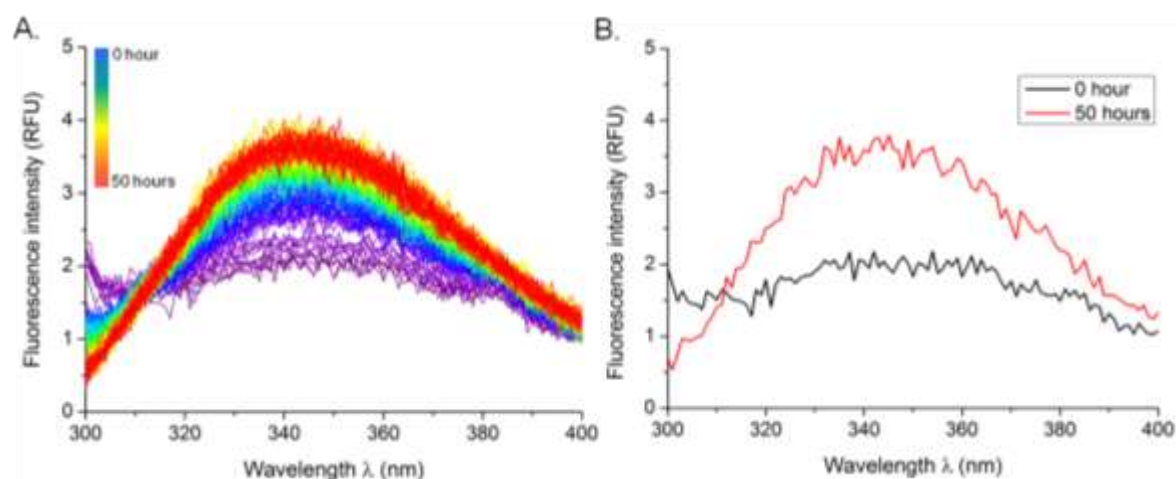


**Figure 5.20** Analysis of the effect of mutations on fluorescence intensity (A), lag time (B) and rate constant (C). Values were extracted from fitting exponential curves to the data and calculating the lag phase, the maximum fluorescence intensity and the calculated rate constants for each condition. Data represent mean and standard error values for 4 replicates. Asterisk represent statistical significance where  $p=0.05$ .

### 5.3.5.2. Aggregation kinetic of mutants using intrinsic fluorescence

Intrinsic fluorescence (IF) measurements were carried out on the two mutant peptides to investigate the effect of the mutations on the tryptophan environment.

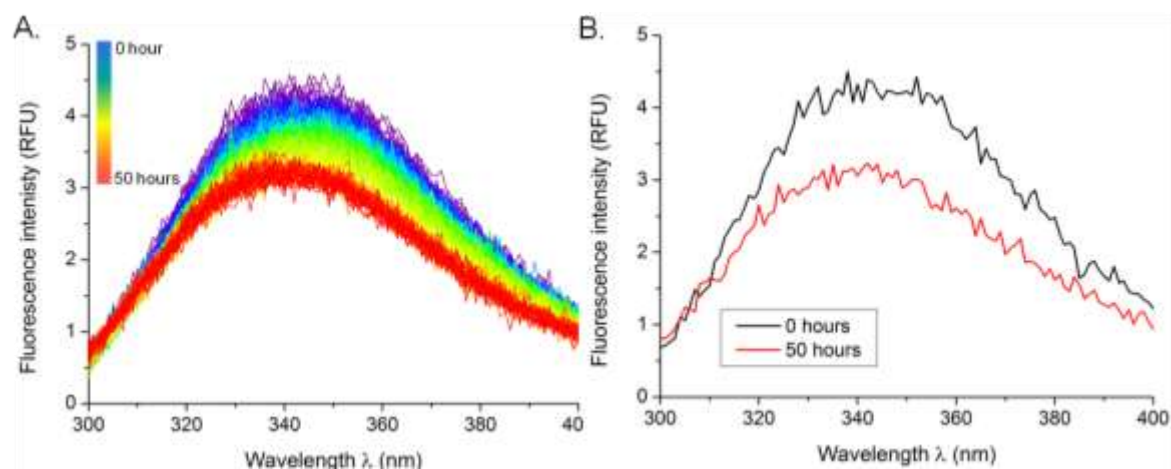
Surprisingly, D25N demonstrated an increasing maximum fluorescence at 355 nm over time, consistent with increasingly solvent exposed tryptophan residues. There is also no change in the emission wavelength, suggesting no change in the local environment. This is in contrast to wild type medin, for which a time-dependant decrease in fluorescence suggests a switch from solvent exposed tryptophan residues, to buried within the hydrophobic core as discussed in section 4.3.1.3.



**Figure 5.21** Changes in intrinsic tryptophan fluorescence of D25N over time at 30 °C, at a protein concentration of 20  $\mu$ M. Tryptophan residues were excited at 279 nm and emission spectra were recorded from 300-400 nm. Data were collected every 15 minutes for 50 hours, raw data shown in (A) and initial and final spectra shown in (B).

K30I demonstrates a decrease in fluorescence maximum at 355 nm suggesting that the tryptophan residues are less solvent exposed (Figure 5.22). This may be a result of precipitation, the aggregation profile obtained from ThT measurements gave the highest overall fluorescence suggesting the greatest overall amount of aggregation (Figure 5.19). K30I, unlike wild type medin, does not undergo the same switch from

solvent exposed to buried within the hydrophobic core. This implies an alternative aggregation mechanism.



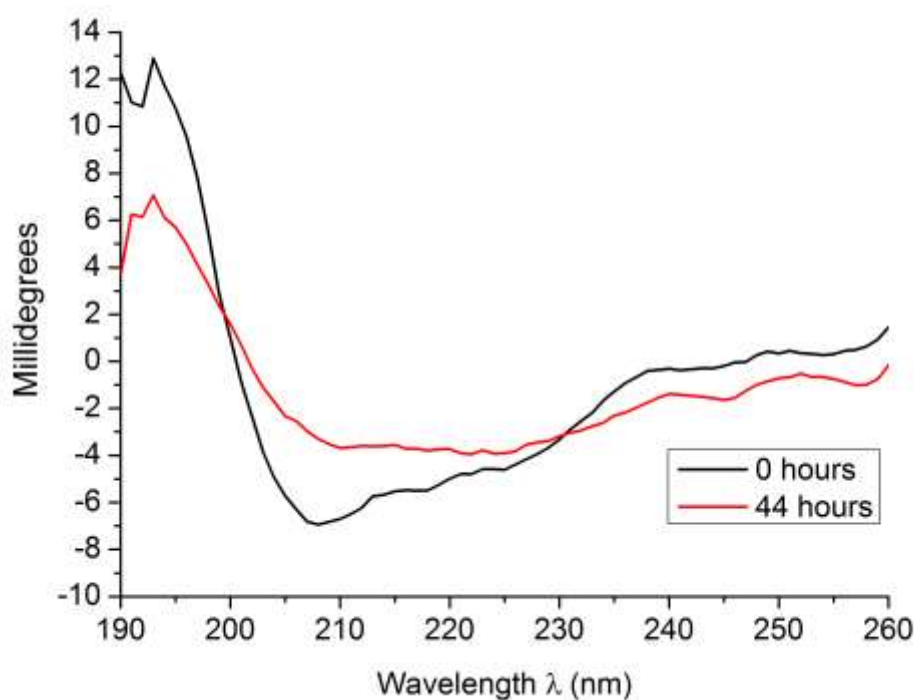
**Figure 5.22** Changes in intrinsic tryptophan fluorescence of K30I over time at 30 °C, at a protein concentration of 20  $\mu$ M. Tryptophan residues were excited at 279 nm and emission spectra were recorded from 300-400 nm. Data were collected every 15 minutes for 50 hours, raw data shown in (A) and initial and final spectra shown in (B).

### 5.3.6. Structure and morphology of mutants

Although the kinetic analysis by ThT indicated that the mutations had little effect on the aggregation potential of medin, it is possible that these mutations have alternative aggregate structures. Accordingly, the secondary structure content and the morphology of the aggregates were examined by circular dichroism and electron microscopy respectively.

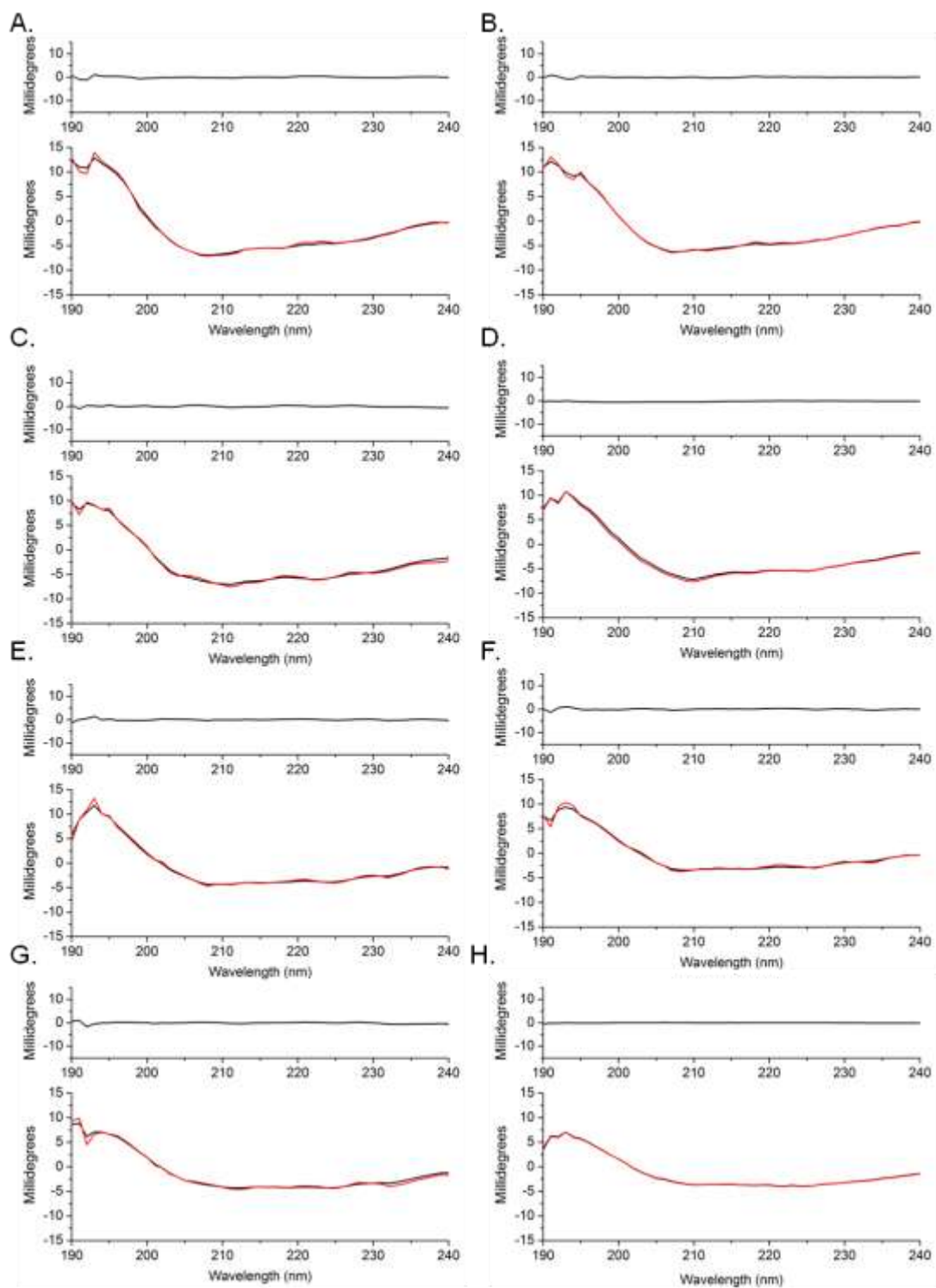
#### 5.3.6.1. Monitoring secondary structure of D25N over time

Circular dichroism was used to monitor the secondary structure content of freshly prepared mutants, over a period of 44 hours, at 2 hourly intervals. Experiments were carried out on Beamline 23 at Diamond light source under the same conditions as for wild type medin (see section 4.2.4 for details).



**Figure 5.23** CD spectra for D25N comparing start and end points (0 and 44 hours). Data were recorded at 30 °C on medin (200 μM, in CD buffer) using far-UV synchrotron radiation circular dichroism at beam line B23 at Diamond light source.

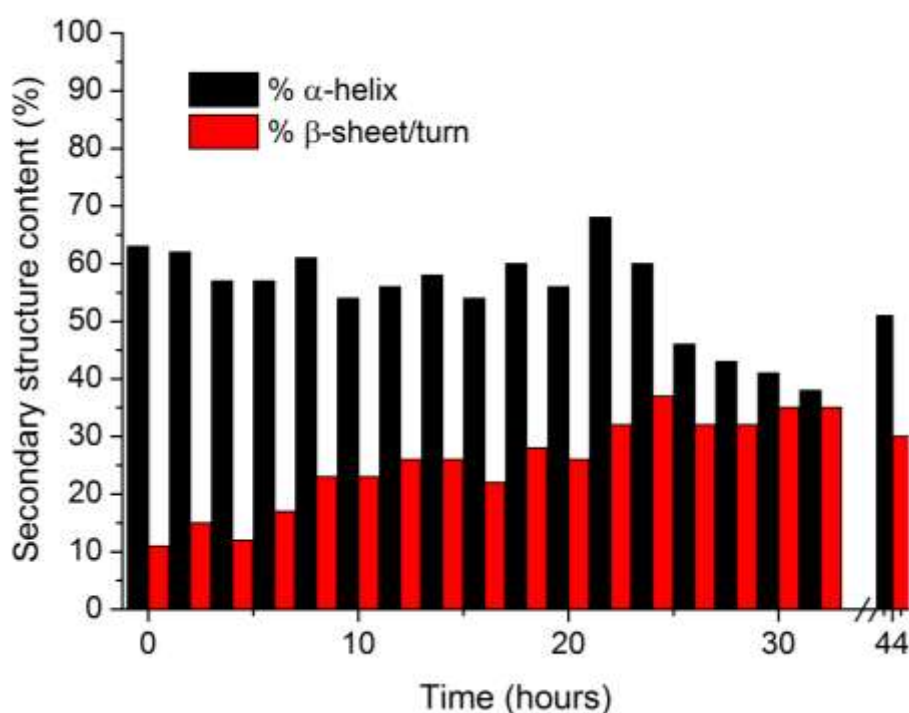
Circular dichroism data for D25N shows that it has a predominantly helical structure (Figure 5.23 and 5.24), approximately 60 % when estimated using CONTILL and CDSTRR fitting methods. The amount of  $\alpha$ -helix present remains relatively constant throughout the time course (Figure 5.25). There is a slight shift towards  $\beta$ -sheet near the end of the time course, from approximately 10 % to 30 % (Figure 5.25). This shift occurs at the same time as the fluorescence intensity begins to increase in ThT measurements; after 25 hours (Figure 5.19). Overall, D25N has more  $\alpha$ -helix than WT medin throughout the time course. There is small conversion to  $\beta$ -sheet, approximately 20%, but it is not as pronounced as the conversion seen in WT medin (65 %).



**Figure 5.24** Circular dichroism spectra of D25N (200  $\mu$ M, in CD buffer) at different time points: 0, 6, 12, 18, 24, 30, 34 and 44 hours, A-H respectively. Experimental data shown in black and secondary structure fits in red. Residual data from the fits are shown above each data set. Data were recorded at 30  $^{\circ}$ C using far-UV synchrotron radiation circular dichroism at beam line B23 at Diamond light source.



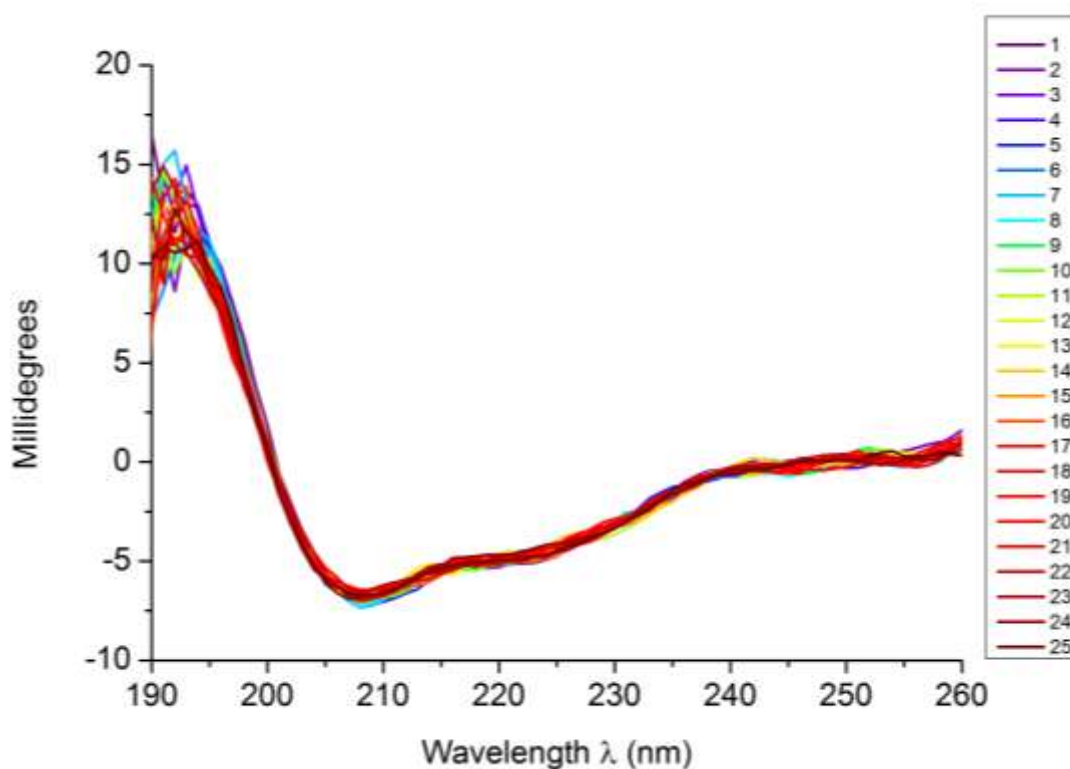
Collectively the ThT, IF and CD data show that D25N has a distinct aggregation pathway to wild type medin. This indicates that D25 has a key role in amyloidogenesis. The preliminary ‘hump’ seen in the ThT, and the large degree of  $\alpha$ -helix observed in the CD may suggest the formation of oligomeric species. This is also consistent with the solvent availability of the Trp residues within D25N suggesting that the species are soluble. These may be off-pathway stable oligomeric species or they may go on to form fibrillar aggregates. The increase in ThT observed at 30 hours and the conformational change seen in CD at the same point implies that there is a change in aggregate structure at this point. The increase in ThT and the conformational change for D25N however is less pronounced than for wild type medin which may suggest that only some of the population undergoes further aggregation.



**Figure 5.25** Secondary structure analysis of D25N over 44 hour time course. CD data were analysed using, CONTIN or CDSSTR secondary structure fitting software, percentage values for  $\alpha$ -helix (black),  $\beta$ -sheet/turn (red), extracted and plotted over time.

### 5.3.6.2. Investigating the stability of D25N in response to UV radiation

To ensure the secondary structure analysis reported changes in the protein conformations rather than degradation, UV stability experiments were carried out. This experiment further provided us with information about the general stability of the peptide. These data show that D25N is stable in response to UV radiation for a total exposure time of 32.5 minutes (Figure 5.26). Secondary structure experiments were carried out for a total UV exposure time of 28.6 minutes, therefore any changes observed are a result of secondary structure changes not degradation.

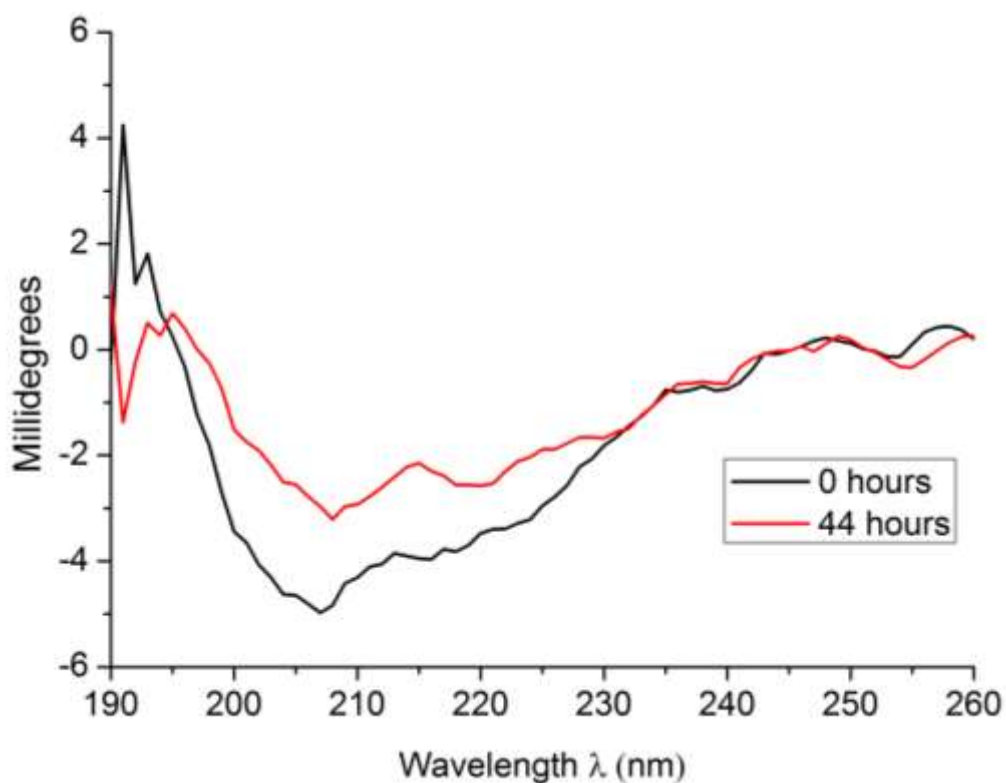


**Figure 5.26** Analysis of D25N stability in response to UV radiation, key refers to scan number. Data were recorded over 25 scans at 30 °C on medin (200 μM, in CD buffer) using far-UV synchrotron radiation circular dichroism at beam line B23 at Diamond light source.

### 5.3.6.3. Monitoring the secondary structure of K30I over time.

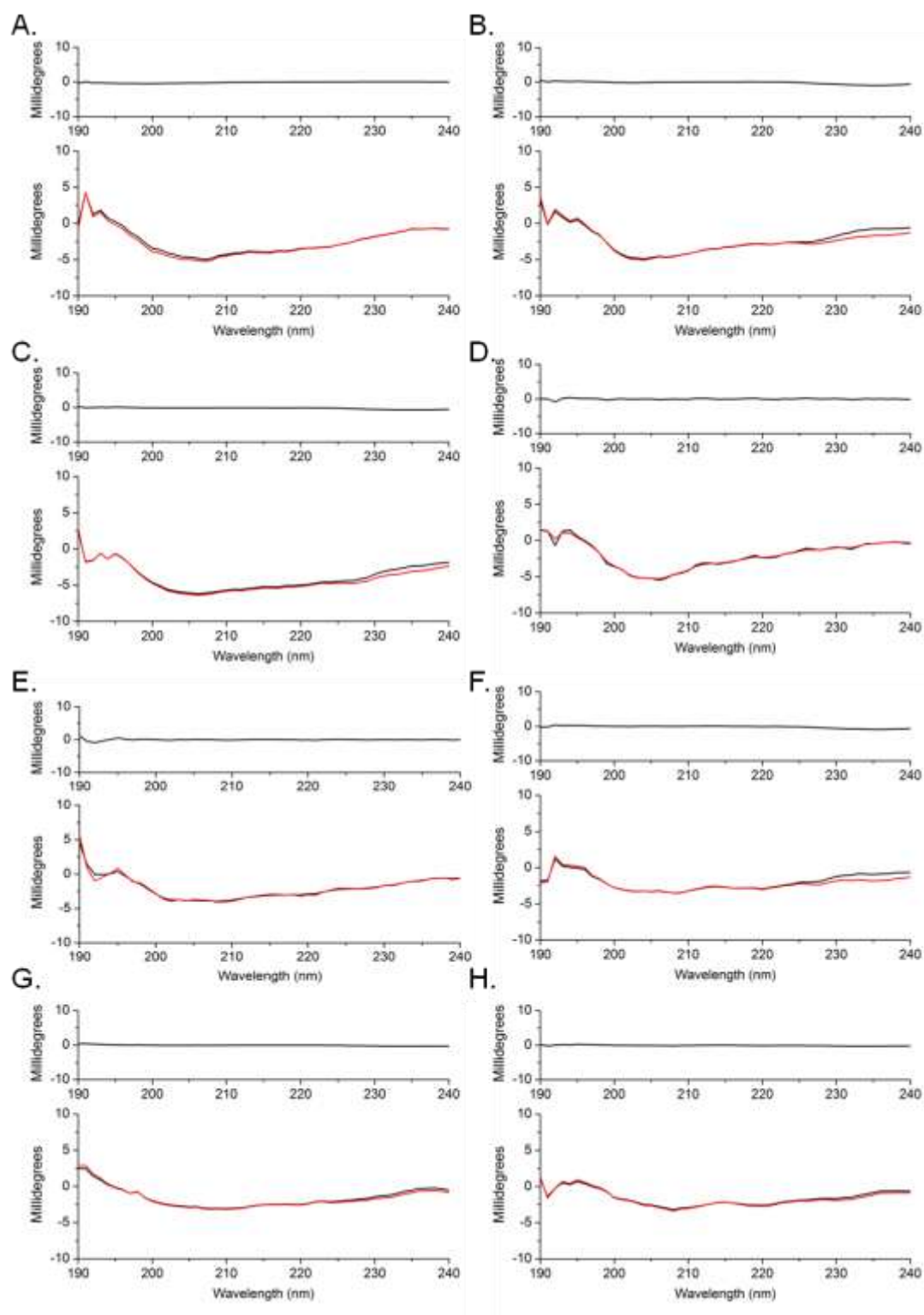
The secondary structure content of K30I was also investigated over 44 hours. Spectra shown in Figure 5.27 and 5.28 do not demonstrate any significant change in secondary structure over time. There is a loss in intensity after 44 hours but the two spectra are virtually superimposable (Figure 5.27). This is confirmed when the data is analysed using secondary structure fitting algorithms (CDSSTR and CONTIN) (Figure 5.29). The amount of each secondary structure element remains constant throughout the time course, approximately 25 %  $\alpha$ -helix, 40 %  $\beta$ -sheet/turn and 35 % random coil (

Figure 5.29).

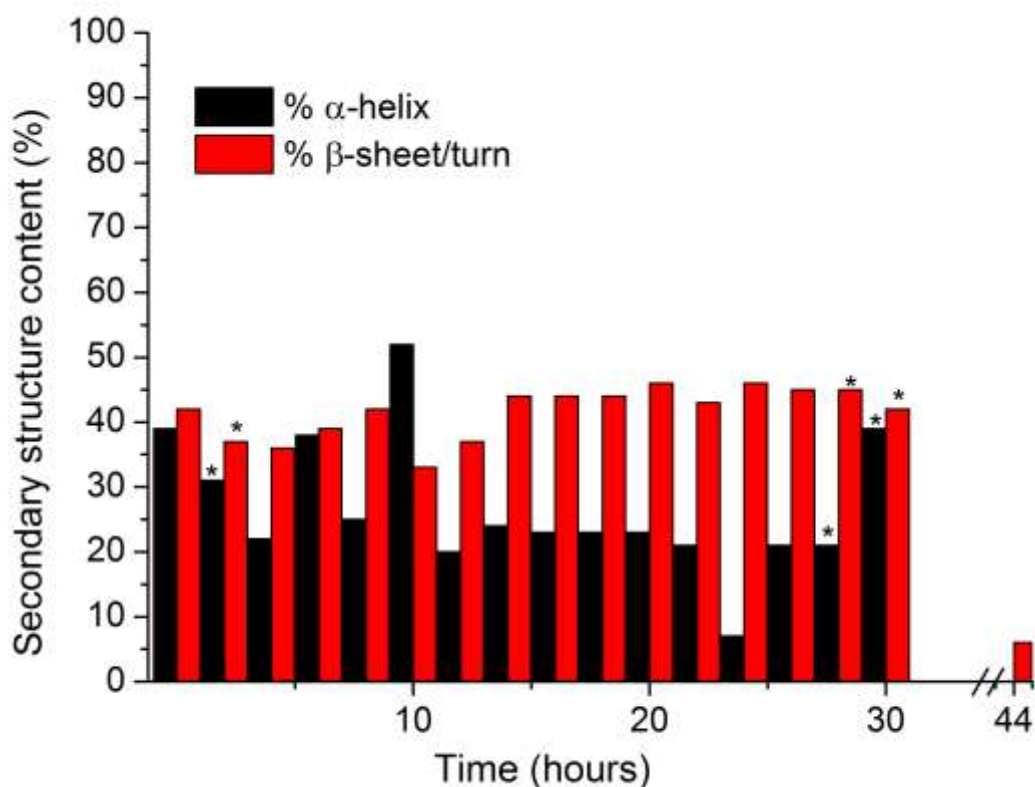


**Figure 5.27** CD spectra for K30I comparing start and end points (0 and 44 hours). Data were recorded at 30 °C on medin (200  $\mu$ M, in CD buffer) using far-UV synchrotron radiation circular dichroism at beam line B23 at Diamond light source.

The end point data for K30I has a higher signal to noise ratio and a lower starting intensity than both D25N and wild type medin suggesting that some of the peptide has precipitated (Figure 5.27). Unlike D25N and wild type medin, the aggregation seen with K30I is not associated with a change in secondary structure that could be measured using this technique. K30I, like D25N, appears to have a distinct aggregation pathway to wild type medin, again suggesting a key role for K30 within amyloidogenesis.



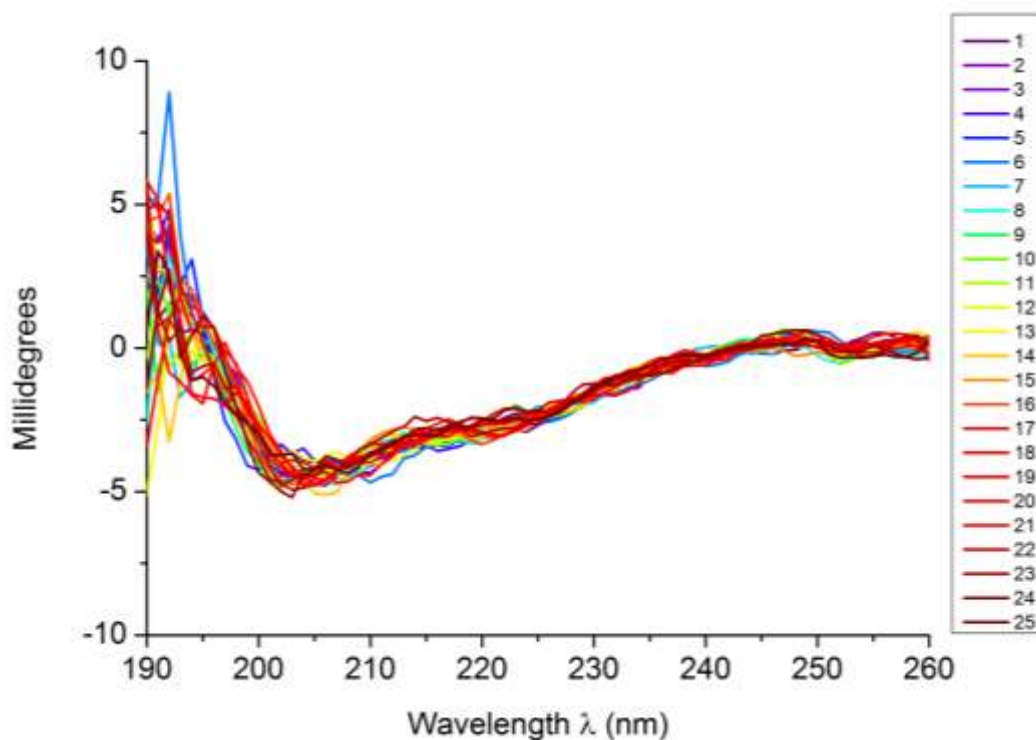
**Figure 5.28** Circular dichroism spectra of K30I (200  $\mu$ M, in CD buffer) at different time points: 0, 6, 12, 18, 24, 30, 34 and 44 hours, A-H respectively. Experimental data shown in black and secondary structure fit in red. Residual data from the fits are shown above each data set. Data were recorded at 30  $^{\circ}$ C using far-UV synchrotron radiation circular dichroism at beam line B23 at Diamond light source.



**Figure 5.29** Secondary structure analysis of K30I over a time course of 44 hours. CD data were analysed using, CONTIN or CDSSTR secondary structure fitting software, percentage values extracted and plotted over time. Asterisks indicate standard deviation values for the fits of over 0.1.

#### 5.3.6.4. Investigating the stability of K30I in response to UV radiation

As with D25N, it was necessary to check the stability of K30I in response to UV radiation. K30I was less stable than D25N following exposure to UV radiation over 32.5 minutes but this was sufficient stable to assess secondary structure over time (28.6 minutes) (Figure 5.30).

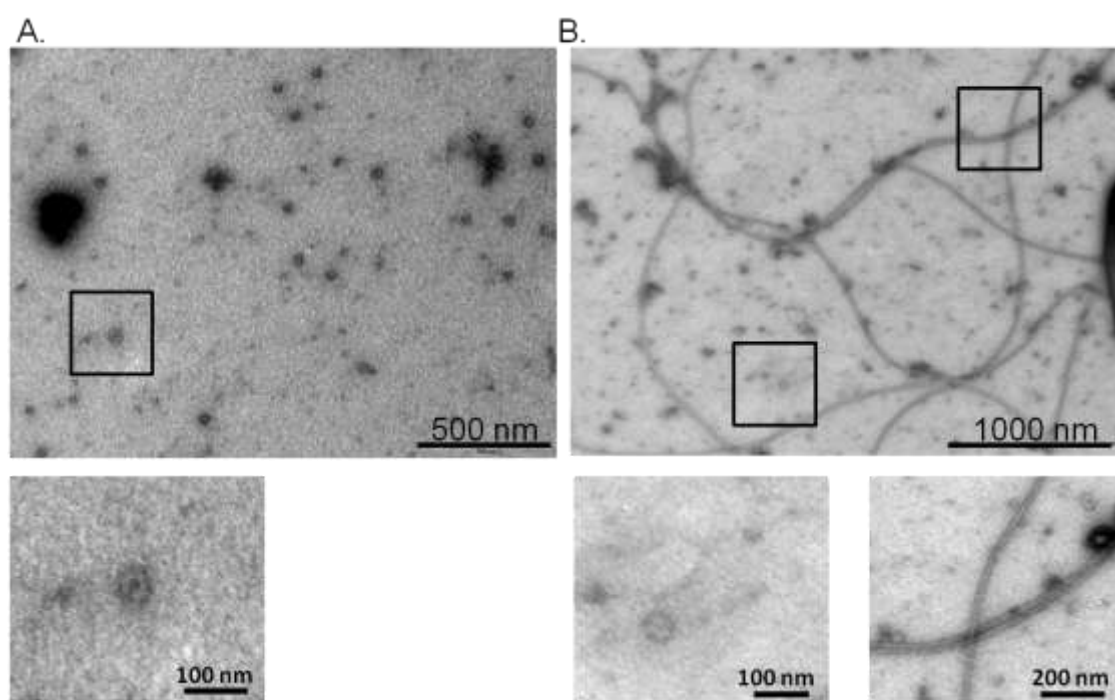


**Figure 5.30 Analysis of K30I stability in response to UV radiation, key refers to scan number. Data were recorded over 25 scans at 30 °C on medin (200 μM, in CD buffer) using far-UV synchrotron radiation circular dichroism at beam line B23 at Diamond light source.**

### **5.3.7. Morphology of mutant aggregates studied by transmission electron microscopy.**

Negative stain electron microscopy was used to examine the morphology of D25N and K30I at the end of the aggregation time course, after 44 hours. There were no fibrillar aggregates visible in the D25N sample (Figure 5.31). There is however a large number of small, ‘doughnut’-shaped assemblies, approximately 50 nm in diameter (Figure 5.31). This correlates with the hypothesised formation of stable oligomers. The subsequent increase in ThT fluorescence seen at 30 hours (Figure 5.19) may be due to a change in oligomer conformation or the association of oligomers into larger assemblies visible in the electron micrograph (Panel A, Figure 5.31).

By contrast K30I showed the presence of long, thin fibres (Figure 5.31), however they do not appear to form the dense networks seen for wild-type medin (Figure 4.16). The presence of large fibrils correlates with the reduction in signal observed in the CD data as the fibrils fall out of solution (Figure 5.27). This also accounts for the large increase in fluorescence intensity observed in the ThT assay (Panel C, Figure 5.20). In addition to the fibrils, there are some small oligomer-like aggregates visible with the same dimensions as those observed for D25N (Figure 5.31).



**Figure 5.31** Negatively stained electron microscopy images of D25N (A) and K30I (B) formed under quiescent conditions, at 30 °C, at a peptide concentration of 20  $\mu$ M in aggregation buffer. Zoomed in sections indicated by boxes are shown below.



## 5.4. Discussion

The recent advances in biophysical techniques have led to huge increase in the number of known protein structures deposited within the Protein Data Bank. The number of depositions however still accounts for only a small portion of all known proteins and the number of solved amyloid protein structures is even fewer. Currently, A $\beta$ <sub>1-40</sub> is the only full length amyloid protein with deposited fibril models in the PDB. Using known structures it is possible to infer information about proteins yet to be solved (Furnham et al., 2013). Molecular modelling is a powerful tool to guide experimental design. A good example of the uses of molecular modelling is in drug/inhibitor design whereby small molecule ligands are modelled in target binding sites of proteins. The results then guide the experimental work (Favia et al., 2013). Conventionally, homology modelling requires a high degree of sequence similarity; however, the low number of amyloid proteins deposited within the PDB and the diversity of amyloid forming proteins makes this challenging. Despite this problem it has been shown that amyloid proteins with low sequence identity, such as IAPP and A $\beta$ , form fibrils that share many structural similarities (Figure 5.2).

The structure of medin is unknown; therefore, a molecular modelling approach was used to identify key features that could be tested experimentally. There are molecular fibril models available for IAPP and A $\beta$ <sub>1-40</sub> either in the PDB or in the literature. Sequence alignment between medin and IAPP and A $\beta$ , revealed closer similarity with A $\beta$  with both a global and local alignment, therefore A $\beta$  was chosen as the template for modelling medin fibrils.

Although the overall sequence alignment of medin with A $\beta$ <sub>1-40</sub> was low, a high level of similarity was identified between residues 25 and 30 of medin, which encompass

the turn region of A $\beta$ <sub>1-40</sub>. Modelling, suggests that it may be possible for ionic interactions to occur in medin if such a turn existed. Despite the lack of experimental evidence for a turn in medin, all amyloid structures composed of peptides over approximately 20 amino acids in length include some form of turn (Iwata et al., 2006, Petkova et al., 2005a, Luca et al., 2007). These models are limited by the relatively low sequence identity. Furthermore, the models represent a truncated form of medin due to the lack of template in these regions. Given these limitations, it was important to test this hypothesis experimentally. This took two forms, direct measurements using solid state NMR and mutagenesis studies.

#### **5.4.1. Solid state NMR measurements of medin**

Initial structural studies of recombinant medin fibrils demonstrate a high degree of heterogeneity within the sample. There is an increasing body of evidence to suggest that there is a large diversity of structural conformations within amyloid deposits (Pedersen et al., 2010). Multiple strains, have not only been seen for prion proteins (Tanaka et al., 2004) but, IAPP (Dzwolak et al., 2004),  $\beta_2$  microglobulin (Yamaguchi et al., 2005) and A $\beta$  (Petkova et al., 2005a). Certain strains or variants can be selected for through the process of seeding (Kim et al., 2007, Yamaguchi et al., 2005, Qiang et al., 2011). It has been shown numerous times that the presence of a 'structural template' in the form of a seed can propagate and select for a specific morphology or structure (Petkova et al., 2005a).

In previous studies on A $\beta$ <sub>1-40</sub>, the solid-state NMR line-widths for un-seeded fibrils were too large to obtain site specific information (Petkova et al., 2005a). In order to overcome the problem of sample heterogeneity, the fibrils were subjected to multiple rounds of seeding to generate a homogenous population of fibrils prior to structural

analysis. This improved the line-widths and led to site-specific information, which in turn led to molecular models (Petkova et al., 2005b). This could be a possible strategy to improve the line-widths and thus resolution of medin data however it was beyond the scope of this work. Despite the overlaps within the DARR spectrum comparison with simulated spectra revealed that it best resembled the simulated spectra for the hairpin model (Figure 5.13, 5.14 and 5.15).

An alternative method to obtain site specific information is to use a synthetic peptide to simplify the NMR experiments. Frequency selective REDOR experiments showed that the inter-nuclear distance between D25 C<sub>γ</sub> and K30 N<sub>ζ</sub> was less than 5 Å. It is therefore likely that medin fibrils are stabilised by a salt bridge between D25 and K30 as seen in Aβ (Petkova et al., 2002).

#### **5.4.2. Mutagenesis studies**

An additional means of testing the hypothesis of a salt bridge is to generate mutants that prevent the formation of a potential salt bridge between D25 and K30. Two mutants were generated in this work, D25N and K30I.

In order to investigate the role of these residues within medin aggregation, D25N and K30I were subject to kinetic, structural and morphological characterisation to observe any differences to wild type medin.

##### **5.4.2.1. Characterisation of D25N**

Kinetic and structural analysis of D25N indicated that it has a distinct aggregation pathway to wild type medin. Collectively the data suggest the formation of oligomeric species. ThT, IF and CD measurements indicate the presence of aggregates that are capable of binding ThT, are likely to be soluble in solution and have a predominantly α-helical conformation that undergoes a slight shift towards β-

sheet. Previous studies have described  $\alpha$ -helical oligomeric species that are capable of later converting to amyloid fibrils (Zhao et al., 2012, Jayaraman et al., 2012). It is possible that the D25N mutation initially favours oligomer formation and there is a partial conversion to  $\beta$ -sheet later on, possibly coinciding with oligomer association. Electron micrographs of D25N following incubation do not show fibrillar structures, they did however, reveal the presence of small, spherical, concave structures similar in appearance to previously described oligomeric species (Figure 5.31) (Ahmed et al., 2010, Fändrich, 2012).

Oligomeric amyloid species are thought to have a key role in the pathology of the amyloidoses. It would be interesting to assess the toxicity of the species formed by D25N to see if they also possess cytotoxic capabilities. The Iowa mutant of A $\beta$  forms fibrils however, it is more toxic than WT A $\beta$  (Van Nostrand et al., 2001). If D25N is more prone to oligomer formation it may be that it too is more toxic than its WT counterpart.

#### **5.4.3. Characterisation of K30I**

ThT analysis, intrinsic fluorescence data, CD measurements are all consistent with rapid conversion to large aggregates (Figure 5.19, 5.22 and 5.27). TEM visualisation confirms the presence of a large quantity of long, slender fibrils (Figure 5.31). Generally, K30I was less stable than both D25N and WT medin, illustrated by the increase degradation in response to UV exposure (Figure 5.30). It has been shown previously that instability greatly accelerates fibril formation (Ionescu-Zanetti et al., 1999) and likely explains the increased number of fibrillar aggregates observed by TEM. It is probable that the increased hydrophobicity of the K30I mutant drives

fibrils formation. In summary, substituting K30 for isoleucine increases the overall aggregation potential of medin

Both D25N and K30I have different kinetic, structural and morphological properties when compared with WT medin. These results indicate that both D25 and K30 have a role in the aggregation of WT medin. Mutations have often been shown to alter protein structure, this is exemplified by the number of amyloid diseases arising from single point mutations such as hereditary forms of TTR (Ericzon et al., 2003) and rarer conditions such as gelsolin related amyloidosis (Solomon et al., 2012). The changes observed here may be no more than a result of altering the intrinsic properties of the protein. However, collectively, the similarity with A $\beta$ , the NMR data and the mutation characterisation may support the hypothesis of a stabilising salt bridge between these residues.

In conclusion, data presented here are consistent with the presence of a salt bridge within medin fibrils. This provides further evidence that ionic interactions could be a common stabilising feature of amyloid structures. This could provide a common therapeutic strategy for targeting amyloid assemblies.

## **6. Expression and purification of SMA for structural studies**

### **6.1. Introduction**

#### **6.1.1. Immunoglobulin light chain amyloidosis**

Immunoglobulin light chains are involved in a variety of protein deposition diseases, including AL amyloidosis. AL amyloidosis is the most common and severe form of systemic amyloidosis. Patients with AL amyloidosis have a very poor prognosis with a median survival time of 40 months following diagnosis (Wechalekar et al., 2008). Approximately 90 % of patients with AL amyloidosis have signs of pathological heart involvement, and approximately 50 % of patients are in heart failure at the time of diagnosis (Desport et al., 2012). In total, approximately 75 % of deaths in patients with AL amyloidosis are attributed to cardiac involvement, namely heart failure and arrhythmia (Desport et al., 2012). Furthermore, subsequent treatment options are very limited (Lachmann et al., 2003). This is in part due to the extremely high level of heterogeneity; no patient diagnosed to date has had the same proteins within deposits as another patient (Enqvist et al., 2009). It has been shown previously in many studies that single point mutations or truncations can convert native non-amyloidogenic proteins into pathogenic amyloid (Van Nostrand et al., 2001, Steward et al., 2008). In the case of AL amyloidosis, it is highly unlikely that a single position or truncation can be responsible due to the diversity of light chains found amongst patients. This has led researchers to explore model systems in an attempt to learn more about this devastating disease.

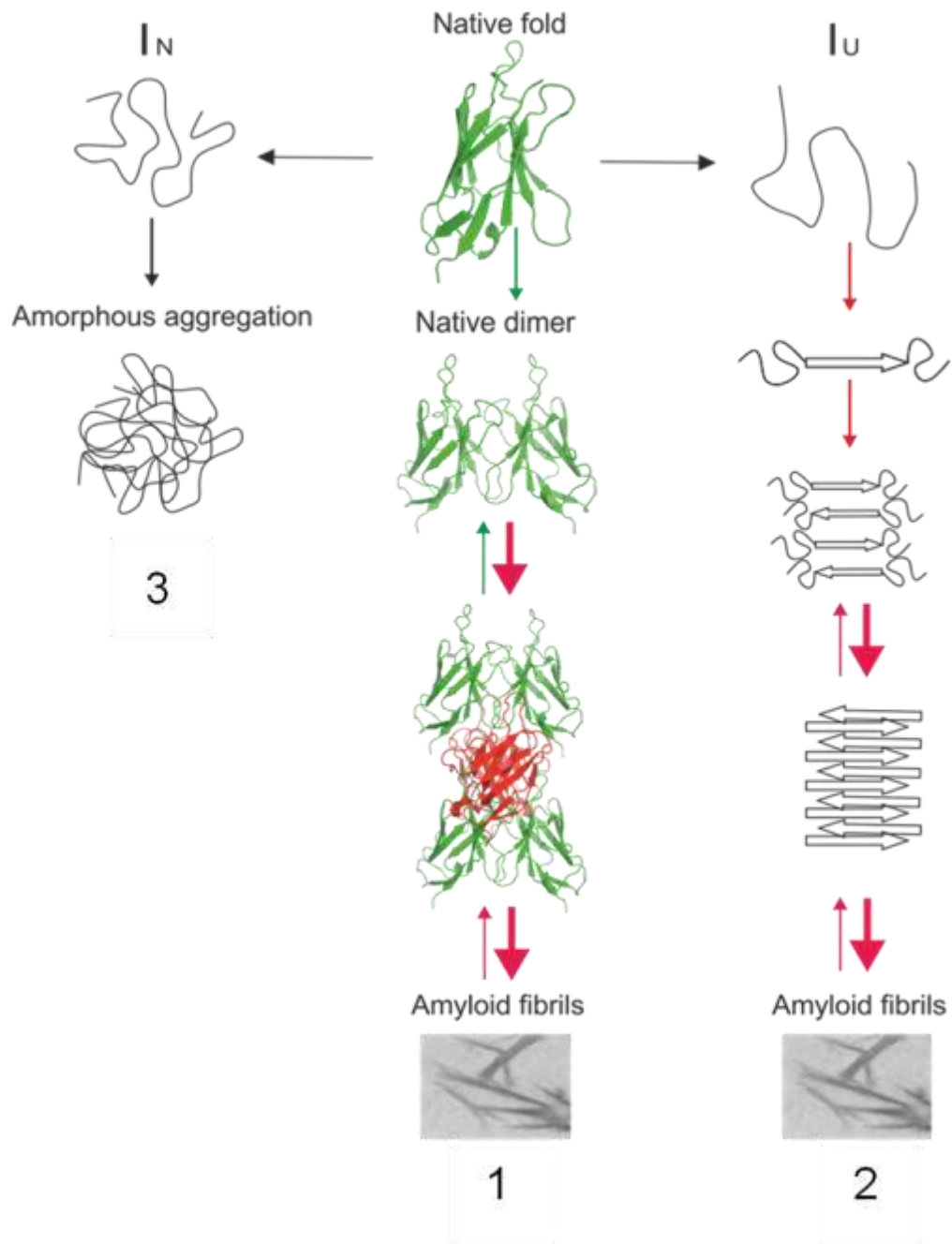
### **6.1.2. SMA as a model system**

As described in section 1.6.4.1, Stevens *et al* (1997) isolated three immunoglobulin light chains from patients which could act as a model system for AL amyloidosis. Two of these proteins, REC and SMA were extracted from the urine and lymph nodes respectively, from patients with amyloid deposits. The third protein, LEN was isolated from the urine of a patient with multiple myeloma with no apparent amyloid deposits. This allowed Stevens *et al.* and others to study the differences between these homologous proteins in an attempt to identify pathogenic features that could be targeted therapeutically.

### **6.1.3. Previous characterisation**

Various studies have investigated the biophysical and morphological properties of LEN, REC and in particular SMA (Khurana *et al.*, 2003). The aggregation characteristics of SMA have been studied over a wide range of pH values (Ionescu-Zanetti *et al.*, 1999). Results have shown that fibril formation is favoured by destabilising environmental conditions such as low pH ( pH 2) or in the presence of denaturants (Khurana *et al.*, 2001). This work led to the identification of two partially folded intermediate conformations,  $I_N$  and  $I_U$ .  $I_N$  is a relatively native-like conformation whereas  $I_U$  was less structured (Khurana *et al.*, 2001). Structural investigation into these intermediate species revealed that  $I_N$  led to amorphous aggregation (Route 1, Figure 6.1) whereas  $I_U$  readily formed fibrils (Route 2, Figure 6.1) (Qin *et al.*, 2007). The data collected on SMA was also compared with data for the non-amyloidogenic light chain LEN (Khurana *et al.*, 2003). This showed that LEN was more stable at all pH values and despite the presence of intermediate species, were very short lived and populated at much lower levels. It was, therefore,

proposed that the presence of long-lived, partially folded intermediate species, greatly contributed the amyloidogenicity of SMA.



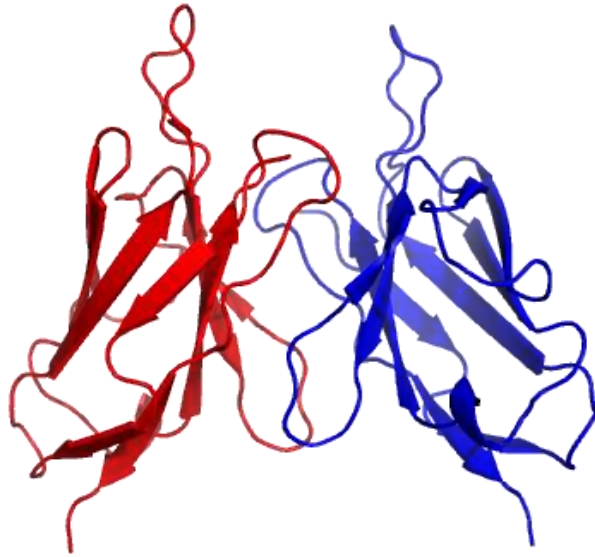
**Figure 6.1 Proposed models of SMA (and light chain in general) amyloid formation. The monomer can form a native-like partially folded intermediate that preferentially forms amorphous aggregates (3.) and a more unfolded intermediate that preferentially forms fibrils (2.) (Qin et al., 2007). A third model suggests that light chain dimers are able to stack with a 90 ° rotation, forming protofibrils and ultimately fibrils (1.) (Stevens et al., 1995a).**



Morphological analysis of SMA at a range of pH values studied by TEM, atomic force microscopy (AFM) and in the presence of different surfaces has shown that it is capable of forming a wide range of morphologies. At low pH large numbers of fibrils are formed rapidly, at slightly acidic pH amorphous aggregation predominates with very few fibrils observed and at physiological pH large amyloid-like fibrils are seen (Ionescu-Zanetti et al., 1999). Fibril growth has been shown to accelerate in the presence of surfaces (Zhu et al., 2002) or lipid membranes (Meng et al., 2008) however they did not seem to alter the morphology.

#### **6.1.4. Structural information**

There does not appear to be any link between the sequential differences between amyloidogenic and non-amyloidogenic proteins. The possibility of intermediate species and the range of observed morphologies emphasise the need for detailed structural information. In 1997, Huang *et al* solved the structure of the non-amyloidogenic LEN using X-ray crystallography (Figure 6.2) (Huang et al., 1997). It has a very similar structure to other variable light chain domains with a typical immunoglobulin fold, a 2-layer  $\beta$ -sandwich with between 7 and 9 anti-parallel  $\beta$ -strands (Figure 6.2) (Bork et al., 1994).



**Figure 6.2 Dimer structure of LEN (PDB ID: 1LVE) solved by X-ray crystallography. Figure adapted from Huang et al (1997) with each monomer individually coloured and visualised in PyMOL.**

LEN, like other immunoglobulin light chains forms a dimer. In their native states SMA and REC also form dimers; it is thought that increased dimer stability may contribute to amyloid formation (Stevens et al., 1995b). A model proposed by Stevens *et al* (1995), suggested that fibril formation is initiated by dimers stacking on top of one another, each rotated 90° around the fibril long axis (Route 2, Figure 6.1). Although the structure of LEN can be used to map differences in sequence to possible structural changes, there remains a need for detailed structural information of the pathogenic light chain aggregates.

Large fibrillar aggregates are unsuitable for crystallisation due to the inherent nature of the fibres and the lack of long range order. Furthermore, they are not amenable to study by solution state NMR due to their insoluble nature and size. The most appropriate technique is currently solid state NMR. This technique enables the study of large insoluble aggregates and can provide structural information about aggregate architecture. This in turn could reveal diagnostic information about the differences

between amyloidogenic and non-amyloidogenic light chains or guide therapeutic strategies.

The principal aim of the work described in this chapter is to develop an expression protocol for the production of isotopically ( $^{13}\text{C}$ ,  $^{15}\text{N}$ ) labelled SMA, which is an essential precursor to detailed solid-state NMR analysis. Methodology is described for the production of SMA aggregates and preliminary solid-state NMR are presented which demonstrate the feasibility of investigating the structure of these protein aggregates using a mixture of  $^{13}\text{C}$  and  $^{15}\text{N}$  isotope labels.

## **6.2. Materials and methods**

### **6.2.1. Vector – pKIV007**

The vector containing SMA was transformed into *E.coli* JM83 cells and supplied as an agar stab from Dr. Fred Stevens (Argonne national laboratory, U.S.A.) An overnight culture was inoculated using the stab and the plasmid extracted using a mini prep kit (Qiagen). The vector was then transformed into XL1 blue and BL21 (DE3) *E.coli* strains to build up vector stocks and carry out expression trails respectively.

The pKIV007 vector was designed in Dr. Fred Stevens' lab and is based on the p40ASK vector and contains ampicillin antibiotic resistance. It contains no affinity tags (Stevens et al., 1995b).

### **6.2.2. Expression and purification of pKIV007**

Expression trials and growth conditions for pKIV007 were the same as carried out for medin. In brief, the vector was transformed into *E.coli* BL21 (DE3) competent cells. A single colony was then selected and grown in 50 ml of Luria broth (LB) (Appendix 1) containing antibiotic overnight at 37 °C. This culture was then spun down at 3000 x g and used to inoculate 2 L of LB. The culture was incubated at 37 °C until an OD<sub>600</sub> = 0.8 was reached. Expression of the fusion protein was induced by the addition of IPTG to a final concentration of 1 mM. Cell growth was then allowed to continue 30 °C for 3 hours prior to harvest by centrifugation. All the samples were then lysed by sonication and then centrifuged at 19000 x g, the supernatant and pellet were collected for analysis by SDS-PAGE.

The resuspended cells were subjected to a single freeze-thaw cycle prior to 2 French press cycles in the presence of DNase at a final concentration of 20 ng/ml and a

cComplete EDTA-free protease inhibitor cocktail tablet (Roche applied sciences). The lysed cells were then centrifuged at 19000 x g to remove the cell debris and the supernatant filtered through a 0.22 µm Acrodisc<sup>TM</sup>. All purification columns were run on an AKTA protein purification system (GE Healthcare) at a flow rate of 2 ml/min unless otherwise stated. Samples were loaded via the peristaltic pump at a flow rate of 2 ml/min.

#### **6.2.2.1. Anion exchange chromatography**

In order to isolate SMA from the crude extract a combination of anion and cation exchange chromatography was performed.

A HighTrap Q FF column (GE Healthcare) was equilibrated with 3 CV of IE A buffer (Appendix 1), 3 CV of IE B buffer followed by a further 3 CV of IE A buffer. The high PI of SMA (8.9) prevented binding to the column and the flow through was therefore collected for further purification.

The flow through was then buffer exchanged using a HiPrep 26/10 desalting column (GE Healthcare). The column was equilibrated with 2 CV IE C buffer and the sample loaded onto the column. The desalted protein was collected in 2 ml fractions and then pooled prior to cation exchange.

Cation exchange chromatography was used to separate SMA from other proteins with a high PI. A HighTrap SP FF column (GE Healthcare) was equilibrated with 3 CV of IE C buffer (Appendix 1), 3 CV of IE D buffer followed by a further 3 CV of IE C buffer. The desalted protein was again loaded onto the column. The proteins were separated and eluted in three steps: 15 % gradient over 5 CV, a further 5 CV at 15 % followed by a gradient from 15-100 % over 10 CV.

### 6.2.2.2. Size exclusion chromatography

The final purification step was size exclusion chromatography. A Superdex 10/300 gel filtration column (GE Healthcare) was equilibrated with 3 CV of GF buffer. Separation was carried out at a flow rate of 1 ml/min and 2 ml fractions were collected.

### 6.2.3. Subcloning

The pKIV007 vector was subcloned into pOPINS in an attempt to improve the expression yield and facilitate purification. Subcloning was carried out according to Berrow *et. al* (2009) as described in section 3.2.1.6 using the following primers:

#### Forward SUMO primer

GCG AAC AGA TCG GTG CAG ACA TCG TGA TGA CCC AG

#### Reverse SUMO primer

ATG GTC TAG AAA GCT TTA TCA CCH TTT GAG TTC CAG TTT

### 6.2.4. Expression and purification of SUMO-SMA

As described in section 3.2.1.6 the pOPINS vector introduces a N-terminal his<sub>6</sub>-tag for affinity purification and a SUMO protein to aid solubility. A schematic of the SUMO-SMA fusion protein is shown in Figure 6.3.



**Figure 6.3 Schematic of the SUMO-SMA construct, red arrow indicates cleavage site.**

#### **6.2.4.1. Expression trials and conditions**

Expression trials of the pOPINS-SMA vector were carried out as described in section 3.2.1.8. The resultant growth conditions were the same as for pKIV007 except the induced culture was allowed to grow overnight prior to harvest instead of 3 hours.

#### **6.2.4.2. SDS PAGE analysis**

The expression and purification of SMA was analysed by SDS-PAGE using the Laemmli method (Laemmli, 1970). Analysis was carried out using 15% polyacrylamide gels in a Bio-Rad gel electrophoresis system. 5  $\mu$ l protein samples were added to 5  $\mu$ l of 2 x sample buffer, boiled and loaded onto the gel. All gels were run at 180 V for 60 minutes and then stained with Coomassie Brilliant Blue G-250 stain and de-stained with destain solution (Appendix 1).

#### **6.2.4.3. Tryptic digest and MALDI -MS**

Purified peptide (approx 2  $\mu$ g) was resuspended in 50 mM ammonium bicarbonate pH 7.8 containing 2 M urea and digested overnight with 0.1  $\mu$ g trypsin. The resulting peptides were cleaned up using C18 Zip Tips (Millipore) in preparation for analysis by either MALDI-MS or ES-MS. MALDI-MS analysis was performed using a MALDI-TOF instrument (Waters-Micromass) on samples dried on to the target plate using a saturated solution of  $\alpha$ -cyano-4 hydroxycinnaminic acid in 50% acetonitrile with 0.1% trifluoroacetic acid.

#### **6.2.4.4. Intact mass analysis using ES-MS**

For intact mass analysis, freshly prepared purified SMA sample was infused into the nano electrospray source of the mass spectrometer (Waters Q-ToF micro), at a flow

rate of 150  $\mu\text{l}/\text{hour}$ , via a gas tight syringe. The positive ion mass spectrum of the sample was recorded.

### 6.2.5. Characterisation of SMA aggregation under different buffer conditions

It has previously been reported that SMA aggregates differently under different pH conditions. This work investigated the aggregation of SMA at pH 2, pH 5.6 and pH 7.4, three previously described conditions detailed in Table 6.1.

**Table 6.1 Different aggregation conditions for SMA**

pH	Buffer components	Temperature ( $^{\circ}\text{C}$ )
2	20 mM HCl	25
5.6	50 mM sodium acetate, 100 mM NaCl	37
7.4	20 mM sodium phosphate, 150 mM NaCl	37

#### 6.2.5.1. Secondary structure analysis by circular dichroism

Changes in secondary structure of SMA at the end of the aggregation time course was assessed using far-UV synchrotron radiation circular dichroism (SRCD). Experiments were carried out at beam line B23 at Diamond light source (Oxford, U.K.). Freshly prepared SMA was incubated at a concentration of 40  $\mu\text{M}$  (as previously described (Zhu et al., 2002)), at the three different pH conditions. After 14 days the aggregate suspension was placed in the cuvette (121.000-QS Hellma, Southend-on-Sea, U.K.). Four scans were recorded between 180 and 260 nm, using a slit width of 0.5 mm, 0.5 cm path length and a scan rate of 1 nm/sec and averaged. The baseline was subtracted prior to secondary structure analysis. Analysis was carried out using Olis<sup>®</sup> GolbalWorks software. Data at each time point was subject to



two fitting methods, CONTILL and CDSSTR, using either basis sets 8 or 11 (Van Stokkum, 1990, Sreerama, 2000, Manavalan, 1987). The best fit, as determined by the normalised spectral fit standard deviation, was selected and the percentage  $\alpha$ -helix,  $\beta$ -sheet, turn and random coil content recorded.

#### **6.2.5.2. Morphology analysis by transmission electron microscopy**

Morphologies of the insoluble SMA aggregates formed under the three pH conditions were visualised using transmission electron microscopy (TEM). Peptide suspensions (10  $\mu$ L) were loaded onto carbon coated copper grids held between tweezers. Excess solution was removed by placing the grid perpendicular to blotting paper. The grid was then inverted onto a droplet of dH<sub>2</sub>O to wash off excess peptide aggregates. The grid was blotted again before being inverted onto a droplet of 4 % uranyl acetate to negatively stain the aggregates. Excess stain was removed by blotting before allow the grid to air dry prior to visualisation. The grids were observed using a Technai 10 electron microscope operating at 100 kV.

#### **6.2.6. NMR studies**

##### **6.2.6.1. Preparation of <sup>13</sup>C and <sup>15</sup>N isotopically labelled SMA mixture**

5 mg of <sup>13</sup>C and 10 mg of <sup>15</sup>N isotopically labelled SMA were mixed and incubated at a final concentration of 40  $\mu$ M, at 37 °C for 3 weeks with agitation. The resultant fibrils were harvested by centrifugation at 21 000 g and transferred to a 3.2 zirconium solid state NMR rotor with a boron nitrate cap (Bruker U.K.)

##### **6.2.6.2. Frequency selective REDOR experiment**

Frequency selective REDOR experiments were conducted as described in section 5.2.2.5.

### **6.2.7. SMA model**

Given the sequence similarities between medin and LEN, models of medin were generated using the structure of LEN derived from X-ray crystallography measurements (PDB ID: 2LVE). Modelling was carried out using Modeller V9.11 on full length SMA (Eswar et al., 2007).

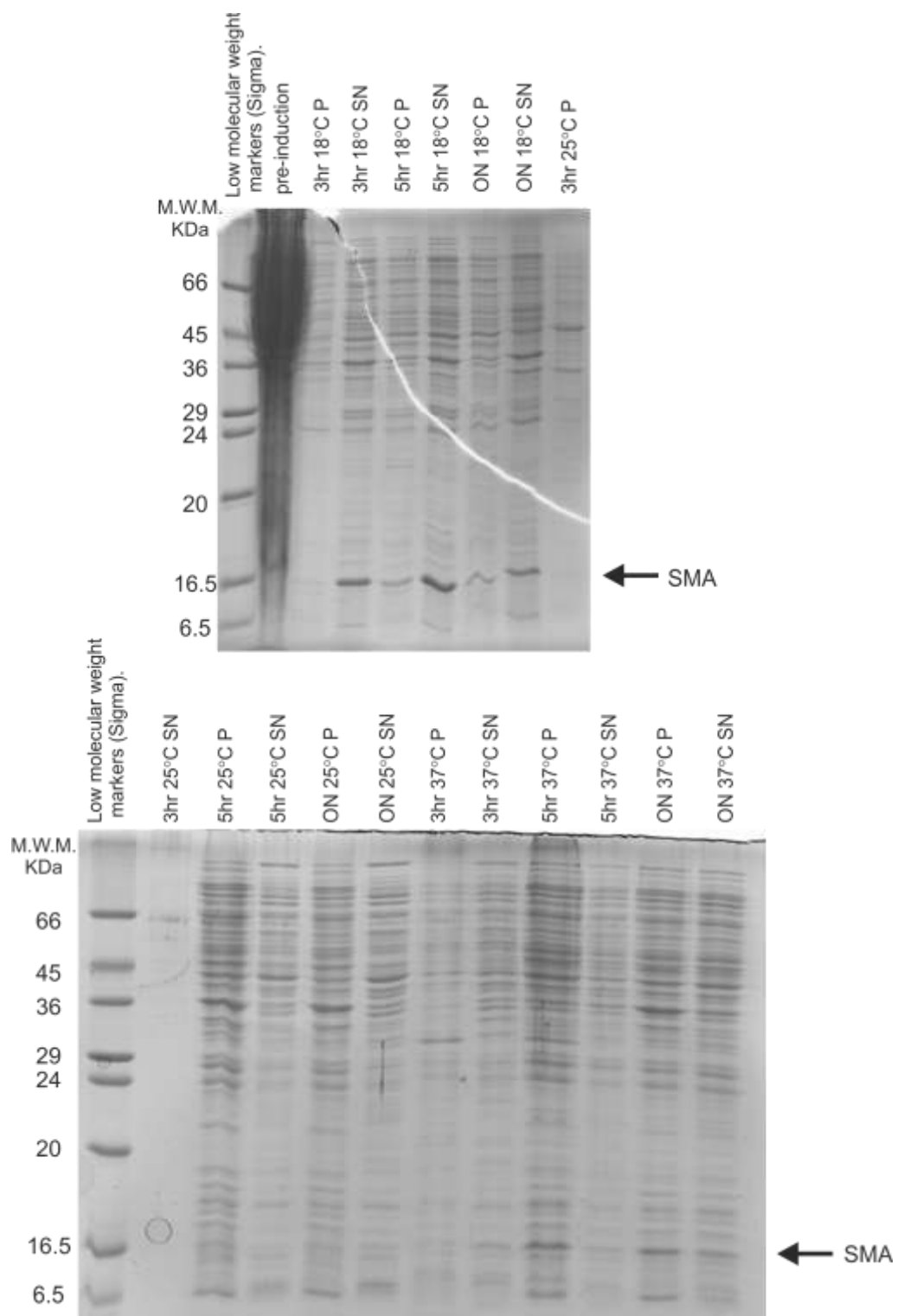
#### **6.2.7.1. Model validation and distance measurements**

An ensemble of 20 models was generated for each template and the best model of each ensemble selected based on the lowest molecular pdf score (molpdf), which is a standard MODELLER scoring function and is simply a sum of all the restraints (Eswar et al., 2007). WHAT IF scores were then calculated for each model and compared with those for the template models (Vriend, 1990). Chimera software was used to identify steric clashes within the models (Yang et al., 2012). Visualisation of models and distances between residues were calculated in PyMOL molecular graphics system version 1.5.0.4 (Schrödinger LLC.).

## **6.3. Results**

### **6.3.1. Expression and purification of pKIV007**

SMA was initially expressed using the clone and expression procedures described by Stevens et al (1997). Where possible, expression and purification conditions, were followed, although some changes were necessary to accommodate differences in laboratory equipment and *E.coli* strain availability. As described in section 3.2.2.4, the vector was provided as an agar stab in *E.coli* JM83 cells. These cells were not available to buy at the time of this work therefore expression tests were carried out in *E.coli* BL21 (DE3) which are used routinely for protein expression in this laboratory. If the protein had over-expressed, a large band would be expected at approximately 12 kDa. Overall tests showed that there was not a high level of over-expression and that the peptide ran slightly above the expected molecular weight upon SDS-PAGE analysis. However, the results indicated that the optimal expression conditions for SMA within this vector were 5 hours at 18 °C (Panel A, Figure 6.4). Due to the lack of affinity tag within this vector, SMA was purified from the crude extract using a combination of ion exchange and size exclusion chromatography.



**Figure 6.4 15 % SDS-PAGE of expression trials of SMA in the pKIV007 vector. Where P denoted pellet sample, SN denotes supernatant and ON means overnight.**

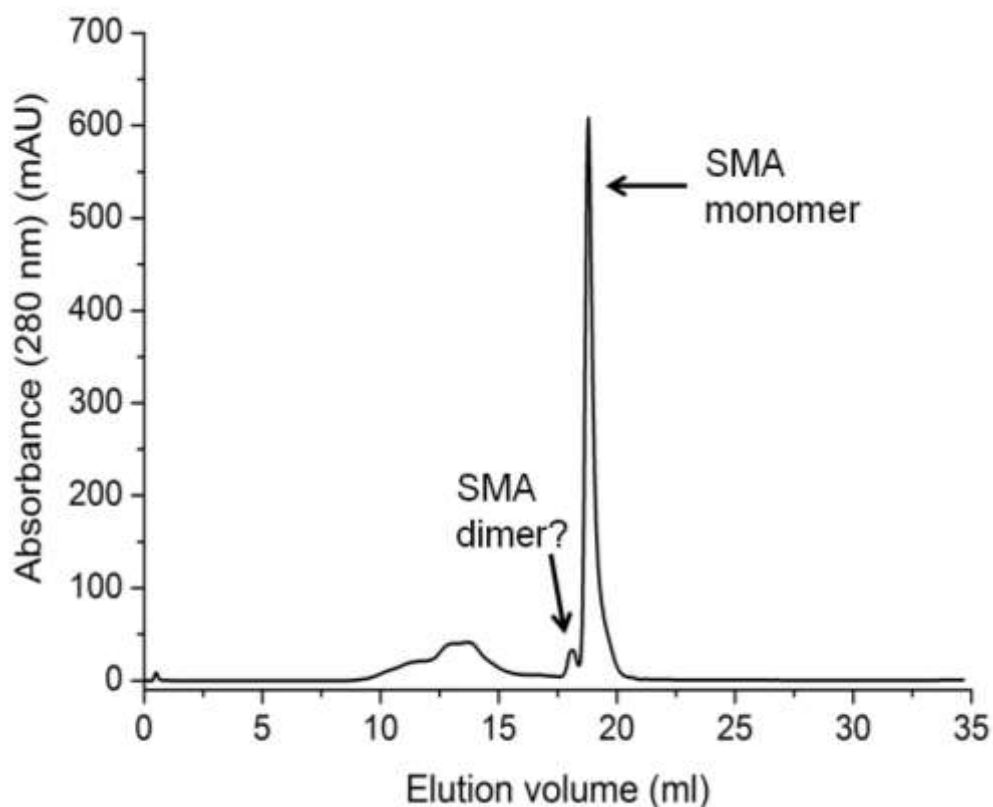
#### **6.3.1.1. Ion exchange chromatography purification**

Initially anion exchange chromatography was carried out pH 8. SMA has a pI of 8.5 and therefore carries an overall positive charge at pH 8 thus did not bind to the quaternary ammonium group on the resin of the Q FF column and eluted in the column flow through. This step quickly removed a large number of contaminating proteins.

The flow through was successfully desalted into IE C buffer (pH 5.6) (buffer composition given in Appendix 1) in preparation for cation exchange chromatography. At pH 5.6 SMA has an overall positive charge and therefore binds to the negatively charged sulphopropyl group of the resin. SMA was eluted during a gradient of NaCl (20 % over 20 CV). SDS-PAGE electrophoresis was performed to identify the fractions containing SMA. These were pooled and concentrated prior to size exclusion chromatography.

#### **6.3.1.2. Size exclusion chromatography purification**

Size exclusion chromatography was used to further purify SMA and ensure that there were no remaining contaminating proteins as a result of peak overlap. The chromatogram shows a large narrow peak eluting after approximately 19 ml consistent with a protein of similar size to SMA (Figure 6.5). There is a small shoulder that precedes the large peak which may represent a dimer species in line with the known dimer properties of immunoglobulin light chains.



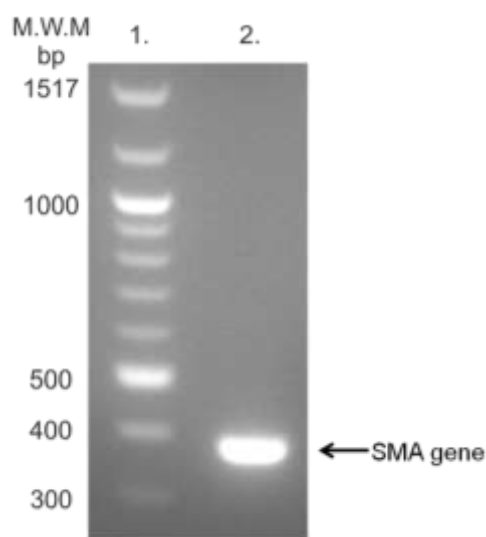
**Figure 6.5** Gel filtration trace chromatogram of S column fractions. Separation was carried out on a Superdex 75 10/300 gel filtration column (GE Healthcare) in 10 mM NaOAc, pH 5.6 at 2 ml/min.

The overall yield of SMA was very low, less than 1 mg/L of rich culture medium. The initial expression yield was low and the large number of purification steps resulted in a very low overall yield. The yield is likely to be much lower still when expressed in a minimal medium required for isotope labelling, the goal of this work. It was thought by sub-cloning into pOPINS might improve the expression yield and reduce the number of purification steps required, both by saving time and potential losses.

### **6.3.2. Sub-cloning of SMA into pOPINS vector**

The SMA sequence was successfully amplified using the vector specific primers and confirmed by agarose gel electrophoresis. The SMA gene is 357 bp in size and a

large band visible at the appropriate size (Lane 2. Figure 6.6). The pOPINS vector was prepared as described previously (section 3.2.1.6). The In-Fusion™ reaction was successfully carried out and a mixture of blue and white colonies was observed. Several white colonies were selected, grown and the plasmids purified and sent for sequencing. The sequencing data confirmed the insertion of the SMA gene into the pOPINS vector (Appendix 2) however there was an additional alanine present in at the N-terminus. Due to the location and nature of this residue it was decided not to sub-clone again and to continue work with this construct.



**Figure 6.6** Agarose gel of amplified SMA gene generated using the pOPINS specific primers. Lane 1. 100 bp DNA ladder (NEB), Lane 2, amplified SMA gene.

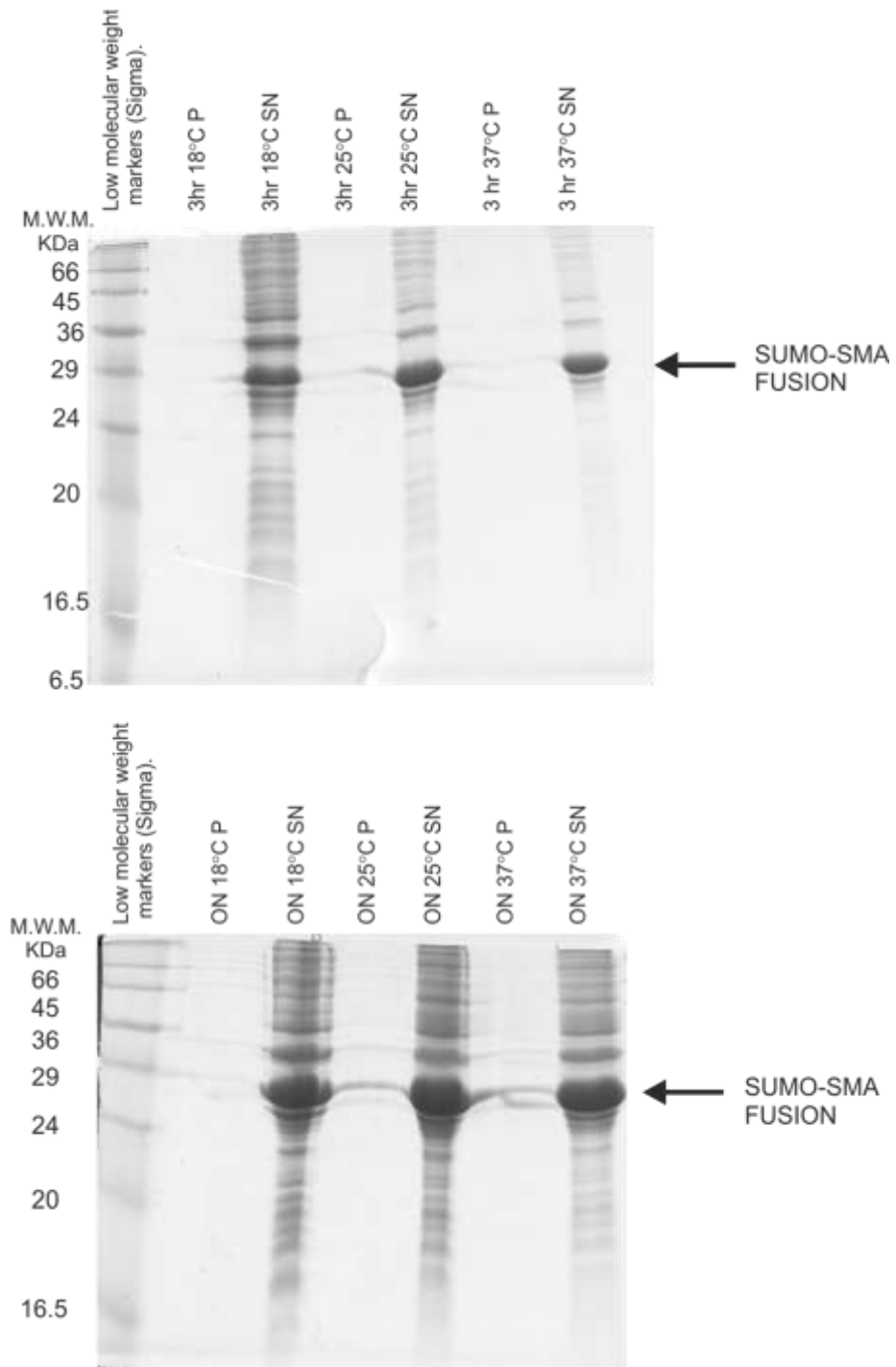
### **6.3.3. Expression and purification of SMA in pOPINS vector.**

#### **6.3.3.1. Expression tests**

The new vector was subject to expression tests to find the optimal expression conditions. The ultimate aim of this work was to develop a protocol for the expression and purification of isotopically labelled SMA in preparation for structural studies. Expression of isotopically labelled protein requires growth in minimal media

with single carbon and nitrogen sources. Consequently it was decided to carry out the expression tests in minimal media from the outset. Results in Figure 6.7, clearly show that the overall expression is substantially higher than that observed for the pKIV007 vector despite the change from enriched to minimal growth media. Optimal conditions were observed following an overnight incubation, although there is little difference between the three temperatures (Panel B, Figure 6.7). Overnight incubation at 30 °C was selected for future growth.

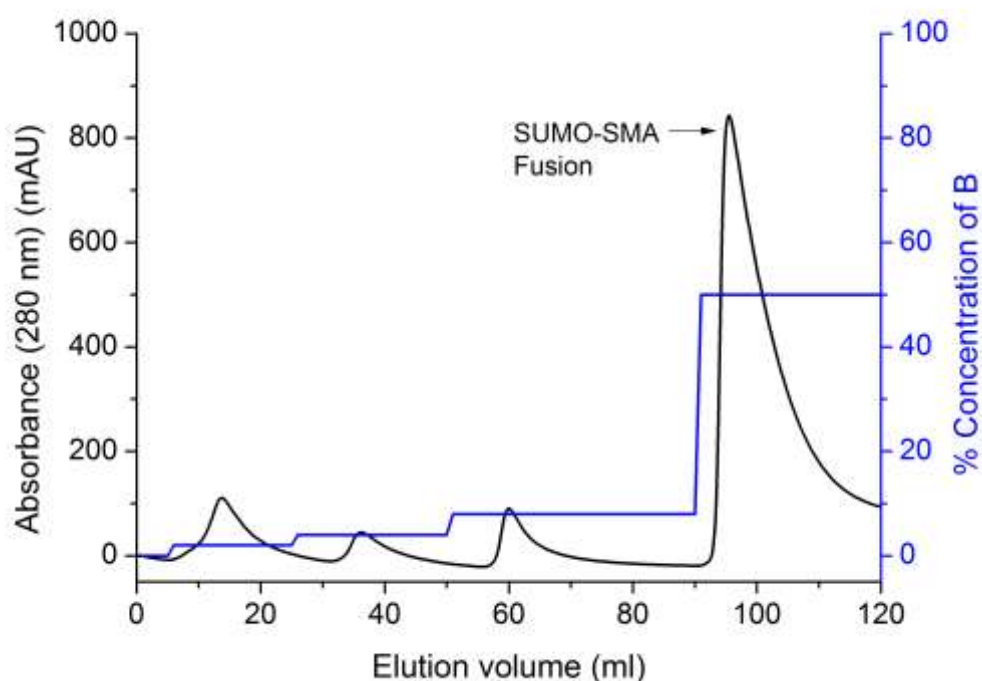




**Figure 6.7 15 % SDS-PAGE of expression trials of SMA in the pOPINS vector. Where P denoted pellet sample, SN denotes supernatant and ON describes overnight incubation.**

### 6.3.3.2. Purification of SMA

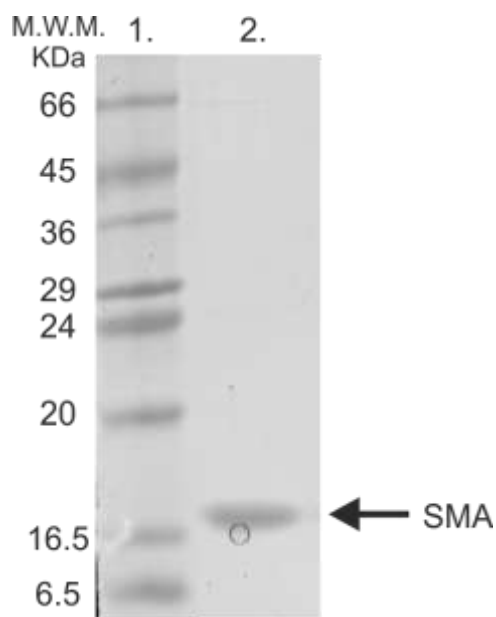
The overall purification strategy for SMA in the pOPINS vector mirrored the steps involved in purifying medin: affinity purification, a desalt step, cleavage of the tag and reverse affinity purification. Full details can be found in sections 3.2.2.4 and 3.2.2.5). Following lysis SMA was isolated from the crude extract using affinity purification.



**Figure 6.8** Affinity purification chromatogram of the fusion protein isolation. Sample was run on a 5 ml HisTrap™ FF column (GE Healthcare) in histrap buffer A. Step gradient of histrap buffer B shown in blue (Appendix 1).

This efficiently isolated SMA from the crude extract (Figure 6.8). Fractions from this stage were collected and pooled prior to desalting. SDS-PAGE analysis confirmed that the tag was then successfully cleaved from SMA using SUMO protease after 3 hours incubation at 25 °C. Cleaved SMA was then passed back through the His-

Trap™ column removing any residual contaminants, the protease and the tag (Lane 2, Figure 6.9).



**Figure 6.9 15 % SDS-PAGE analysis of purified SMA. Lane 1. Low molecular weight markers (Sigma Aldrich), Lane 2. Purified SMA.**

This proved to be a successful purification method for SMA and resulted in an improved yield when compared with the pKIV007 vector – approximately 5 mg/L of culture.

#### **6.3.4. Confirmation of SMA**

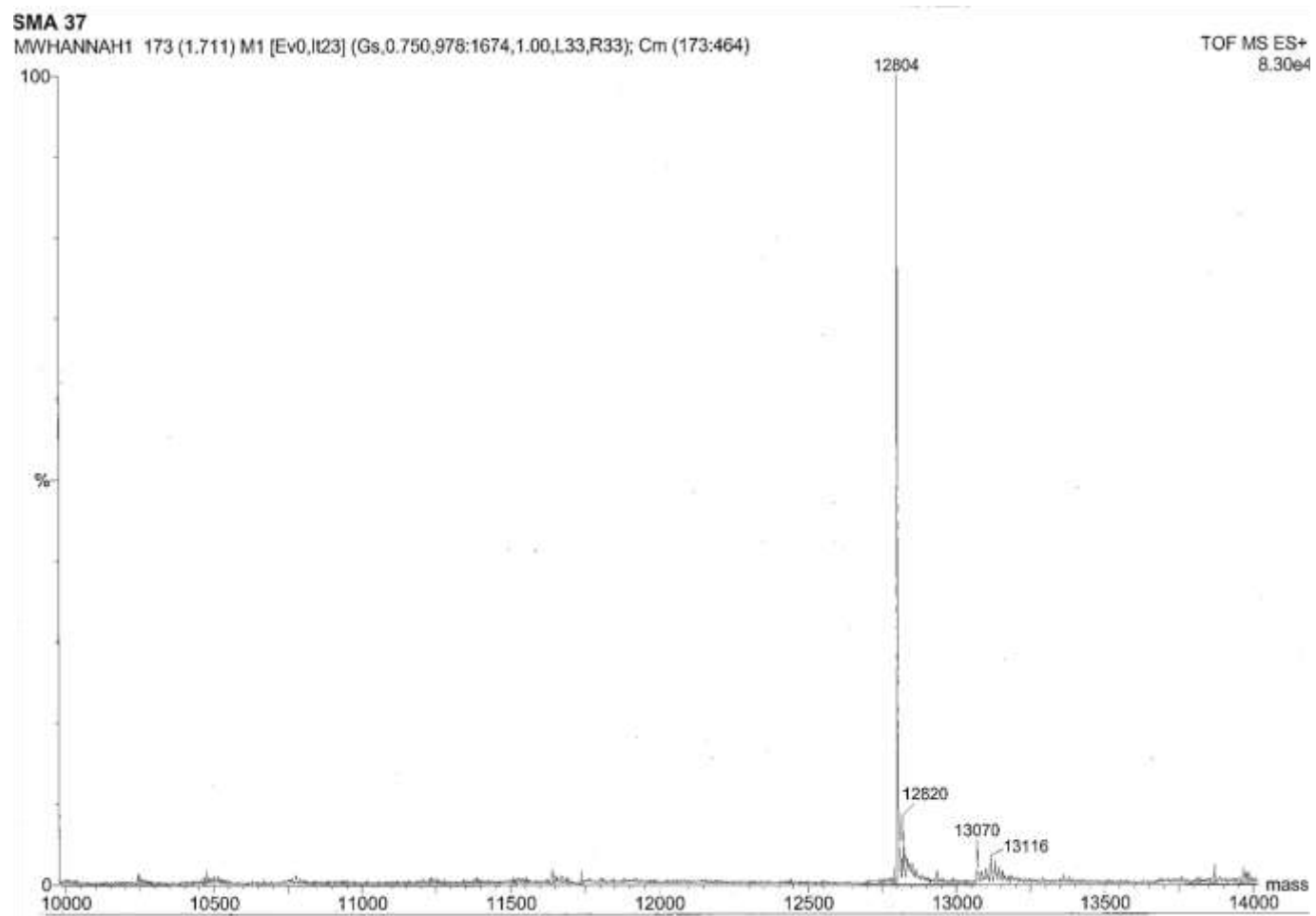
In order to confirm isolation of the correct protein both tryptic digest and intact mass spectrometry analysis were performed.

Tryptic digest fragments were subject to MALDI mass spectrometry to obtain mass data. Digest prediction was performed using MASTCOT peptide mass fingerprint (Perkins et al., 1999) and detected 9 possible fragments, 7 of which were obtained experimentally suggesting that the purified protein has the correct sequence for SMA (Table 6.2).

**Table 6.2 Tryptic digest fragments of SMA, detailing the sequence, expected mass from digest and actual masses obtained from MALDI.**

<b>Fragment sequence</b>	<b>Expected mass</b>	<b>MALDI</b>
ADIVMTQSPDSLAVSLGER	1988.9	1990
ATINCK	649.3	649.0
SSQSVLYSSNNR	1341.6	1341.7
NYLAWYQQK	1213.6	1213.6
LGQPPK	639.3	638.2
LLIYWASTR	1122.6	1122.7
ESGVPDR	759.3	759.0
FSGSGSGTDFLTISLQAEDVAVYYCHQYYSHPTFGQGTK	4578.0	N/A
LELK	502.3	-

MS-MS data for the intact protein shows a large peak at 12804 Da, the exact expected mass of SMA (+ alanine), confirming purification of the correct protein (Figure 6.10).



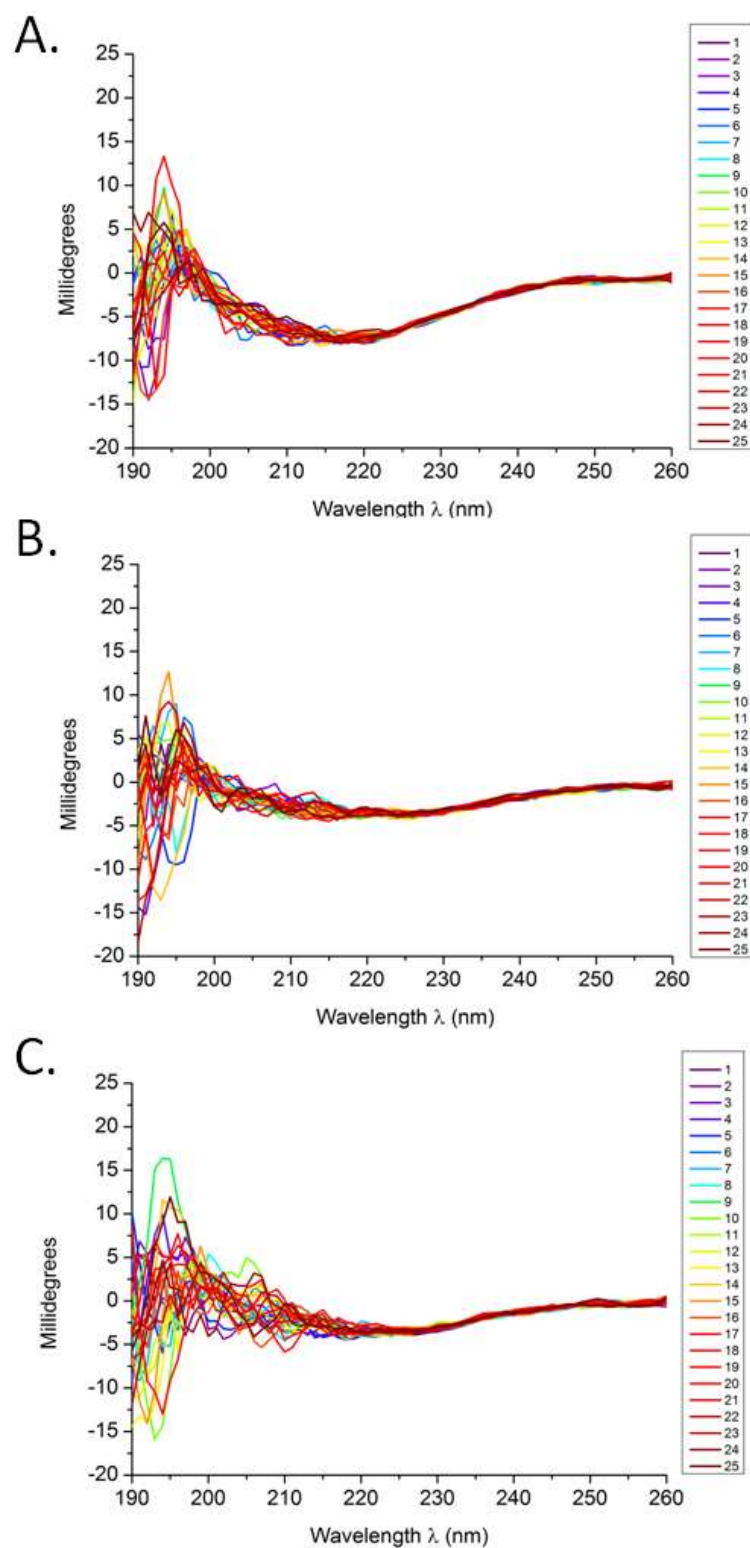
**Figure 6.10 MS-MS data of purified protein peak, showing a large peak at 12804 Da, consistent with the correct mass for SMA.**

### **6.3.5. Characterisation of SMA aggregation under different buffer conditions**

The aggregation of SMA has previously been studied at a wide range of pH values primarily using ThT and AFM (Ionescu-Zanetti et al., 1999, Meng et al., 2008). In order to ensure that the SMA produced in the pOPINS vector exhibited similar properties, it was characterised under three previously described conditions (Ionescu-Zanetti et al., 1999) detailed in Table 6.1. This was done with a view to identifying conditions to produce aggregates with defined morphology and in sufficient (mg) quantities for NMR analysis.

#### **6.3.5.1. Secondary structure and stability analysis by circular dichroism**

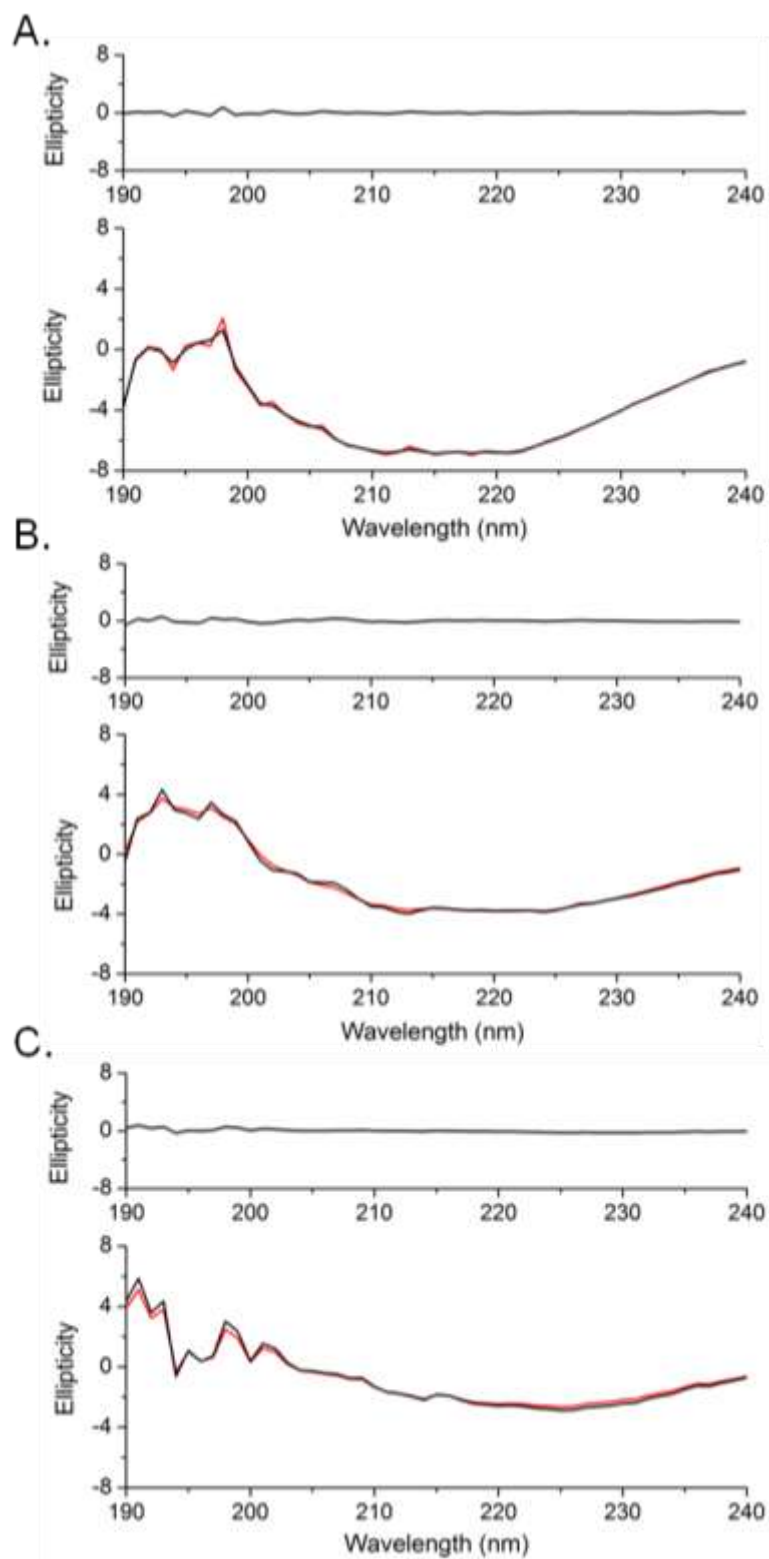
It has been proposed that the stability of early folding intermediates play a key role in determining the propensity of SMA, and immunoglobulin light chains in general, to self-assemble and the nature of the resultant aggregates (Qin et al., 2007). It was therefore decided to assess the stability of SMA in response to UV radiation under the three different buffer conditions. UV radiation can alter protein structure through ionising radiation. The primary photochemical process splits the water and the fragments can then attack the amide bond resulting in cleavage of the polypeptide chain (Johnson et al., 2011). Secondly, it has been suggested that UV radiation, can heat the internal water molecules within a protein resulting in conformational change and unfolding (Wallace, 2009). SMA was relatively unstable in response to UV radiation under all three conditions. At pH 7.4 SMA seemed the most unstable with a large amount of degeneration in the signal (Figure 6.12). UV is merely a model for destabilising conditions but it is interesting that SMA is the most unstable at physiologically relevant pH conditions.



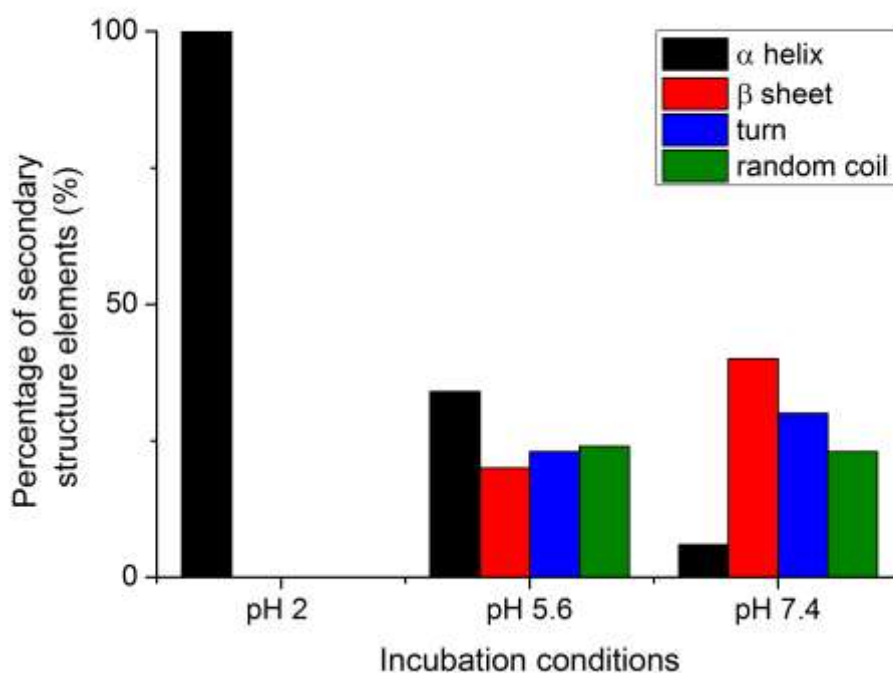
**Figure 6.11** Analysis of SMA stability at pH2 (A.), pH 5.6 (B.) and pH 7.4 (C.) in response to UV radiation. Data were recorded over 25 scans at 30 °C on freshly prepared SMA (40  $\mu$ M, in pH2 buffer) using far-UV synchrotron radiation circular dichroism at beam line B23 at Diamond light source. Numbers represent number of repeated exposures to UV beam.

CD analysis of the secondary structure of freshly prepared SMA at the three pH values showed that different incubation conditions altered the secondary structure content (Figure 6.12). At pH 2, secondary structure analysis software reported 100 %  $\alpha$ -helix (Figure 6.13). At pH 5.6 there was a mixture of secondary structure elements with a slight preference for  $\alpha$ -helix (approximately 50 %) (Figure 6.13). At pH 7.4, there is much less  $\alpha$ -helix but approximately 40 %  $\beta$ -sheet (Figure 6.13).





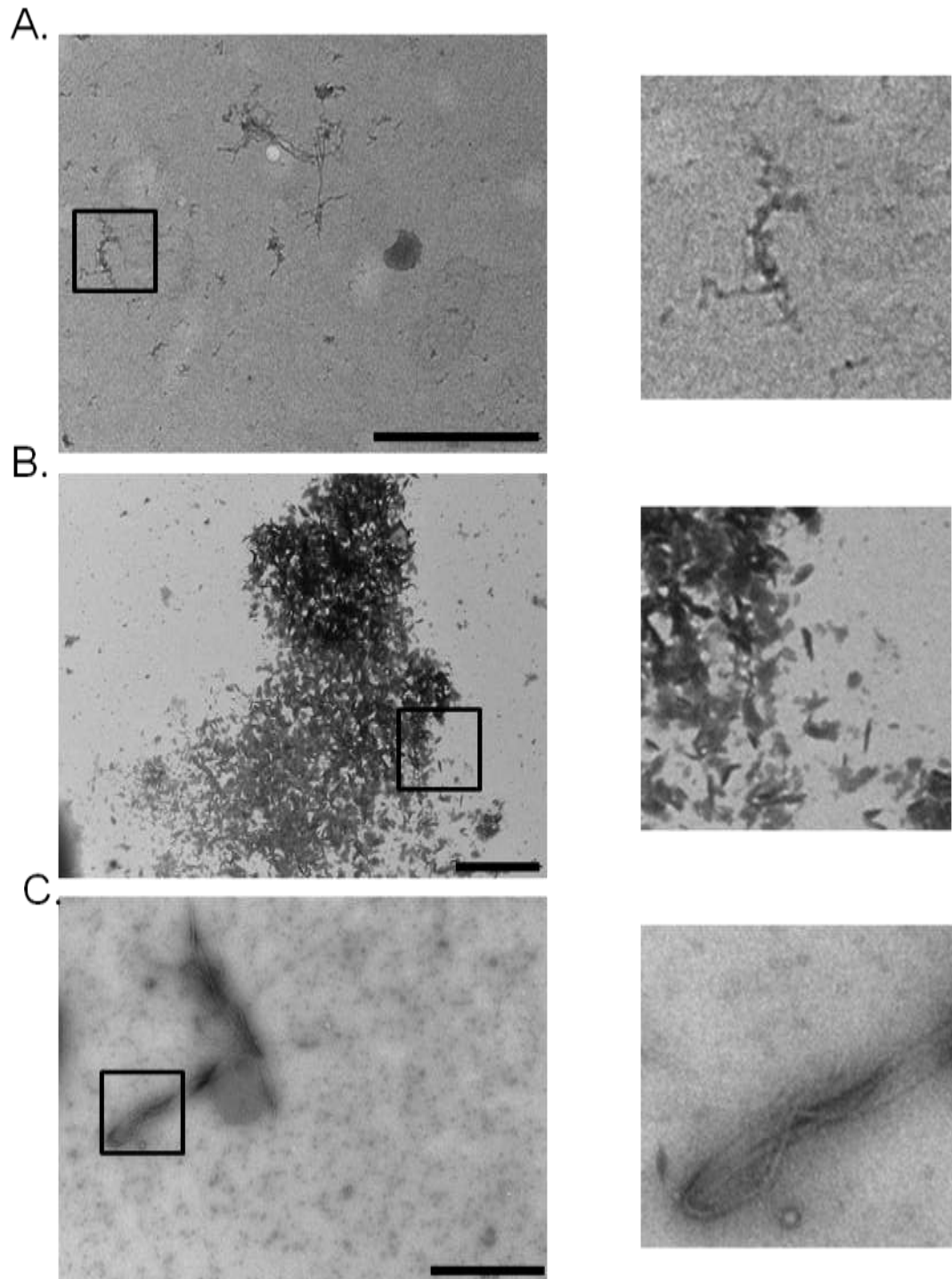
**Figure 6.12** CD spectra for freshly prepared SMA at pH 2, 5.6 and 7.4 (A, B and C respectively). Experimental data shown in black, fit in red, residuals are shown above.



**Figure 6.13 Secondary structure comparison of SMA under three different buffer conditions, calculated by CDSSTR or CONTILL algorithms.**

### **6.3.5.2. Morphology analysis by transmission electron microscopy**

Analysis of aggregates, formed from SMA following incubation at 30 °C with agitation, under three different pH conditions by transmission electron microscopy showed very different morphologies. At pH 2, a sparse deposition of small curly fibrillar species are seen similar to previously described protofibril species (Flach et al., 2012, Kumar et al., 2007) (Panel A, Figure 6.14). By contrast, a very large number of amorphous aggregates were formed a pH 5.6. These species were granular and quite uniform in size (Panel B, Figure 6.14) At pH 7.4 large twisted fibrils are formed with typical amyloid-like dimensions (Panel C, Figure 6.14). At pH 7.4 small uniform spherical aggregates are observed similar to oligomer species described in several studies (Fändrich, 2012).

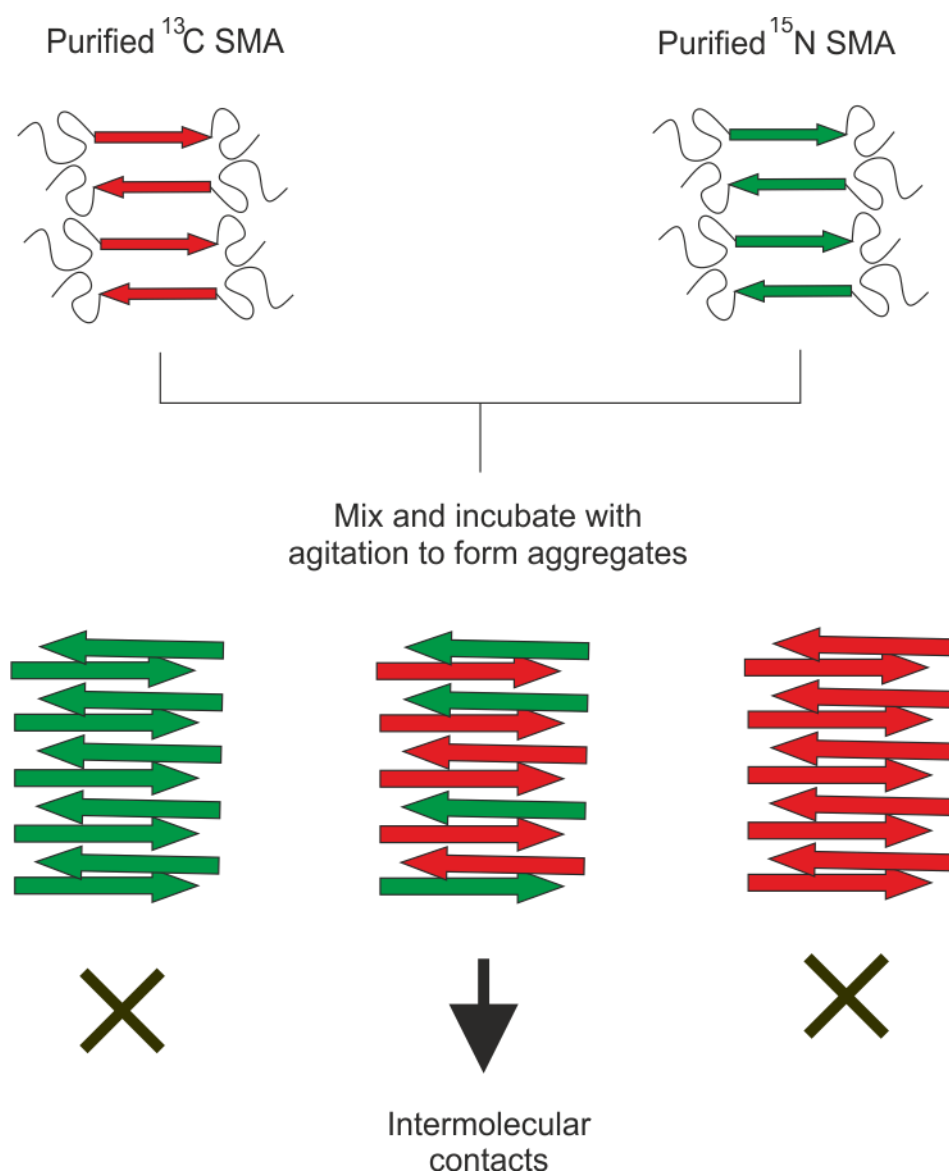


**Figure 6.14** Negatively stained electron microscopy images of SMA aggregates formed under different conditions: pH 2, pH 5.6 and pH 7.4 (A, B and C respectively). Higher magnification images of boxed areas shown on the right. All scale bars represent 1000 nm.

### 6.3.6. NMR experiments

The final stage of this work was to investigate the feasibility of obtaining solid-state NMR information from  $^{13}\text{C}$  and  $^{15}\text{N}$  labelled SMA aggregates. It was necessary to design a suitable NMR experiment for this purpose, taking into account the amount of material produced, the available access to a high-field spectrometer (necessary to obtain the highest sensitivity and resolution) and the remaining time and resources available for the project. With these considerations in mind, it was decided that aggregates produced at pH 5.6 should be examined, as this condition gave rise to the highest yield of material. In addition, the amount of access available on the 850 MHz National High Field Solid-state NMR spectrometer at the University of Warwick (7 days) limited the number and type of experiments that could be performed.

Consequently, it was decided that an experiment on aggregates formed from a mixture of uniformly  $^{15}\text{N}$  labelled SMA and uniformly  $^{13}\text{C}$  labelled SMA would be performed. By using an appropriate heteronuclear dipolar recoupling experiment (in this case frequency-selective REDOR), it would, in principle, be possible to detect intermolecular contacts between residues at the interfaces of protein monomers in the aggregates. Two batches of SMA, one with uniform  $^{13}\text{C}$  labels and a second with uniform  $^{15}\text{N}$  labels were successfully expressed and purified, then mixed and allowed to aggregate at pH 5.6 (Figure 6.15). In these aggregates 50 % of the interfacial sites between monomers correspond to interactions between a  $^{15}\text{N}$  labelled monomer and a  $^{13}\text{C}$  labelled monomer.

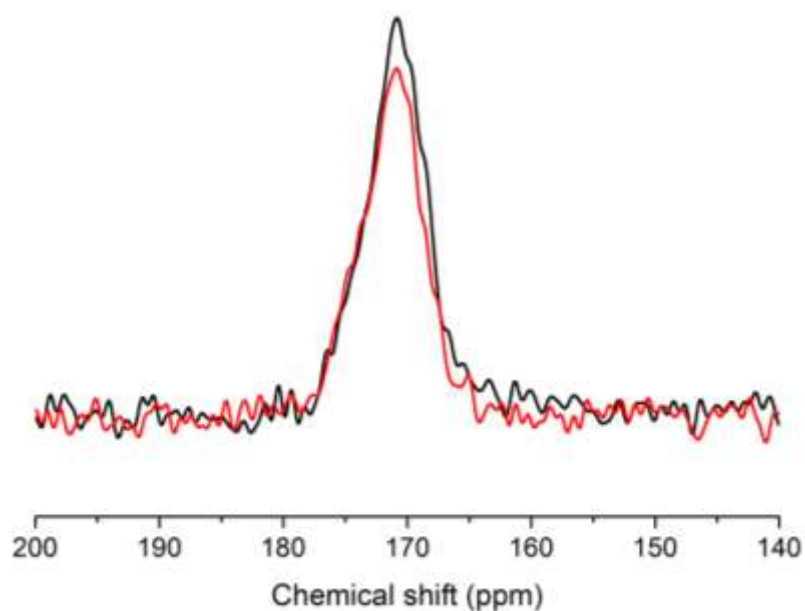


**Figure 6.15 Schematic illustrating the process of forming SMA aggregates with mixed isotope labels.**

Frequency selective REDOR experiments allow the selective measurement of specific inter-atomic distances from uniformly labelled solids. Two experiments are carried out: a  $^{13}\text{C}$  spectrum of the carbonyl region ( $\sim 170$  ppm) is observed after a dephasing period in which rotationally synchronous radiofrequency pulses are applied at a selective  $^{15}\text{N}$  frequency (i.e., in the backbone amide frequency range at around 120 ppm or at a specific side chain  $^{15}\text{N}$  frequency). A second, control  $^{13}\text{C}$

spectrum of the carbonyl region is obtained without pulses applied at the  $^{15}\text{N}$  frequency. If the first spectrum shows diminished intensity compared to the control spectrum, then this indicates coupling between the selective  $^{15}\text{N}$  sites and carbonyl  $^{13}\text{C}$  sites. For  $^{13}\text{C}$ - $^{15}\text{N}$ , such a dephasing is consistent with intermolecular  $^{13}\text{C}$ - $^{15}\text{N}$  distances of less than 4.5 Å.

Here the  $^{15}\text{N}$  dephasing frequency was set to recouple dipolar interactions between side-chain nitrogen atoms of lysine and arginine, and carbon atoms of all backbone and side chain carbonyl groups.



**Figure 6.16** Frequency selective REDOR experiment of mixed aggregates formed at pH 5.6 from singly labelled  $^{13}\text{C}$  and  $^{15}\text{N}$  SMA. There is a decrease in signal between spectra with (black) and without dephasing (red) on the  $^{15}\text{N}$  channel consistent with an intermolecular distance of less than 5 Å.

The REDOR spectra show that a small (approximately 10 %) dephasing occurs when the  $^{15}\text{N}$  pulses are applied (Figure 6.16) providing evidence of close inter-molecular contacts (< 5 Å)

This preliminary result demonstrates that it is possible to detect couplings between backbone and side chain carbonyl groups and K/R amino groups of SMA aggregates. Furthermore, the results provide initial constraints on the molecular organisation of SMA molecules within the aggregates.

## 6.4. Discussion

Initial expression and purification of SMA in the pKIV007 vector resulted in low protein yield and was therefore unsuitable for the production of large quantities of isotopically labelled protein required for ssNMR analysis.

Sub-cloning of SMA into the pOPINS vector led to increased protein yield and reduced purification time. This enabled the production of  $^{13}\text{C}$  and  $^{15}\text{N}$  isotopically labelled protein in preparation for ssNMR studies.

Secondary structure and morphological analysis was carried out on recombinantly produced SMA at three different pH values. Earlier studies on SMA and other amyloidogenic light chains have noted that aggregate formation occurs more rapidly upon destabilisation, including low pH. Here, this work investigated the effect of UV radiation on the stability of SMA at three different pHs. SMA was unstable under all three conditions and seemed the most unstable at pH 7.4. This is consistent with previous data on the stability of SMA. Khurana et al, (2003) used urea induced denaturation to demonstrate that SMA was unstable at a wide range of pH values (Khurana et al., 2003). Notably it was much more unstable than LEN, the non-amyloidogenic immunoglobulin light chain with only 8 residues different to SMA. This instability is likely to contribute to the amyloidogenicity of SMA.

Secondary structure analysis revealed differences in secondary structure content between the three conditions (Figure 6.12). This variation is broadly consistent with previous data (Figure 6.14). At pH 2, 100 %  $\alpha$ -helix was observed (Figure 6.13).

A mixture of secondary structure elements was recorded at pH 5.6 with a slight preference for  $\alpha$ -helix (Figure 6.13). At pH 7.4, there is a predominance of  $\beta$ -sheet,



consistent with previous reports and in line with the proposed structure of a  $\beta$ -sandwich (Figure 6.13) (Khurana et al., 2003)

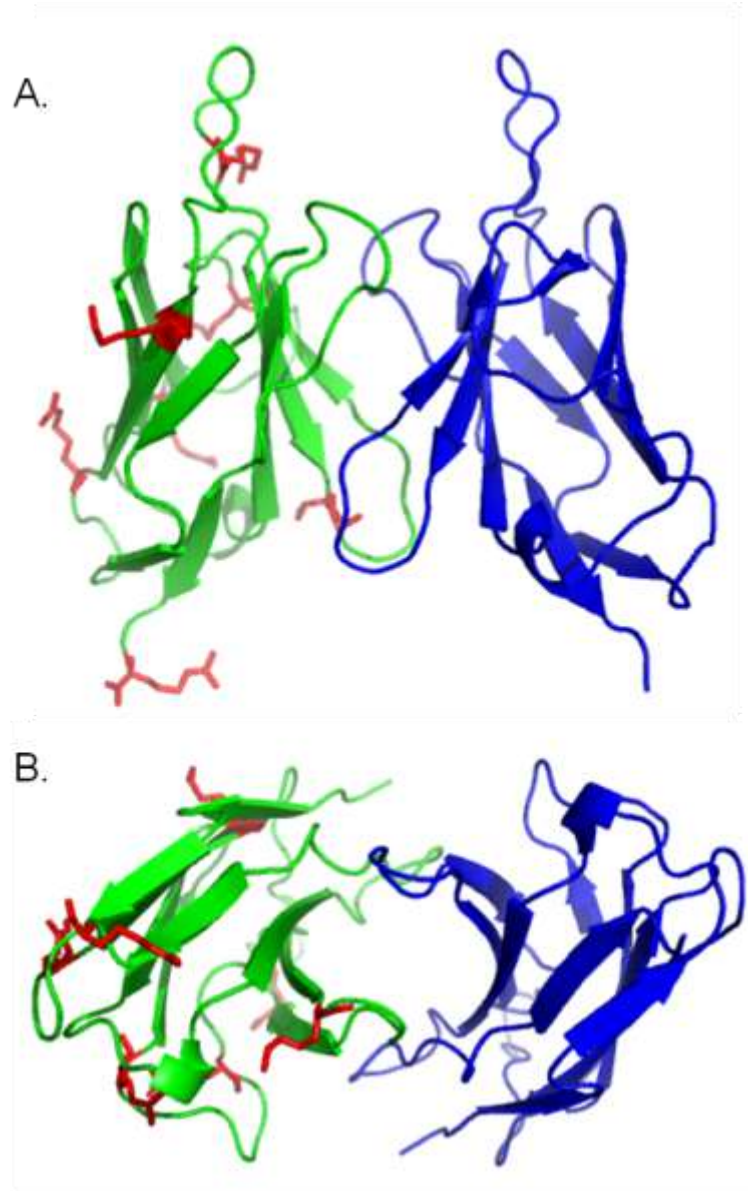
It must be noted that the CD data in general has a high signal to noise ratio which may produce artificial variation in the data and thus compromise the accuracy of the secondary structure fitting. However, this data is broadly consistent with other secondary structure studies on SMA.

At pH 2, small curly fibrillar species are seen which resemble previously described protofibrillar species (Panel A, Figure 6.14) (Lashuel et al., 2003). At pH 5.6 a large number of granular amorphous aggregates were formed. This data fits with previous descriptions of SMA aggregated at pH 5.6, where amorphous aggregation is dominant (Panel B, Figure 6.14) (Khurana et al., 2001). At pH 7.4 fibrils, similar in appearance to the 'braided' SMA fibrils described by Ionescu-Zanetti *et al* (1999) were formed (Panel C, Figure 6.14).

In summary, these data are consistent with previously described characteristics of SMA under these conditions. Experiments investigating the stability of SMA in response to UV radiation support the hypothesis that destabilisation of monomers is likely to contribute to amyloidogenicity.

Initial solid state NMR studies on SMA aggregates demonstrated that a mixed labelling scheme was effective for detecting inter-molecular contacts between different monomers. FS-REDOR experiments indicated the presence of close inter-atomic distances ( $< 5\text{\AA}$ ) between lys/arg amino groups and carbonyl groups (Figure 6.16). A model of an SMA dimer generated by homology modelling of LEN is shown in Figure 6.17, arginine and lysine side chains are shown as sticks in red. These side chains are too far away from any carbonyl groups in the second monomer (in blue) to enable detection in the FS-REDOR experiment (Figure 6.17). Previous

models regarding fibril formation of SMA have suggested that aggregate formation is initiated by dimer formation and the subsequent stacking of dimers, each rotated 90° around the fibril long axis (Figure 6.1) (Stevens et al., 1995a).



**Figure 6.17 Homology model of SMA dimer generated using Modeller, viewed from the side (A) and down the proposed fibril long axis (B). Modelled using the non-amyloidogenic light chain LEN (PDB ID: 2LVE) as a template. Lys and arg residues shown in red. Model visualised in PyMol.**

These data are not consistent with this model of aggregate formation suggesting a distinct pathway. NMR experiments were carried out pH 5.6 as this condition gave

rise to the highest aggregate yield and therefore it was hoped the best signal to noise ratio. This work demonstrates that it is possible to detect inter-molecular contacts between monomers.

Further structural analysis was beyond the scope of this work but the labelling scheme and subsequent FS-REDOR experiments could be used to investigate the structure of aggregates formed under a wide range of conditions. This could lead to site-specific information about the assembly of SMA monomers into the different aggregates and intermediates formed by this immunoglobulin light chain. In turn this information could be used to modulate aggregation in potential therapeutic strategies.

## 7. Discussion and future work

The overall aim of this thesis has been to study the amyloidogenic and structural properties of medin and SMA in order to enhance the understanding of cardiovascular amyloid and to design molecules for use as therapeutics or diagnostics.

### 7.1. Medin

Aortic medial amyloid is the most common form of localised amyloid recorded to date (Cornwell, 1982, Westermark and Westermark, 2011). At the inception of this project, little was known about the aggregation, structural and toxicological properties of medin when compared with the information available for other amyloidogenic peptides such as A $\beta$  and IAPP. It is thought that medin is excised from the precursor protein lactadherin by enzymatic cleavage but hitherto, there is no information available about how this process occurs, or about the protease responsible (Larsson et al., 2007). Furthermore, it is not known what affect medin deposition has on the cells and tissues within arteries. There is some evidence to suggest that medin may have a role in aortic aneurysm and dissection, through the weakening of the elastic structures within the arteries, however, this data is primarily circumstantial (Peng et al., 2007).

Previous biophysical studies on medin have identified eight residues at the C-terminus as being highly amyloidogenic in isolation and this region is thought to drive aggregation of full length medin (Larsson et al., 2007). Madine *et al* (2009) studied the biophysical properties of this peptide and suggested a structural model for the arrangement of this short peptide within fibrils (Madine et al., 2009a). Chapter 2 aimed to continue efforts to identify other key amyloidogenic regions (self-

recognition elements) of medin that could be targeted therapeutically. Targeting SRE of amyloid proteins has previously been shown to be a successful strategy for inhibiting amyloid proteins (Hughes et al., 2000, Madine et al., 2008a). Amyloidogenic prediction servers have been shown to identify effectively regions of amyloidogenic peptides that drive aggregation (Conchillo-Solé et al., 2007, Tartaglia and Vendruscolo, 2008). These servers highlighted three potential self recognition areas of medin. Consequently, three synthetic peptides were generated, incorporating one, two and three predicted regions (Med<sub>42-49</sub>, Med<sub>30-50</sub> and Med<sub>1-50</sub> respectively). Initial ThT analysis indicated that Med<sub>30-50</sub> and Med<sub>1-50</sub> both gave rise to a characteristic increase in fluorescence consistent with cross- $\beta$  sheet structure, and TEM analysis indicated that all three peptides had different morphologies.

Solid-state NMR is a very powerful tool to investigate the structure of amyloid assemblies (Heise, 2008, Tycko, 2006b). Here, ssNMR was used to probe the structural differences between aggregates derived from the three different peptides. Chemical shift information was consistent with a  $\beta$ -sheet conformation at all positions with the exception of F43 in Med<sub>42-49</sub>. Rotational resonance experiments investigated the packing arrangement of the assemblies; a close contact between F42 and V46 observed in Med<sub>42-49</sub> by Madine *et al* (2009) was not seen for Med<sub>30-50</sub> and Med<sub>1-50</sub>. Analysis of F43 ring dynamics indicated that F43 within Med<sub>42-49</sub> fibrils are rigid whereas it is increasingly flexible in Med<sub>30-50</sub> and Med<sub>1-50</sub>. There is ongoing debate about the role of aromatic residues in amyloid assemblies. Results from many different studies show that aromatic residues promote aggregation; early studies suggest this was mediated through the formation of  $\pi$ - $\pi$  stacking interactions (Gazit, 2002). However more recently there are a growing number of reports that

suggest hydrophobicity,  $\beta$ -sheet propensity, and planar geometry play a more important role (Berhanu and Hansmann, 2012a, Cukalevski et al., 2012, Doran et al., 2012b). Reches and Gazit (2004) showed that substituting phenylalanine at position 43 for either alanine or isoleucine in the octapeptide dramatically reduced its amyloidogenic potential (Reches and Gazit, 2004). In contrast, previous studies on full length medin substituted F43 and 48 for A in the full length peptide. This study reported that these substitutions did not affect the amyloidogenicity of medin (Larsson et al., 2007). The data presented in chapter 2 are consistent with both studies; F43 in Med<sub>42-49</sub> was rigid whereas in Med<sub>1-50</sub> it was much more flexible. In the context of these results, it may be that  $\pi$ - $\pi$  stacking does play a role within Med<sub>42-49</sub> in the absence of other driving forces. In the larger peptides, this interaction is probably superseded by a multitude of electrostatic and hydrophobic interactions at different points within the sequence.

In summary, these results suggest that the three peptides have distinct aggregate morphologies and at least two of the predicted self-recognition regions are required for fibril formation observed in the disease state. This data suggests that the prediction servers were able to accurately identify aggregation-prone regions and the use of multiple servers generates more accurate predictions.

This research has also emphasised that short peptides are often not good structural models of full length proteins, as it is apparent that Med<sub>42-49</sub> has different structural properties to Med<sub>1-50</sub>. Short peptides can often have different strand orientations and packing arrangements, for example A $\beta$ <sub>16-22</sub> has an anti-parallel sheet arrangement whereas A $\beta$ <sub>1-40</sub> has a parallel sheet arrangement (Balbach et al., 2000). This has been observed for several different amyloid proteins (Balbach et al., 2000, Goldsbury et

al., 2000, Jaroniec et al., 2002) and underlines the need to study the full length protein to obtain disease relevant results.

The three regions identified here, NH<sub>2</sub>-NAWVA-COOH, NH<sub>2</sub>-VTGII-COOH and NH<sub>2</sub>-FGSV-COOH, could now be used in the rational design of inhibitory peptides, as described previously for other amyloid proteins (Madine et al., 2008a, Hughes et al., 2000). This strategy aims to inhibit aggregation by preventing the association of subsequent monomers to the assembly through the incorporation of  $\beta$ -sheet breaking residues. Previous studies have designed inhibitors to the highly amyloidogenic C-terminal region of medin. One inhibitor in particular, F( $\beta$ A)S prevented the formation of large aggregates and demonstrated cyto-protective effect in toxicity studies (Madine and Middleton, 2010). This same peptide was tested alongside two others, Y( $\beta$ A)S and ( $\beta$ A)SV in chapter 4. There are several experimental differences between the work described and that by Madine *et al* (2010) namely the use of recombinant medin as opposed to synthetic medin, and continual monitoring by ThT, in place of discreet time point measurements. The preliminary results shown in section 4.3.4.1 suggest that F( $\beta$ A)S and ( $\beta$ A)SV had an effect on the aggregation kinetics of recombinant medin, significantly extending the lag time. They did not however prevent aggregation as seen by Madine *et al.* (2010). These results, and many previous studies, indicate that targeting the SRE of an amyloidogenic peptide with peptide inhibitors can be a successful strategy for therapeutic design. However, with regards to medin, further characterisation and optimisation of the current peptides is required.

The work in chapter 3 represents the culmination of many optimisation procedures to establish a reproducible expression and purification protocol for medin. Moreover,

this protocol has been adapted to produce isotopically labelled medin for NMR structural studies. Biophysical and structural characterisation of amyloid proteins generally requires relatively large quantities of protein. Furthermore, structural studies using NMR require the incorporation of isotope labels. A popular method to achieve this is the use of a bacterial expression host, most commonly *E.coli*. One of the many challenges of working with amyloidogenic proteins is the transient nature of monomeric species and subsequent formation of insoluble aggregates. Some studies have used this inherent insolubility as an advantage during purification (Finder et al., 2010), whilst others employ the use of solubility tags (Chen et al., 2012). This work used the latter to express and purify recombinant medin for the first time in sufficient quantities for subsequent structural and biophysical characterisation. Although the protocol can now be used for the production of medin, there are still some areas that could be optimised further, such as cleavage to improve yield and visualisation.

Previous studies on the characteristics of medin aggregation are limited, and have almost exclusively, been carried out on synthetically produced medin peptides (Madine et al., 2009a, Madine and Middleton, 2010, Gazit et al., 2007). Synthetically produced proteins have been shown to have different aggregation properties to their recombinant counterparts (Finder et al., 2010). Furthermore, they often require dissolution in organic solvents which do not reflect physiological conditions. The work described in chapter 4 is, to date, the most comprehensive *in vitro* characterisation of medin aggregation. Kinetic analysis of medin aggregation using ThT assay revealed that medin aggregates in a typical nucleation dependent manner, demonstrating a characteristic lag phase, followed by a rapid period of elongation and growth (Cabriolu et al., 2010). This approach uses continual monitoring



throughout the aggregation time course and can therefore give rise to accurate measurement of lag times, elongation rate and maximum fluorescence intensity. It is not possible to rule out any contribution from the presence of ThT but control experiments indicate that ThT fluorescence does not change over time in the absence of medin and it is currently a standard method for studying amyloid kinetics (Levine, 1999). Recombinant medin exhibits a lag phase of approximately 30 hours, supporting previous data obtained by Larsson *et al* (2007) who recorded a lag phase of 33 hours (Larsson *et al.*, 2007). This lag phase was reduced by the introduction of agitation, in agreement with previous observations of other amyloid proteins (Nielsen *et al.*, 2001, Serio *et al.*, 2000, Collins *et al.*, 2004). This is significantly quicker than the data obtained for synthetic medin which did not reach equilibrium after 21 days of incubation.

In the wider context of amyloid aggregation, medin aggregation is relatively rapid although it has a longer lag time than similarly sized amyloid peptides A $\beta$  and IAPP. Direct comparisons are difficult to make as there is such large range of aggregation conditions used in the literature. These include differences between recombinant and synthetic peptides (Finder *et al.*, 2010), seeded and unseeded growth (Kim *et al.*, 2007) denaturing conditions (Platt *et al.*, 2008) and whether proteins are agitated or not (Sasahara *et al.*, 2008).

In addition to the wide variety of conditions used to characterise amyloid proteins there exists the complication of co-localising factors. Amyloid deposits within the body are invariably found associated with a variety of additional factors which can include other proteins, GAGs and proteoglycans (Papy-Garcia *et al.*, 2011). In order to better understand disease processes the role of these additional components needs to be taken into consideration. The characterisation of medin aggregation shown in

chapter 4, under relatively physiological conditions provides a platform for future work into the effect of co-localising factors and toxicity studies.

Many studies within the amyloid field focus on the identification of structural or sequential amyloidogenic motifs used in therapeutics and diagnostics as discussed previously. As discussed in detail in chapter 5, medin shares some remarkable sequential similarities with A $\beta$ , particularly within the turn region of A $\beta$ , stabilised by a salt bridge. It was therefore hypothesised, that medin may also contain a stabilising salt bridge. Mutagenesis studies showed that residues D25 and K30 have an important role in the aggregation of medin as their substitution alters the kinetics and morphology of medin. TEM analysis of D25N resulted in the formation of oligomer-like species whereas K30I produced an increase in the number of fibrils. There are many reports on the toxic potential of oligomeric species and it would be critical to assess the toxicity of WT, D25N and K30I aggregates.

In order to investigate the possibility of a salt bridge further, solid state NMR experiments were used to measure the inter-nuclear distance between D25 and K30. This was initially attempted with uniformly  $^{13}\text{C}/^{15}\text{N}$  double labelled recombinant medin but the fibrils were too heterogeneous resulting in insufficient resolution to identify the resonances for D25 and K30. As discussed in chapter 6, many amyloid proteins have been shown to form polymorphic fibrils (Paravastu et al., 2006, Madine et al., 2008c). Distinct structural morphologies of some of these proteins have been documented, predominantly for A $\beta$  (Petkova et al., 2002, Paravastu et al., 2008). In order to study these morphologies in isolation, repetitive seeded growth was required (Petkova et al., 2005a). Previous studies have required between 12 and 15 generations of seeded growth to obtain a homogenous morphology (Paravastu et

al., 2008). Seeding medin in a similar way was beyond the possibilities of this work due to time constraints and availability of high field NMR instrumentation, but is a key consideration for future studies. Alternatively, a synthetically prepared medin sample with selective isotope labels was prepared to address this aim. Fibrils derived from synthetically prepared medin were investigated and revealed an inter-nuclear distance between D25 and K30 of less than 5 Å consistent with salt bridge formation. As highlighted in the general introduction, salt bridges occur in several other amyloid assemblies (Petkova et al., 2002) and may be a key interaction to target therapeutically. However as seen with the salt bridge mutations, one must be cautious as with all therapeutics, not to generate a more toxic alternative. It is generally accepted that oligomeric species are more toxic than mature fibrils (Kayed et al., 2003). Targeting salt bridge interactions may destabilise fibril formation but promote oligomer formation, which could result in a more severe pathology.

The remarkable similarity between medin and A $\beta$  may indicate that there are common features or motifs within amyloid proteins. Medin has also been shown to share similarity with IAPP (Reches and Gazit, 2004). Likewise, there is also data that investigate the similarities between A $\beta$  and IAPP. There is evidence that A $\beta$  can activate IAPP receptors (Fu et al., 2012) and they share structural and sequential similarities (Luca et al., 2007). Further analysis of these common regions and comparison with additional amyloid proteins could assist in the design of common therapeutics and diagnostics.

## **7.2. SMA**

Information from the optimisation of medin production was subsequently used to optimise the expression and purification of a second cardiovascular amyloid protein

– SMA. AL amyloidosis is a rare yet devastating disease. One of the greatest research challenges in the study of AL amyloidosis is to elucidate the mechanism that causes such a diverse range of immunoglobulin light chains to form fibrillar aggregates. There is a high degree of sequence homology between these amyloidogenic light chains and it is possible therefore that structural differences govern their amyloidogenicity. Due to the variety of amyloidogenic light chain fragments a model protein SMA is often used in the study of AL amyloidosis. Chapter 7 discusses the production of isotopically labelled SMA in preparation of structural studies by solid state NMR. Furthermore, SSNMR experiments demonstrate that isotopically labelled proteins coupled with high field selective NMR experiments can provide site-specific information about the structure of light chain aggregates. This work acts as a proof of principle and can be used in the study of many different amyloidogenic light chains and other amyloid proteins in general.

The importance of establishing a recombinant expression protocol for amyloid proteins is not limited to economic and efficiency factors. Studies concerning other amyloid proteins such as A $\beta$  have shown that synthetic and recombinant proteins can have different aggregation properties (Finder et al., 2010).

Overall, this work presents the first comprehensive expression and purification protocol for the production of medin. This has enabled the characterisation of its aggregation under relatively physiological conditions which can be used to assess the role of additional co-localising factors or aggregation inhibitors in future studies. This work has also highlighted structural similarities between medin and A $\beta$ , opening up the debate about common structural motifs/interactions, their role in the aggregation process and implications for therapeutic design and diagnostics. In

conclusion this work has enhanced our understanding of medin aggregation and provided a platform for future studies into both medin and SMA.

## References

- ADACHI, E., NAKAJIMA, H., MIZUGUCHI, C., DHANASEKARAN, P., KAWASHIMA, H., NAGAO, K., AKAJI, K., LUND-KATZ, S., PHILLIPS, M. C. & SAITO, H. 2013. Dual Role of an N-terminal Amyloidogenic Mutation in Apolipoprotein A-I: DESTABILIZATION OF HELIX BUNDLE AND ENHANCEMENT OF FIBRIL FORMATION. *Journal of Biological Chemistry*, 288, 2848-2856.
- AHMED, M., DAVIS, J., AUCOIN, D., SATO, T., AHUJA, S., AIMOTO, S., ELLIOTT, J. I., VAN NOSTRAND, W. E. & SMITH, S. O. 2010. Structural conversion of neurotoxic amyloid-[beta]1-42 oligomers to fibrils. *Nat Struct Mol Biol*, 17, 561-567.
- ALVAREZ-MARTINEZ, M. T., FONTES, P., ZOMOSA-SIGNORET, V., ARNAUD, J. D., HINGANT, E., PUJO-MENJOUET, L. & LIAUTARD, J. P. 2011. Dynamics of polymerization shed light on the mechanisms that lead to multiple amyloid structures of the prion protein. *Biochimica Et Biophysica Acta-Proteins and Proteomics*, 1814, 1305-1317.
- AMARO, M., BIRCH, D. J. S. & ROLINSKI, O. J. 2011. Beta-amyloid oligomerisation monitored by intrinsic tyrosine fluorescence. *Physical Chemistry Chemical Physics*, 13, 6434-6441.
- AMIJEE, H., BATE, C., WILLIAMS, A., VIRDEE, J., JEGGO, R., SPANSWICK, D., SCOPES, D. I. C., TREHERNE, J. M., MAZZITELLI, S., CHAWNER, R., EYERS, C. E. & DOIG, A. J. 2012. The N-Methylated Peptide SEN304 Powerfully Inhibits A beta(1-42) Toxicity by Perturbing Oligomer Formation. *Biochemistry*, 51, 8338-8352.
- AMIJEE, H., MADINE, J., MIDDLETON, D. A. & DOIG, A. J. 2009. Inhibitors of protein aggregation and toxicity. *Biochemical Society Transactions*, 37, 692-696.
- ANDERSEN, M. H., GRAVERSEN, H., FEDOSOV, S. N., PETERSEN, T. E. & RASMUSSEN, J. T. 2000. Functional Analyses of Two Cellular Binding Domains of Bovine Lactadherin†. *Biochemistry*, 39, 6200-6206.
- ANDREWS, L. J. & FORSTER, L. S. 1972. Protein difference spectra. Effect of solvent and charge on tryptophan. *Biochemistry*, 11, 1875-1879.
- ARIGA, T., MIYATAKE, T. & YU, R. K. 2010. Role of Proteoglycans and Glycosaminoglycans in the Pathogenesis of Alzheimer's Disease and Related Disorders: Amyloidogenesis and Therapeutic Strategies-A Review. *Journal of Neuroscience Research*, 88, 2303-2315.
- BALBACH, J. J., ISHII, Y., ANTZUTKIN, O. N., LEAPMAN, R. D., RIZZO, N. W., DYDA, F., REED, J. & TYCKO, R. 2000. Amyloid Fibril Formation by Aβ16-22, a Seven-Residue Fragment of the Alzheimer's β-Amyloid Peptide, and Structural Characterization by Solid State NMR†. *Biochemistry*, 39, 13748-13759.
- BALBACH, J. J., PETKOVA, A. T., OYLER, N. A., ANTZUKIN, O. N., GORDON, D. J., MEREDITH, S. C., TYCKO, R. 2002. Supramolecular structure in full-length Alzheimer's beta-amyloid fibrils: evidence for a parallel beta-sheet organization from solid-state nuclear magnetic resonance. *Biophys J.*, 83, 1205-1216.

- BANYPERSAD, S. M., MOON, J. C., WHELAN, C., HAWKINS, P. N. & WECHALEKAR, A. D. 2012. Updates in Cardiac Amyloidosis: A Review. *Journal of the American Heart Association*, 1.
- BEMPORAD, F., TADDEI, N., STEFANI, M. & CHITI, F. 2006. Assessing the role of aromatic residues in the amyloid aggregation of human muscle acylphosphatase. *Protein Science*, 15, 862-870.
- BENNETT, A. E., RIENSTRA, C. M., AUGER, M., LAKSHMI, K. V. & GRIFFIN, R. G. 1995. Heteronuclear decoupling in rotating solids. *The Journal of Chemical Physics*, 103, 6951-6958.
- BERHANU, W. M. & HANSMANN, U. H. E. 2012a. Side-chain hydrophobicity and the stability of A beta(16-22) aggregates. *Protein Science*, 21, 1837-1848.
- BERHANU, W. M. & HANSMANN, U. H. E. 2012b. Structure and Dynamics of Amyloid- $\beta$  Segmental Polymorphisms. *PLoS ONE*, 7, e41479.
- BERROW, N. S., ALDERTON, D. & OWENS, R. J. 2009. The Precise Engineering of Expression Vectors Using High-Throughput In-Fusion (TM) PCR Cloning. In: DOYLE, S. A. D. S. A. (ed.) *Methods in Molecular Biology*.
- BERROW, N. S., ALDERTON, D., SAINSBURY, S., NETTLESHIP, J., ASSENBERG, R., RAHMAN, N., STUART, D. I. & OWENS, R. J. 2007. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Research*, 35.
- BERSON, J. F., THEOS, A. C., HARPER, D. C., TENZA, D., RAPOSO, G. & MARKS, M. S. 2003. Proprotein convertase cleavage liberates a fibrillogenic fragment of a resident glycoprotein to initiate melanosome biogenesis. *The Journal of Cell Biology*, 161, 521-533.
- BERTHELOT, K., CULLIN, C. & LECOMTE, S. 2013. What does make an amyloid toxic: Morphology, structure or interaction with membrane? *Biochimie*, 95, 12-19.
- BODIN, K., ELLMERICH, S., KAHAN, M. C., TENNENT, G. A., LOESCH, A., GILBERTSON, J. A., HUTCHINSON, W. L., MANGIONE, P. P., GALLIMORE, J. R., MILLAR, D. J., MINOGUE, S., DHILLON, A. P., TAYLOR, G. W., BRADWELL, A. R., PETRIE, A., GILLMORE, J. D., BELLOTTI, V., BOTTO, M., HAWKINS, P. N. & PEPYS, M. B. 2010. Antibodies to human serum amyloid P component eliminate visceral amyloid deposits. *Nature*, 468, 93-97.
- BORK, P., HOLM, L. & SANDER, C. 1994. The immunoglobulin fold - structural classification, sequence patterns and common core. *Journal of Molecular Biology*, 242, 309-320.
- BREDESEN, D. 2009. Neurodegeneration in Alzheimer's disease: caspases and synaptic element interdependence. *Molecular Neurodegeneration*, 4, 27.
- BU, Z., SHI, Y., CALLAWAY, D. J. E. & TYCKO, R. 2007. Molecular alignment within  $\beta$ -sheets in A $\beta$ 14-23 fibrils: Solid-state NMR experiments and theoretical predictions. *Biophysical Journal*, 92, 594-602.
- BUCCIANTINI, M., GIANNONI, E., CHITI, F., BARONI, F., FORMIGLI, L., ZURDO, J., TADDEI, N., RAMPONI, G., DOBSON, C. M. & STEFANI, M. 2002. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature*, 416, 507-511.

- BUCHETE, N. V., TYCKO, R. & HUMMER, G. 2005. Molecular dynamics simulations of Alzheimer's  $\beta$ -amyloid protofilaments. *Journal of Molecular Biology*, 353, 804-821.
- BULAWA, C. E., CONNELLY, S., DEVIT, M., WANG, L., WEIGEL, C., FLEMING, J. A., PACKMAN, J., POWERS, E. T., WISEMAN, R. L., FOSS, T. R., WILSON, I. A., KELLY, J. W. & LABAUDINIÈRE, R. 2012. Tafamidis, a potent and selective transthyretin kinetic stabilizer that inhibits the amyloid cascade. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 9629-9634.
- BURKOTH, T. S., BENZINGER, T. L. S., URBAN, V., MORGAN, D. M., GREGORY, D. M., THIYAGARAJAN, P., BOTTO, R. E., MEREDITH, S. C. & LYNN, D. G. 2000. Structure of the beta-amyloid((10-35)) fibril. *Journal of the American Chemical Society*, 122, 7883-7889.
- BUTTERFIELD, S. M. & LASHUEL, H. A. 2010. Amyloidogenic Protein Membrane Interactions: Mechanistic Insight from Model Systems. *Angewandte Chemie-International Edition*, 49, 5628-5654.
- BUXBAUM, J. 1992. Mechanisms of disease: Monoclonal immunoglobulin deposition. *Hematology/Oncology Clinics of North America*, 6, 323-346.
- BUXBAUM, J. N. & LINKE, R. P. 2012. A Molecular History of the Amyloidoses. *Journal of Molecular Biology*, 421, 142-159.
- CABRIOLU, R., KASHCHIEV, D. & AUER, S. 2010. Atomistic theory of amyloid fibril nucleation. *The Journal of Chemical Physics*, 133, 225101.
- CELIS, J. E., CARTER, N., HUNTER, T., SIMMONS, K., SMALL, J. V. AND SHORLTON D. (ed.) 2006. *Cell Biology: A laboratory handbook*: Elsevier Academic Press.
- CERIANI, R. L. & BLANK, E. W. 1988. EXPERIMENTAL-THERAPY OF HUMAN-BREAST TUMORS WITH I-131-LABELED MONOCLONAL-ANTIBODIES PREPARED AGAINST THE HUMAN-MILK FAT GLOBULE. *Cancer Research*, 48, 4664-4672.
- CERIANI, R. L., THOMPSON, K., PETERSON, J. A. & ABRAHAM, S. 1977. SURFACE DIFFERENTIATION ANTIGENS OF HUMAN MAMMARY EPITHELIAL-CELLS CARRIED ON HUMAN MILK-FAT GLOBULE. *Proceedings of the National Academy of Sciences of the United States of America*, 74, 582-586.
- CHAPMAN, M. R., ROBINSON, L. S., PINKNER, J. S., ROTH, R., HEUSER, J., HAMMAR, M., NORMARK, S. & HULTGREN, S. J. 2002. Role of Escherichia coli Curli Operons in Directing Amyloid Fiber Formation. *Science*, 295, 851-855.
- CHEN, W., GAMACHE, E., RICHARDSON, D., DU, Z. & WANG, C. 2012. Expression, purification, and reconstitution of the transmembrane domain of the human amyloid precursor protein for NMR studies. *Protein Expression and Purification*, 81, 11-17.
- CHITI, F., BUCCIANTINI, M., CAPANNI, C., TADDEI, N., DOBSON, C. M. & STEFANI, M. 2001. Solution conditions can promote formation of either amyloid protofilaments or mature fibrils from the HypF N-terminal domain. *Protein Science*, 10, 2541-2547.
- CHITI, F. & DOBSON, C. M. 2006. Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry*, 75, 333-366.



- CHITI, F., STEFANI, M., TADDEI, N., RAMPONI, G. & DOBSON, C. M. 2003. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature*, 424, 805-808.
- CHIU, J., MARCH, P. E., LEE, R. & TILLET, D. 2004. Site-directed, Ligase-Independent Mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4 h. *Nucleic Acids Research*, 32, e174.
- CHOOSMITH, L. P., GARZONRODRIGUEZ, W., GLABE, C. G. & SUREWICZ, W. K. 1997. Acceleration of amyloid fibril formation by specific binding of A beta-(1-40) peptide to ganglioside-containing membrane vesicles. *Journal of Biological Chemistry*, 272, 22987-22990.
- COHLBERG, J. A., LI, J., UVERSKY, V. N. & FINK, A. L. 2002. Heparin and Other Glycosaminoglycans Stimulate the Formation of Amyloid Fibrils from  $\alpha$ -Synuclein in Vitro†. *Biochemistry*, 41, 1502-1511.
- COLLINS, S. R., DOUGLASS, A., VALE, R. D. & WEISSMAN, J. S. 2004. Mechanism of prion propagation: Amyloid growth occurs by monomer addition. *PLoS Biology*, 2, 1582-1590.
- COME, J. H., FRASER, P. E. & LANSBURY, P. T. 1993. A kinetic model for amyloid formation in the prion diseases: importance of seeding. *Proceedings of the National Academy of Sciences*, 90, 5959-5963.
- CONCHILLO-SOLÉ, O., DE GROOT, N. S., AVILÉS, F. X., VENDRELL, J., DAURA, X. & VENTURA, S. 2007. AGGREGSCAN: A server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. *BMC Bioinformatics*, 8.
- CONNORS, L. H., RICHARDSON, A. M., THEBERGE, R. & COSTELLO, C. E. 2000. Sequence Communication: Tabulation of transthyretin (TTR) variants as of 1/1/2000. *Amyloid*, 7, 54-69.
- CORNWELL, G. G. & WESTERMARK, P. 1980. Senile amyloidosis: a protean manifestation of the aging process. *Journal of Clinical Pathology*, 33, 1146-1152.
- CORNWELL, G. G., WESTERMARK, P., MURDOCH, W., PITKANEN, P. 1982. Senile aortic amyloid - A third distinctive type of age-related cardiovascular amyloid. *American Journal of Pathology*, 108, 135-139.
- CORNWELL III, G. G., MURDOCH, W. L., KYLE, R. A., WESTERMARK, P. & PITKÄNEN, P. 1983. Frequency and distribution of senile cardiovascular amyloid: A clinicopathologic correlation. *The American Journal of Medicine*, 75, 618-623.
- COUTO, J. R., TAYLOR, M. R., GODWIN, S. G., CERIANI, R. L. & PETERSON, J. A. 1996. Cloning and sequence analysis of human breast epithelial antigen BA46 reveals an RGD cell adhesion sequence presented on an epidermal growth factor-like domain. *DNA and Cell Biology*, 15, 281-286.
- CRIBBS, D. H., AZIZEH, B. Y., COTMAN, C. W. & LAFERLA, F. M. 2000. Fibril Formation and Neurotoxicity by a Herpes Simplex Virus Glycoprotein B Fragment with Homology to the Alzheimer's A $\beta$  Peptide†. *Biochemistry*, 39, 5988-5994.
- CRUZ, L., RAO, J. S., TELOW, D. B. & URBANC, B. 2012. Dynamics of Metastable beta-Hairpin Structures in the Folding Nucleus of Amyloid beta-Protein. *Journal of Physical Chemistry B*, 116, 6311-6325.

- CUKALEVSKI, R., BOLAND, B., FROHM, B., THULIN, E., WALSH, D. & LINSE, S. 2012. Role of Aromatic Side Chains in Amyloid beta-Protein Aggregation. *Acs Chemical Neuroscience*, 3, 1008-1016.
- DAVIES, H. A., MADINE, J. & MIDDLETON, D. A. 2012. Solid-state NMR reveals differences in the packing arrangements of peptide aggregates derived from the aortic amyloid polypeptide medin. *Journal of Peptide Science*, 18, 65-72.
- DAVIS, J., XU, F., DEANE, R., ROMANOV, G., PREVITI, M. L., ZEIGLER, K., ZLOKOVIC, B. V. & VAN NOSTRAND, W. E. 2004. Early-onset and robust cerebral microvascular accumulation of amyloid beta-protein in transgenic mice expressing low levels of a vasculotropic Dutch/Iowa mutant form of amyloid beta-protein precursor. *Journal of Biological Chemistry*, 279, 20296-20306.
- DEMURO, A., SMITH, M. & PARKER, I. 2011. Single-channel Ca<sup>2+</sup> imaging implicates A beta 1-42 amyloid pores in Alzheimer's disease pathology. *Journal of Cell Biology*, 195, 515-524.
- DESPOIT, E., BRIDOUX, F., SIRAC, C., DELBES, S., BENDER, S., FERNANDEZ, B., QUELLARD, N., LACOMBE, C., GOUJON, J. M., LAVERGNE, D., ABRAHAM, J., TOUCHARD, G., FERMAND, J. P., JACCARD, A. & CTR NATL REFERENCE POUR AMYLOSE, A. L. 2012. AL Amyloidosis. *Orphanet Journal of Rare Diseases*, 7.
- DIVREY, P. F., M. 1927. Sur les proprietes optiques de l'amyloide. *Comptes Rendus Soc. Biol*, 97, 1808-1810.
- DOBSON, C. M. 2003. Protein folding and misfolding. *Nature*, 426, 884-890.
- DOBSON, C. M. 2006. Protein aggregation and its consequences for human disease. *Protein and Peptide Letters*, 13, 219-227.
- DORAN, T. M., ANDERSON, E. A., LATCHNEY, S. E., OPANASHUK, L. A. & NILSSON, B. L. 2012a. Turn Nucleation Perturbs Amyloid  $\beta$  Self-Assembly and Cytotoxicity. *Journal of Molecular Biology*, 421, 315-328.
- DORAN, T. M., KAMENS, A. J., BYRNES, N. K. & NILSSON, B. L. 2012b. Role of amino acid hydrophobicity, aromaticity, and molecular volume on IAPP(20-29) amyloid self-assembly. *Proteins-Structure Function and Bioinformatics*, 80, 1053-1065.
- DUBREY, S. W., HAWKINS, P. N. & FALK, R. H. 2011. Amyloid diseases of the heart: assessment, diagnosis, and referral. *Heart*, 97, 75-84.
- DUNGU, J. N., ANDERSON, L. J., WHELAN, C. J. & HAWKINS, P. N. 2012. Cardiac transthyretin amyloidosis. *Heart*.
- DZWOLAK, W., SMIRNOVAS, V., JANSEN, R. & WINTER, R. 2004. Insulin forms amyloid in a strain-dependent manner: An FT-IR spectroscopic study. *Protein Science*, 13, 1927-1932.
- EISENBERG, D. & JUCKER, M. 2012. The Amyloid State of Proteins in Human Diseases. *Cell*, 148, 1188-1203.
- EMADI, S., BARKHORDARIAN, H., WANG, M. S., SCHULZ, P. & SIERKS, M. R. 2007. Isolation of a human single chain antibody fragment against oligomeric alpha-synuclein that inhibits aggregation and prevents alpha-synuclein-induced toxicity. *Journal of Molecular Biology*, 368, 1132-1144.
- ENQVIST, S., SLETTEN, K. & WESTERMARK, P. 2009. Fibril protein fragmentation pattern in systemic AL-amyloidosis. *The Journal of Pathology*, 219, 473-480.

- ERICZON, B.-G., LARSSON, M., HERLENIUS, G. & WILCZEK, H. E. 2003. Report from the Familial Amyloidotic Polyneuropathy World Transplant Registry (FAPWTR) and the Domino Liver Transplant Registry (DLTR). *Amyloid: The International Journal Of Experimental And Clinical Investigation: The Official Journal Of The International Society Of Amyloidosis*, 10 Suppl 1, 67-76.
- ESWAR, N., WEBB, B., MARTI-RENOM, M. A., MADHUSUDHAN, M. S., ERAMIAN, D., SHEN, M.-Y., PIEPER, U. & SALI, A. 2007. Comparative protein structure modeling using MODELLER. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]*, Chapter 2, Unit 2.9.
- FALK, R. H. & DUBREY, S. W. 2005. Amyloid Heart Disease. *Progress in Cardiovascular Diseases*, 52, 347-361.
- FALK, R. H., RAYMOND, M D, COMENZO, R L, SKINNER, M. 1998. The Systemic Amyloidoses. *New England Journal of Medicine*, 337, 898-909.
- FÄNDRICH, M. 2012. Oligomeric Intermediates in Amyloid Formation: Structure Determination and Mechanisms of Toxicity. *Journal of Molecular Biology*, 421, 427-440.
- FAVIA, A. D., NICOLOTTI, O., STEFANACHI, A., LEONETTI, F. & CAROTTI, A. 2013. Computational methods for the design of potent aromatase inhibitors. *Expert Opinion on Drug Discovery*, 8, 395-409.
- FAWZI, N. L., PHILLIPS, A. H., RUSCIO, J. Z., DOUCLEFF, M., WEMMER, D. E. & HEAD-GORDON, T. 2008. Structure and dynamics of the A<sub>ss</sub>(21-30) peptide from the interplay of NMR experiments and molecular simulations. *Journal of the American Chemical Society*, 130, 6145-6158.
- FINDER, V. H., VODOPIVEC, I., NITSCH, R. M. & GLOCKSHUBER, R. 2010. The Recombinant Amyloid-beta Peptide A<sub>beta</sub> 1-42 Aggregates Faster and Is More Neurotoxic than Synthetic A<sub>beta</sub> 1-42. *Journal of Molecular Biology*, 396, 9-18.
- FLACH, K., HILBRICH, I., SCHIFFMANN, A., GARTNER, U., KRUGER, M., LEONHARDT, M., WASCHIPKY, H., WICK, L., ARENDT, T. & HOLZER, M. 2012. Tau Oligomers Impair Artificial Membrane Integrity and Cellular Viability. *Journal of Biological Chemistry*, 287, 43223-43233.
- FU, W., RUANGKITISAKUL, A., MACTAVISH, D., SHI, J. Y., BALLANYI, K. & JHAMANDAS, J. H. 2012. Amyloid  $\beta$  (A $\beta$ ) Peptide Directly Activates Amylin-3 Receptor Subtype by Triggering Multiple Intracellular Signaling Pathways. *Journal of Biological Chemistry*, 287, 18820-18830.
- FURNHAM, N., LASKOWSKI, R. A. & THORNTON, J. M. 2013. Abstracting Knowledge from the Protein Data Bank. *Biopolymers*, 99, 183-188.
- GAL, R., KORZETS, A., SCHWARTZ, A., RATH-WOLFSON, L. & GAFTER, U. 1994. Systemic distribution of  $\beta$ 2-microglobulin-derived amyloidosis in patients who undergo long-term hemodialysis: Report of seven cases and review of the literature. *Archives of Pathology and Laboratory Medicine*, 118, 718-721.
- GAST, K. & MODLER, A. J. 2005. *Studying Protein Folding and Aggregation by Laser Light Scattering*, Wiley-V C H Verlag Gmbh, Pappelallee 3, W-69469 Weinheim, Germany.
- GAZIT, E. 2002. A possible role for pi-stacking in the self-assembly of amyloid fibrils. *FASEB Journal*, 16, 77-83.

- GAZIT, E., DELLA BRUNA, P., PIERACCINI, S. & COLOMBO, G. 2007. The molecular dynamics of assembly of the ubiquitous aortic medial amyloidoid medin fragment. *Journal of Molecular Graphics and Modelling*, 25, 903-911.
- GEJYO, F., YAMADA, T., ODANI, S., NAKAGAWA, Y., ARAKAWA, M., KUNITOMO, T., KATAOKA, H., SUZUKI, M., HIRASAWA, Y., SHIRAHAMA, T., COHEN, A. S. & SCHMID, K. 1985. A new form of amyloid protein associated with chronic hemodialysis was identified as  $\beta$ 2-microglobulin. *Biochemical and Biophysical Research Communications*, 129, 701-706.
- GENDOO, D. M. A. & HARRISON, P. M. 2011. Origins and Evolution of the HET-s Prion-Forming Protein: Searching for Other Amyloid-Forming Solenoids. *PLoS ONE*, 6, e27342.
- GILLMORE, J. D. & HAWKINS, P. N. 2006. Drug insight: emerging therapies for amyloidosis. *Nature Clinical Practice Nephrology*, 2, 263-270.
- GOLDBURG, W. I. 1999. Dynamic light scattering. *American Journal of Physics*, 67, 1152-1160.
- GOLDSBURY, C., GOLDIE, K., PELLAUD, J., SEELIG, J., FREY, P., MÜLLER, S. A., KISTLER, J., COOPER, G. J. S. & AEBI, U. 2000. Amyloid Fibril Formation from Full-Length and Fragments of Amylin. *Journal of Structural Biology*, 130, 352-362.
- GRABOWSKI, T. J., CHO, H. S., VONSATTEL, J. P. G., REBECK, G. W. & GREENBERG, S. M. 2001. Novel amyloid precursor protein mutation in an Iowa family with dementia and severe cerebral amyloid angiopathy. *Annals of Neurology*, 49, 697-705.
- GRAS, S., WADDINGTON, L. & GOLDIE, K. 2011. Transmission Electron Microscopy of Amyloid Fibrils. In: HILL, A. F., BARNHAM, K. J., BOTTOMLEY, S. P. & CAPPAL, R. (eds.) *Protein Folding, Misfolding, and Disease*. Humana Press.
- GROENNING, M. 2010. Binding mode of Thioflavin T and other molecular probes in the context of amyloid fibrils—current status. *Journal of Chemical Biology*, 3, 1-18.
- GROSS, M. 2000. Proteins that Convert from alpha Helix to beta Sheet: Implications for Folding and Disease. *Current Protein & Peptide Science*, 1, 339-347.
- HAGGQVIST, B., MUCCHIANO, G., SLETTEN, K., NASLUND, J. & WESTERMARK, P. 1999. Amyloid protein in aortic media is partly characterized. *Amyloid and Amyloidosis 1998*, 562-564.
- HÄGGQVIST, B., NÄSLUND, J., SLETTEN, K., WESTERMARK, G. T., MUCCHIANO, G., TJERNBERG, L. O., NORDSTEDT, C., ENGSTRÖM, U. & WESTERMARK, P. 1999. Medin: An integral fragment of aortic smooth muscle cell-produced lactadherin forms the most common human amyloid. *Proceedings of the National Academy of Sciences*, 96, 8669-8674.
- HAN, B., LIU, Y., GINZINGER, S. & WISHART, D. 2011. SHIFTX2: significantly improved protein chemical shift prediction. *Journal of Biomolecular NMR*, 50, 43-57.
- HANAHAHAN, D., JESSEE, J. & BLOOM, F. R. 1991. Plasmid transformation of Escherichia coli and other bacteria. In: JEFFREY, H. M. (ed.) *Methods in Enzymology*. Academic Press.

- HANAYAMA, R., TANAKA, M., MIWA, K., SHINOHARA, A., IWAMATSU, A. & NAGATA, S. 2002. Identification of a factor that links apoptotic cells to phagocytes. *Nature*, 417, 182-187.
- HASSAN, W., AL-SERGANI, H., MOURAD, W. & TABBAA, R. 2005. Amyloid heart disease. New frontiers and insights in pathophysiology, diagnosis, and management. *Texas Heart Institute Journal / From The Texas Heart Institute Of St. Luke's Episcopal Hospital, Texas Children's Hospital*, 32, 178-184.
- HATTERS, D. M., MINTON, A. P. & HOWLETT, G. J. 2002. Macromolecular Crowding Accelerates Amyloid Formation by Human Apolipoprotein C-II. *Journal of Biological Chemistry*, 277, 7824-7830.
- HAWKINS, P. N., APRILE, C., CAPRIA, G., VIGANO, L., MUNZONE, E., GIANNI, L., PEPYS, M. B. & MERLINI, G. 1998. Scintigraphic imaging and turnover studies with iodine-131 labelled serum amyloid P component in systemic amyloidosis. *European Journal of Nuclear Medicine*, 25, 701-708.
- HEISE, H. 2008. Solid-state NMR spectroscopy of amyloid proteins. *Chembiochem*, 9, 179-189.
- HELMUS, J. J., SUREWICZ, K., APOSTOL, M. I., SUREWICZ, W. K. & JARONIEC, C. P. 2011. Intermolecular Alignment in Y145Stop Human Prion Protein Amyloid Fibrils Probed by Solid-State NMR Spectroscopy. *Journal of the American Chemical Society*, 133, 13934-13937.
- HIROAKI, H., UMETSU, Y., NABESHIMA, Y. I., HOSHI, M. & KOHDA, D. 2011. A simplified recipe for assigning amide NMR signals using combinatorial <sup>14</sup>N amino acid inverse-labeling. *Journal of Structural and Functional Genomics*, 12, 167-174.
- HONG, M., GROSS, J. D. & GRIFFIN, R. G. 1997. Site-Resolved Determination of Peptide Torsion Angle  $\phi$  from the Relative Orientations of Backbone N-H and C-H Bonds by Solid-State NMR. *The Journal of Physical Chemistry B*, 101, 5869-5874.
- HOPWOOD, D. 1972. Theoretical and practical aspects of glutaraldehyde fixation. *The Histochemical Journal*, 4, 267-303.
- HUANG, D. B., CHANG, C. H., AINSWORTH, C., JOHNSON, G., SOLOMON, A., STEVENS, F. J. & SCHIFFER, M. 1997. Variable domain structure of kappa IV human light chain Len: High homology to the murine light chain McPC603. *Molecular Immunology*, 34, 1291-1301.
- HUANG, X. Q. & MILLER, W. 1991. A TIME-EFFICIENT, LINEAR-SPACE LOCAL SIMILARITY ALGORITHM. *Advances in Applied Mathematics*, 12, 337-357.
- HUGHES, E., BURKE, R. M. & DOIG, A. J. 2000. Inhibition of toxicity in the beta-amyloid peptide fragment beta-(25-35) using N-methylated derivatives - A general strategy to prevent amyloid formation. *Journal of Biological Chemistry*, 275, 25109-25115.
- HUIYUAN, L., FARID, R., SHARMISTHA, S., PANCHANAN, M. AND GAL, B. 2009. Amyloids and protein aggregation - analytical methods. In: MEYERS, R. A. (ed.) *Encyclopedia of analytical chemistry*. Wiley & Sons Ltd.
- HUSTER, D., XIAO, L. S. & HONG, M. 2001. Solid-state NMR investigation of the dynamics of the soluble and membrane-bound colicin Ia channel-forming domain. *Biochemistry*, 40, 7662-7674.
- IONESCU-ZANETTI, C., KHURANA, R., GILLESPIE, J. R., PETRICK, J. S., TRABACHINO, L. C., MINERT, L. J., CARTER, S. A. & FINK, A. L.

1999. Monitoring the assembly of Ig light-chain amyloid fibrils by atomic force microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 13175-13179.
- IWATA, K., FUJIWARA, T., MATSUKI, Y., AKUTSU, H., TAKAHASHI, S., NAIKI, H. & GOTO, Y. 2006. 3D structure of amyloid protofilaments of  $\beta$ 2-microglobulin fragment probed by solid-state NMR. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 18119-18124.
- IWATA, T., KAMEI, T., UCHINO, F., MIMAYA, H., YANAGAKI, T. & ETOH, H. 1978. PATHOLOGICAL STUDY ON AMYLOIDOSIS — RELATIONSHIP OF AMYLOID DEPOSITS IN THE AORTA TO AGING—. *Pathology International*, 28, 193-203.
- JACOBSON, D. R., PASTORE, R., POOL, S., MALENDOWICZ, S., KANE, I., SHIVJI, A., EMBURY, S. H., BALLAS, S. K. & BUXBAUM, J. N. 1996. Revised transthyretin Ile 122 allele frequency in African-Americans. *Human Genetics*, 98, 236-238.
- JAHN, T. R., MAKIN, O. S., MORRIS, K. L., MARSHALL, K. E., TIAN, P., SIKORSKI, P. & SERPELL, L. C. 2010. The Common Architecture of Cross-beta Amyloid. *Journal of Molecular Biology*, 395, 717-727.
- JARONIEC, C. P., MACPHEE, C. E., ASTROF, N. S., DOBSON, C. M. & GRIFFIN, R. G. 2002. Molecular conformation of a peptide fragment of transthyretin in an amyloid fibril. *Proceedings of the National Academy of Sciences*, 99, 16748-16753.
- JARONIEC, C. P., TOUNGE, B. A., HERZFELD, J. & GRIFFIN, R. G. 2001. Frequency Selective Heteronuclear Dipolar Recoupling in Rotating Solids: Accurate  $^{13}\text{C}$ - $^{15}\text{N}$  Distance Measurements in Uniformly  $^{13}\text{C}$ , $^{15}\text{N}$ -labeled Peptides. *Journal of the American Chemical Society*, 123, 3507-3519.
- JARRETT, J. T. & LANSBURY JR, P. T. 1993. Seeding “one-dimensional crystallization” of amyloid: A pathogenic mechanism in Alzheimer's disease and scrapie? *Cell*, 73, 1055-1058.
- JAYARAMAN, M., KODALI, R., SAHOO, B., THAKUR, A. K., MAYASUNDARI, A., MISHRA, R., PETERSON, C. B. & WETZEL, R. 2012. Slow Amyloid Nucleation via alpha-Helix-Rich Oligomeric Intermediates in Short Polyglutamine-Containing Huntingtin Fragments. *Journal of Molecular Biology*, 415, 881-899.
- JOHNSON, P. S., COOK, P. L., LIU, X., YANG, W., BAI, Y., ABBOTT, N. L. & HIMPSEL, F. J. 2011. Universal mechanism for breaking amide bonds by ionizing radiation. *The Journal of Chemical Physics*, 135, 044702.
- JONES, S., MANNING, J., KAD, N. M. & RADFORD, S. E. 2003. Amyloid-forming Peptides from  $\beta$ 2-Microglobulin—Insights into the Mechanism of Fibril Formation in Vitro. *Journal of Molecular Biology*, 325, 249-257.
- JUCKER, M. & WALKER, L. C. 2011. Pathogenic Protein Seeding in Alzheimer Disease and Other Neurodegenerative Disorders. *Annals of Neurology*, 70, 532-540.
- KAPURNIOTU, A., SCHMAUDER, A. & TENIDIS, K. 2002. Structure-based design and study of non-amyloidogenic, double N-methylated IAPP amyloid core sequences as inhibitors of IAPP amyloid formation and cytotoxicity. *Journal of Molecular Biology*, 315, 339-350.

- KAPUST, R. B. & WAUGH, D. S. 2000. Controlled Intracellular Processing of Fusion Proteins by TEV Protease. *Protein Expression and Purification*, 19, 312-318.
- KAYED, R., CANTO, I., BREYDO, L., RASOOL, S., LUKACSOVICH, T., WU, J., ALBAY, R., III, PENSALFINI, A., YEUNG, S., HEAD, E., MARSH, J. L. & GLABE, C. 2010. Conformation dependent monoclonal antibodies distinguish different replicating strains or conformers of prefibrillar A beta oligomers. *Molecular Neurodegeneration*, 5.
- KAYED, R., HEAD, E., THOMPSON, J. L., MCINTIRE, T. M., MILTON, S. C., COTMAN, C. W. & GLABE, C. G. 2003. Common Structure of Soluble Amyloid Oligomers Implies Common Mechanism of Pathogenesis. *Science*, 300, 486-489.
- KENNEY, J. M., KNIGHT, D., WISE, M. J. & VOLLRATH, F. 2002. Amyloidogenic nature of spider silk. *European Journal of Biochemistry*, 269, 4159-4163.
- KHOLOVÁ, I. & NIESSEN, H. W. M. 2005. Amyloid in the cardiovascular system: a review. *Journal of Clinical Pathology*, 58, 125-133.
- KHURANA, R., GILLESPIE, J. R., TALAPATRA, A., MINERT, L. J., IONESCU-ZANETTI, C., MILLETT, I. & FINK, A. L. 2001. Partially folded intermediates as critical precursors of light chain amyloid fibrils and amorphous aggregates. *Biochemistry*, 40, 3525-3535.
- KHURANA, R., SOUILLAC, P. O., COATS, A. C., MINERT, L., IONESCU-ZANETTI, C., CARTER, S. A., SOLOMON, A. & FINK, A. L. 2003. A model for amyloid fibril formation in immunoglobulin light chains based on comparison of amyloidogenic and benign proteins and specific antibody binding. *Amyloid-Journal of Protein Folding Disorders*, 10, 97-109.
- KIM, H. J., CHATANI, E., GOTO, Y. & PAIK, S. R. 2007. Seed-dependent accelerated fibrillation of alpha-synuclein induced by periodic ultrasonication treatment. *Journal of Microbiology and Biotechnology*, 17, 2027-2032.
- KIRKITADZE, M. D., CONDRON, M. M. & TEPLow, D. B. 2001. Identification and characterization of key kinetic intermediates in amyloid  $\beta$ -protein fibrillogenesis. *Journal of Molecular Biology*, 312, 1103-1119.
- KLABUNDE, T., PETRASSI, H. M., OZA, V. B., RAMAN, P., KELLY, J. W. & SACCHETTINI, J. C. 2000. Rational design of potent human transthyretin amyloid disease inhibitors. *Nat Struct Mol Biol*, 7, 312-321.
- KNIGHT, J. D. & MIRANKER, A. D. 2004. Phospholipid catalysis of diabetic amyloid assembly. *Journal of Molecular Biology*, 341, 1175-1187.
- KREBS, M. R. H., MOROZOVA-ROCHE, L. A., DANIEL, K., ROBINSON, C. V. & DOBSON, C. M. 2004. Observation of sequence specificity in the seeding of protein amyloid fibrils. *Protein Science*, 13, 1933-1938.
- KUMAR, S., MOHANTY, S. K. & UDGAONKAR, J. B. 2007. Mechanism of Formation of Amyloid Protofibrils of Barstar from Soluble Oligomers: Evidence for Multiple Steps and Lateral Association Coupled to Conformational Conversion. *Journal of Molecular Biology*, 367, 1186-1204.
- KUMAR, S. & NUSSINOV, R. 2002. Close-Range Electrostatic Interactions in Proteins. *Chembiochem*, 3, 604-617.
- KYLE, R., LINOS, A., BEARD, C., LINKE, R., GERTZ, M., O'FALLON, W. & KURLAND, L. 1992. Incidence and natural history of primary systemic

- amyloidosis in Olmsted County, Minnesota, 1950 through 1989 [see comments]. *Blood*, 79, 1817-1822.
- KYLE, R. A. & GERTZ, M. A. 1995. Primary systemic amyloidosis: clinical and laboratory features in 474 cases. *Seminars in hematology*, 32, 45-59.
- LACHMANN, H. J., GALLIMORE, R., GILLMORE, J. D., CARR-SMITH, H. D., BRADWELL, A. R., PEPYS, M. B. & HAWKINS, P. N. 2003. Outcome in systemic AL amyloidosis in relation to changes in concentration of circulating free immunoglobulin light chains following chemotherapy. *British Journal of Haematology*, 122, 78-84.
- LAEMMLI, U. K. 1970. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, 227, 680-685.
- LAKOWICZ, J. R. 2006. *Protein Fluorescence*, Springer US.
- LAKSHMANAN, A., CHEONG, D. W., ACCARDO, A., DI FABRIZIO, E., RIEKEL, C. & HAUSER, C. A. E. 2013. Aliphatic peptides show similar self-assembly to amyloid core sequences, challenging the importance of aromatic interactions in amyloidosis. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 519-24.
- LAROCCA, D., PETERSON, J. A., URREA, R., KUNIYOSHI, J., BISTRAN, A. M. & CERIANI, R. L. 1991. A Mr 46,000 Human Milk Fat Globule Protein That Is Highly Expressed in Human Breast Tumors Contains Factor VIII-like Domains. *Cancer Research*, 51, 4994-4998.
- LARSSON, A., MALMSTRÖM, S. & WESTERMARK, P. 2011. Signs of cross-seeding: aortic median amyloid as a trigger for protein AA deposition. *Amyloid*, 18, 229-234.
- LARSSON, A., PENG, S., PERSSON, H., ROSENBLOOM, J., ABRAMS, W. R., WASSBERG, E., THELIN, S., SLETTEN, K., GERWINS, P. & WESTERMARK, P. 2006. Lactadherin binds to elastin - A starting point for median amyloid formation? *Amyloid*, 13, 78-85.
- LARSSON, A., SÖDERBERG, L., WESTERMARK, G. T., SLETTEN, K., ENGSTRÖM, U., TJERNBERG, L. O., NÄSLUND, J. & WESTERMARK, P. 2007. Unwinding fibril formation of median, the peptide of the most common form of human amyloid. *Biochemical and Biophysical Research Communications*, 361, 822-828.
- LASHUEL, H. A., HARTLEY, D. M., PETRE, B. M., WALL, J. S., SIMON, M. N., WALZ, T. & LANSBURY JR, P. T. 2003. Mixtures of Wild-type and a Pathogenic (E22G) Form of A $\beta$ 40 in Vitro Accumulate Protofibrils, Including Amyloid Pores. *Journal of Molecular Biology*, 332, 795-808.
- LAZO, N. D., GRANT, M. A., CONDRON, M. C., RIGBY, A. C. & TEPLow, D. B. 2005. On the nucleation of amyloid beta-protein monomer folding. *Protein Science*, 14, 1581-1596.
- LEE, C.-C., NAYAK, A., SETHURAMAN, A., BELFORT, G. & MCRAE, G. J. 2007. A Three-Stage Kinetic Model of Amyloid Fibrillation. *Biophysical Journal*, 92, 3448-3458.
- LEVINE, H. 1993. Thioflavine T interaction with synthetic Alzheimer's disease  $\beta$ -amyloid peptides: Detection of amyloid aggregation in solution. *Protein Science*, 2, 404-410.
- LEVINE, H. 1999. Quantification of  $\beta$ -sheet amyloid fibril structures with thioflavin T. In: RONALD, W. (ed.) *Methods in Enzymology*. Academic Press.



- LEWANDOWSKI, J. R., VAN DER WEL, P. C. A., RIGNEY, M., GRIGORIEFF, N. & GRIFFIN, R. G. 2011. Structural Complexity of a Composite Amyloid Fibril. *Journal of the American Chemical Society*, 133, 14686-14698.
- LIAN, L. Y. & MIDDLETON, D. A. 2001. Labelling approaches for protein structural studies by solution-state and solid-state NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 39, 171-190.
- LIN, L., HUAI, Q., HUANG, M., FURIE, B. & FURIE, B. C. 2007. Crystal Structure of the Bovine Lactadherin C2 Domain, a Membrane Binding Motif, Shows Similarity to the C2 Domains of Factor V and Factor VIII. *Journal of Molecular Biology*, 371, 717-724.
- LINK, A. J. & LABAER, J. 2011. Trichloroacetic Acid (TCA) Precipitation of Proteins. *Cold Spring Harbor Protocols*, 2011, pdb.prot5651.
- LOVAT, L. B., O'BRIEN, A. A. J., ARMSTRONG, S. F., MADHOO, S., BULPITT, C. J., ROSSOR, M. N., PEPYS, M. B. & HAWKINS, P. N. 1998a. Scintigraphy with I-123-serum amyloid P component in Alzheimer disease. *Alzheimer Disease & Associated Disorders*, 12, 208-210.
- LOVAT, L. B., PERSEY, M. R., MADHOO, S., PEPYS, M. B. & HAWKINS, P. N. 1998b. The liver in systemic amyloidosis: insights from I-123 serum amyloid P component scintigraphy in 484 patients. *Gut*, 42, 727-734.
- LUCA, S., YAU, W. M., LEAPMAN, R. & TYCKO, R. 2007. Peptide conformation and supramolecular organization in amylin fibrils: Constraints from solid-state NMR. *Biochemistry*, 46, 13505-13522.
- MABBOTT, N. A. & MACPHERSON, G. G. 2006. Prions and their lethal journey to the brain. *Nat Rev Micro*, 4, 201-211.
- MACEDO-RIBEIRO, S., BODE, W., HUBER, R., QUINN-ALLEN, M. A., KIM, S. W., ORTEL, T. L., BOURENKOV, G. P., BARTUNIK, H. D., STUBBS, M. T., KANE, W. H. & FUENTES-PRIOR, P. 1999. Crystal structures of the membrane-binding C2 domain of human coagulation factor V. *Nature*, 402, 434-439.
- MADINE, J., COPLAND, A., SERPELL, L. C. & MIDDLETON, D. A. 2009a. Cross- $\beta$  Spine architecture of fibrils formed by the amyloidogenic segment NFGSVQFV of medin from solid-state NMR and X-ray fiber diffraction measurements. *Biochemistry*, 48, 3089-3099.
- MADINE, J., DOIG, A. J. & MIDDLETON, D. A. 2008a. Design of an N-methylated peptide inhibitor of alpha-synuclein aggregation guided by solid-state NMR. *Journal of the American Chemical Society*, 130, 7873-7881.
- MADINE, J., DOIG, A. J. & MIDDLETON, D. A. 2008b. Design of an N-Methylated Peptide Inhibitor of  $\alpha$ -Synuclein Aggregation Guided by Solid-State NMR. *Journal of the American Chemical Society*, 130, 7873-7881.
- MADINE, J., JACK, E., STOCKLEY, P. G., RADFORD, S. E., SERPELL, L. C. & MIDDLETON, D. A. 2008c. Structural Insights into the Polymorphism of Amyloid-Like Fibrils Formed by Region 20-29 of Amylin Revealed by Solid-State NMR and X-ray Fiber Diffraction. *Journal of the American Chemical Society*, 130, 14990-15001.
- MADINE, J. & MIDDLETON, D. A. 2010. Comparison of aggregation enhancement and inhibition as strategies for reducing the cytotoxicity of the aortic amyloid polypeptide medin. *European Biophysics Journal*, 39, 1281-1288.

- MADINE, J., WANG, X. Y., BROWN, D. R. & MIDDLETON, D. A. 2009b. Evaluation of beta-Alanine- and GABA-Substituted Peptides as Inhibitors of Disease-Linked Protein Aggregation. *ChemBiochem*, 10, 1982-1987.
- MAJI, S. K., AMSDEN, J. J., ROTHSCILD, K. J., CONDRON, M. M. & TELOW, D. B. 2005. Conformational Dynamics of Amyloid  $\beta$ -Protein Assembly Probed Using Intrinsic Fluorescence†. *Biochemistry*, 44, 13365-13376.
- MAJI, S. K., PERRIN, M. H., SAWAYA, M. R., JESSBERGER, S., VADODARIA, K., RISSMAN, R. A., SINGRU, P. S., NILSSON, K. P. R., SIMON, R., SCHUBERT, D., EISENBERG, D., RIVIER, J., SAWCHENKO, P., VALE, W. & RIEK, R. 2009. Functional Amyloids As Natural Storage of Peptide Hormones in Pituitary Secretory Granules. *Science*, 325, 328-332.
- MAKIN, O. S., SERPELL, L. C. 2005. Structures for amyloid proteins. *FEBS Journal*, 2005, 5950-5961.
- MANAVALAN, P. A. J., W.C., JR. 1987. Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Analytical Biochemistry*, 167, 76-85.
- MANENTI, L., TANSINDA, P. & VAGLIO, A. 2008. Eprodinate in amyloid A amyloidosis: a novel therapeutic approach? *Expert Opinion on Pharmacotherapy*, 9, 2175-2180.
- MANTYH, P. W., GHILARDI, J. R., ROGERS, S., DEMASTER, E., ALLEN, C. J., STIMSON, E. R. & MAGGIO, J. E. 1993. Aluminum, Iron, and Zinc Ions Promote Aggregation of Physiological Concentrations of  $\beta$ -Amyloid Peptide. *Journal of Neurochemistry*, 61, 1171-1174.
- MARION M, B. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72, 248-254.
- MCINTOSH, L. P. & DAHLQUIST, F. W. 1990. Biosynthetic Incorporation of  $^{15}\text{N}$  and  $^{13}\text{C}$  for Assignment and Interpretation of Nuclear Magnetic Resonance Spectra of Proteins. *Quarterly Reviews of Biophysics*, 23, 1-38.
- MENG, X. Y., FINK, A. L. & UVERSKY, V. N. 2008. The effect of membranes on the in vitro fibrillation of an amyloidogenic light-chain variable-domain SMA. *Journal of Molecular Biology*, 381, 989-999.
- MILANESI, L., SHEYNIS, T., XUE, W. F., ORLOVA, E. V., HELLEWELL, A. L., JELINEK, R., HEWITT, E. W., RADFORD, S. E. & SAIBIL, H. R. 2012. Direct three-dimensional visualization of membrane disruption by amyloid fibrils. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 20455-20460.
- MUCCHIANO, G., CORNWELL, G. G. & WESTERMARK, P. 1992. Senile aortic amyloid - evidence for 2 distinct forms of localized deposits. *American Journal of Pathology*, 140, 871-877.
- MUCCHIANO, G. I., JONASSON, L., HÄGGQVIST, B., EINARSSON, E. & WESTERMARK, P. 2001. Apolipoprotein A-I-Derived Amyloid in Atherosclerosis. *American Journal of Clinical Pathology*, 115, 298-303.
- MUNISHKINA, L. A. & FINK, A. L. 2007. Fluorescence as a method to reveal structures and membrane-interactions of amyloidogenic proteins. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1768, 1862-1885.

- NAIKI, H., HIGUCHI, K., HOSOKAWA, M. & TAKEDA, T. 1989. Fluorometric determination of amyloid fibrils in vitro using the fluorescent dye, thioflavine T. *Analytical Biochemistry*, 177, 244-249.
- NÄSLUND, J., HAROUTUNIAN, V., MOHS, R., DAVIS, K. L., DAVIES, P., GREENGARD, P. & BUXBAUM, J. D. 2000. Correlation Between Elevated Levels of Amyloid  $\beta$ -Peptide in the Brain and Cognitive Decline. *JAMA: The Journal of the American Medical Association*, 283, 1571-1577.
- NECULA, M., CHIRITA, C. N. & KURET, J. 2003. Rapid anionic micelle-mediated alpha-synuclein fibrillization in vitro. *Journal of Biological Chemistry*, 278, 46674-46680.
- NELSON, R. & EISENBERG, D. 2006. Recent atomic models of amyloid fibril structure. *Current Opinion in Structural Biology*, 16, 260-265.
- NELSON, R., SAWAYA, M. R., BALBIRNIE, M., MADSEN, A. O., RIEKEL, C., GROTHE, R. & EISENBERG, D. 2005. Structure of the cross-beta spine of amyloid-like fibrils. *Nature*, 435, 773-778.
- NEUBURG, D. S., PETERSON, J. A., RUIZ-PALACIOS, G. M., MATSON, D. O., MORROW, A. L., SHULTS, J., GUERRERO, M. D. L., CHATURVEDI, P., NEUBURG, S. O., SCALLAN, C. D., TAYLOR, M. R., CERIANI, R. L. & PICKERING, L. K. 1998. Role of human-milk lactadherin in protectoin against symptomatic rotavirus infection. *The Lancet*, 351, 1160-1164.
- NIELSEN, L., KHURANA, R., COATS, A., FROKJAER, S., BRANGE, J., VYAS, S., UVERSKY, V. N. & FINK, A. L. 2001. Effect of environmental factors on the kinetics of insulin fibril formation: Elucidation of the molecular mechanism. *Biochemistry*, 40, 6036-6046.
- O'NUALLAIN, B., WILLIAMS, A. D., WESTERMARK, P. & WETZEL, R. 2004. Seeding Specificity in Amyloid Growth Induced by Heterologous Fibrils. *Journal of Biological Chemistry*, 279, 17490-17499.
- OHISHI, H., SKINNER, M., SATO-ARAKI, N., OKUYAMA, T., GEJYO, F., KIMURA, A., COHEN, A. S. & SCHMID, K. 1990. Glycosaminoglycans of the hemodialysis-associated carpal synovial amyloid and of amyloid-rich tissues and fibrils of heart, liver, and spleen. *Clinical Chemistry*, 36, 88-91.
- OLSEN, K. E., SLETTEN, K. & WESTERMARK, P. 1998. Fragments of the Constant Region of Immunoglobulin Light Chains Are Constituents of AL-Amyloid Proteins. *Biochemical and Biophysical Research Communications*, 251, 642-647.
- ONUCHIC, J. N. & WOLYNES, P. G. 2004. Theory of protein folding. *Current Opinion in Structural Biology*, 14, 70-75.
- PAN, K. M., BALDWIN, M., NGUYEN, J., GASSET, M., SERBAN, A., GROTH, D., MEHLHORN, I., HUANG, Z. W., FLETTERICK, R. J., COHEN, F. E. & PRUSINER, S. B. 1993. CONVERSION OF ALPHA-HELICES INTO BETA-SHEETS FEATURES IN THE FORMATION OF THE SCRAPIE PRION PROTEINS. *Proceedings of the National Academy of Sciences of the United States of America*, 90, 10962-10966.
- PAPY-GARCIA, D., MORIN, C., HUYNH, M. B., SENERIZ, F., SISOEFF, L., SEPUVEDA-DIAZ, J. E. & RAISMAN-VOZARI, R. 2011. Glycosaminoglycans, Protein Aggregation and Neurodegeneration. *Current Protein & Peptide Science*, 12, 258-268.
- PARAVASTU, A. K., LEAPMAN, R. D., YAU, W. M. & TYCKO, R. 2008. Molecular structural basis for polymorphism in Alzheimer's  $\beta$ -amyloid fibrils.

- Proceedings of the National Academy of Sciences of the United States of America*, 105, 18349-18354.
- PARAVASTU, A. K., PETKOVA, A. T. & TYCKO, R. 2006. Polymorphic fibril formation by residues 10-40 of the Alzheimer's  $\beta$ -amyloid peptide. *Biophysical Journal*, 90, 4618-4629.
- PARAVASTU, A. K., QAHWASH, I., LEAPMAN, R. D., MEREDITH, S. C. & TYCKO, R. 2009. Seeded growth of  $\beta$ -amyloid fibrils from Alzheimer's brain-derived fibrils produces a distinct fibril structure. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 7443-7448.
- PEDERSEN, J. S., ANDERSEN, C. B. & OTZEN, D. E. 2010. Amyloid structure – one but not the same: the many levels of fibrillar polymorphism. *FEBS Journal*, 277, 4591-4601.
- PENG, S., GLENNERT, J. & WESTERMARK, P. 2005. Medin-amyloid: A recently characterized age-associated arterial amyloid form affects mainly arteries in the upper part of the body. *Amyloid*, 12, 96-102.
- PENG, S., LARSSON, A., WASSBERG, E., GERWINS, P., THELIN, S., FU, X. & WESTERMARK, P. 2007. Role of aggregated medin in the pathogenesis of thoracic aortic aneurysm and dissection. *Laboratory Investigation*, 87, 1195-1205.
- PENG, S., WESTERMARK, G. T., NÄSLUND, J., HÄGGQVIST, B., GLENNERT, J. & WESTERMARK, P. 2002. Medin and medin-amyloid in ageing inflamed and non-inflamed temporal arteries. *The Journal of Pathology*, 196, 91-96.
- PEPYS, M. B. 2001. Pathogenesis, diagnosis and treatment of systemic amyloidosis. *Philos Trans R Soc Lond B Biol Sci.*, 356, 203-211.
- PEPYS, M. B., BOOTH, D. R., HUTCHINSON, W. L., GALLIMORE, J. R., COLLINS, I. M. & HOHENESTER, E. 1997. Amyloid P component. A critical review. *Amyloid*, 4, 274-295.
- PEPYS, M. B., DYCK, R. F., DEBEER, F. C., SKINNER, M. & COHEN, A. S. 1979. BINDING OF SERUM AMYLOID P-COMPONENT (SAP) BY AMYLOID FIBRILS. *Clinical and Experimental Immunology*, 38, 284-293.
- PEPYS, M. B., HERBERT, J., HUTCHINSON, W. L., TENNENT, G. A., LACHMANN, H. J., GALLIMORE, J. R., LOVAT, L. B., BARTFAI, T., ALANINE, A., HERTEL, C., HOFFMANN, T., JAKOB-ROETNE, R., NORCROSS, R. D., KEMP, J. A., YAMAMURA, K., SUZUKI, M., TAYLOR, G. W., MURRAY, S., THOMPSON, D., PURVIS, A., KOLSTOE, S., WOOD, S. P. & HAWKINS, P. N. 2002. Targeted pharmacological depletion of serum amyloid P component for treatment of human amyloidosis. *Nature*, 417, 254-259.
- PERKINS, D. N., PAPPIN, D. J. C., CREASY, D. M. & COTTRELL, J. S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, 20, 3551-3567.
- PERRIN, R. J., FAGAN, A. M. & HOLTZMAN, D. M. 2009. Multimodal techniques for diagnosis and prognosis of Alzheimer's disease. *Nature*, 461, 916-922.
- PERTINHEZ, T. A., BOUCHARD, M., TOMLINSON, E. J., WAIN, R., FERGUSON, S. J., DOBSON, C. M. & SMITH, L. J. 2001. Amyloid fibril formation by a helical cytochrome. *FEBS Letters*, 495, 184-186.

- PETKOVA, A. T., BUNTKOWSKY, G., DYDA, F., LEAPMAN, R. D., YAU, W. M. & TYCKO, R. 2004. Solid state NMR reveals a pH-dependent antiparallel  $\beta$ -sheet registry in fibrils formed by a  $\beta$ -amyloid peptide. *Journal of Molecular Biology*, 335, 247-260.
- PETKOVA, A. T., ISHII, Y., BALBACH, J. J., ANTZUTKIN, O. N., LEAPMAN, R. D., DELAGLIO, F. & TYCKO, R. 2002. A structural model for Alzheimer's  $\beta$ -amyloid fibrils based on experimental constraints from solid state NMR. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 16742-16747.
- PETKOVA, A. T., LEAPMAN, R. D., GUO, Z. H., YAU, W. M., MATTSON, M. P. & TYCKO, R. 2005a. Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils. *Science*, 307, 262-265.
- PETKOVA, A. T., YAU, W.-M. & TYCKO, R. 2005b. Experimental Constraints on Quaternary Structure in Alzheimer's  $\beta$ -Amyloid Fibrils†. *Biochemistry*, 45, 498-512.
- PITKÄNEN, P., WESTERMARK, PER, CORNWELL III, GIBBONS G. 1984. Senile systemic amyloidosis. *American Journal of Pathology*, 117, 391-399.
- PLATT, G. W., ROUTLEDGE, K. E., HOMANS, S. W. & RADFORD, S. E. 2008. Fibril Growth Kinetics Reveal a Region of  $\beta$ 2-microglobulin Important for Nucleation and Elongation of Aggregation. *Journal of Molecular Biology*, 378, 251-263.
- PRADO, M. A. M. & BARON, G. 2012. Seeding plaques in Alzheimer's disease. *Journal of Neurochemistry*, 120, 641-643.
- PRAS, M. S., M. ZUCKER-FRANKLIN, D. RIMON, A. AND FRANKLIN, E. C. 1968. The characterisation of soluble amyloid prepared in water. *J. Clin. Invest*, 47, 924-933.
- PUCHTLER, H., SWEAT, F. & LEVINE, M. 1962. On the binding of Congo Red. *Journal of Histochemistry & Cytochemistry*, 10, 355-364.
- QIANG, W., YAU, W.-M., LUO, Y., MATTSON, M. P. & TYCKO, R. 2012. Antiparallel  $\beta$ -sheet architecture in Iowa-mutant  $\beta$ -amyloid fibrils. *Proceedings of the National Academy of Sciences*, 109, 4443-4448.
- QIANG, W., YAU, W.-M. & TYCKO, R. 2011. Structural Evolution of Iowa Mutant  $\beta$ -Amyloid Fibrils from Polymorphic to Homogeneous States under Repeated Seeded Growth. *Journal of the American Chemical Society*, 133, 4018-4029.
- QIN, Z. J., HU, D. M., ZHU, M. & FINK, A. L. 2007. Structural characterization of the partially folded intermediates of an immunoglobulin light chain leading to amyloid fibrillation and amorphous aggregation. *Biochemistry*, 46, 3521-3531.
- RECHES, M. & GAZIT, E. 2004. Amyloidogenic hexapeptide fragment of medin: Homology to functional islet amyloid polypeptide fragments. *Amyloid*, 11, 81-89.
- RENNER, M., LACOR, P. N., VELASCO, P. T., XU, J., CONTRACTOR, A., KLEIN, W. L. & TRILLER, A. 2010. Deleterious Effects of Amyloid  $\beta$  Oligomers Acting as an Extracellular Scaffold for mGluR5. *Neuron*, 66, 739-754.
- ROSEN, R. F., FRITZ, J. J., DOOYEMA, J., CINTRON, A. F., HAMAGUCHI, T., LAH, J. J., LEVINE, H., JUCKER, M. & WALKER, L. C. 2012. Exogenous seeding of cerebral beta-amyloid deposition in beta APP-transgenic rats. *Journal of Neurochemistry*, 120, 660-666.

- SACKEWITZ, M., SCHEIDT, H. A., LODDERSTEDT, G., SCHIERHORN, A., SCHWARZ, E. & HUSTER, D. 2008. Structural and dynamical characterization of fibrils from a disease-associated alanine expansion domain using proteolysis and solid-state NMR spectroscopy. *Journal of the American Chemical Society*, 130, 7172-+.
- SAMBROOK, J., FRITSCH, E. F., MANIATIS, T. 1989. Molecular Cloning - A Laboratory Handbook. In: NOLAN, C. (ed.) 2 ed.: Cold Spring Harbour Laboratory Press.
- SANCHORAWALA, V. 2006. Light-Chain (AL) Amyloidosis: Diagnosis and Treatment. *Clinical Journal of the American Society of Nephrology*, 1, 1331-1341.
- SASAHARA, K., YAGI, H., SAKAI, M., NAIKI, H. & GOTO, Y. 2008. Amyloid Nucleation Triggered by Agitation of  $\beta$ 2-Microglobulin under Acidic and Neutral pH Conditions. *Biochemistry*, 47, 2650-2660.
- SAWAYA, M. R., SAMBASHIVAN, S., NELSON, R., IVANOVA, M. I., SIEVERS, S. A., APOSTOL, M. I., THOMPSON, M. J., BALBIRNIE, M., WILTZIUS, J. J. W., MCFARLANE, H. T., MADSEN, A. O., RIEKEL, C. & EISENBERG, D. 2007. Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature*, 447, 453-457.
- SCHÄGGER, H. & VON JAGOW, G. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry*, 166, 368-379.
- SCHEIDT, H. A., MORGADO, I., ROTHEMUND, S., HUSTER, D. & FAENDRICH, M. 2011. Solid-State NMR Spectroscopic Investigation of A beta Protofibrils: Implication of a beta-Sheet Remodeling upon Maturation into Terminal Amyloid Fibrils. *Angewandte Chemie-International Edition*, 50, 2837-2840.
- SCHUBERT, D., BEHL, C., LESLEY, R., BRACK, A., DARGUSCH, R., SAGARA, Y. & KIMURA, H. 1995. Amyloid peptides are toxic via a common oxidative mechanism. *Proceedings of the National Academy of Sciences*, 92, 1989-1993.
- SCHWARTZ, P. (ed.) 1970. *Cardiovascular amyloidosis - amyloidosis, cause and manifestation of senile deterioration*: Springfield.
- SCIARRETTA, K. L., GORDON, D. J., PETKOVA, A. T., TYCKO, R. & MEREDITH, S. C. 2005. A beta 40-Lactam(D23/K28) models a conformation highly favorable for nucleation of amyloid. *Biochemistry*, 44, 6003-6014.
- SERIO, T. R., CASHIKAR, A. G., KOWAL, A. S., SAWICKI, G. J., MOSLEHI, J. J., SERPELL, L., ARNSDORF, M. F. & LINDQUIST, S. L. 2000. Nucleated Conformational Conversion and the Replication of Conformational Information by a Prion Determinant. *Science*, 289, 1317-1321.
- SERPELL, L. C., SUNDE, M., BENSON, M. D., TENNENT, G. A., PEPYS, M. B. & FRASER, P. E. 2000. The protofilament substructure of amyloid fibrils. *Journal of Molecular Biology*, 300, 1033-1039.
- SHAO, C., NOVAKOVIC, V. A., HEAD, J. F., SEATON, B. A. & GILBERT, G. E. 2008. Crystal Structure of Lactadherin C2 Domain at 1.7Å Resolution with Mutational and Computational Analyses of Its Membrane-binding Motif. *Journal of Biological Chemistry*, 283, 7230-7241.

- SHI, J. & GILBERT, G. E. 2003. Lactadherin inhibits enzyme complexes of blood coagulation by competing for phospholipid-binding sites. *Blood*, 101, 2628-2636.
- SIPE, J. D., BENSON, M. D., BUXBAUM, J. N., IKEDA, S.-I., MERLINI, G., SARAIVA, M. J. M. & WESTERMARK, P. 2012. Amyloid fibril protein nomenclature: 2012 recommendations from the Nomenclature Committee of the International Society of Amyloidosis. *Amyloid-Journal of Protein Folding Disorders*, 19, 167-170.
- SKLENAR, V., PIOTTO, M., LEPIK, R. & SAUDEK, V. 1993. Gradient-Tailored Water Suppression for 1H-15N HSQC Experiments Optimized to Retain Full Sensitivity. *Journal of Magnetic Resonance, Series A*, 102, 241-245.
- SOLOMON, J. P., PAGE, L. J., BALCH, W. E. & KELLY, J. W. 2012. Gelsolin amyloidosis: genetics, biochemistry, pathology and possible strategies for therapeutic intervention. *Critical Reviews in Biochemistry and Molecular Biology*, 47, 282-296.
- SOTO, C., SIGURDSSON, E. M., MORELLI, L., ASOK KUMAR, R., CASTANO, E. M. & FRANGIONE, B. 1998. [beta]-sheet breaker peptides inhibit fibrillogenesis in a rat brain model of amyloidosis: Implications for Alzheimer's therapy. *Nat Med*, 4, 822-826.
- SREERAMA, N. A. W., R.W 2000. Estimation of protein secondary structure from CD spectra: Comparison of CONTIN, SELCON and CDSSTR methods with an expanded reference set. *Analytical Biochemistry*, 287, 252-260.
- STEFANI, M. 2007. Generic Cell Dysfunction in Neurodegenerative Disorders: Role of Surfaces in Early Protein Misfolding, Aggregation, and Aggregate Cytotoxicity. *The Neuroscientist*, 13, 519-531.
- STEVENS, F. J. & KISILEVSKY, R. 2000. Immunoglobulin light chains, glycosaminoglycans, and amyloid. *Cellular and Molecular Life Sciences*, 57, 441-449.
- STEVENS, F. J., MYATT, E. A., CHANG, C. H., WESTHOLM, F. A., EULITZ, M., WEISS, D. T., MURPHY, C., SOLOMON, A. & SCHIFFER, M. 1995a. A molecular-model for the self-assembly of amyloid fibrils-immunoglobulin light chains. *Biochemistry*, 34, 10697-10702.
- STEVENS, P. W., RAFFEN, R., HANSON, D. K., DENG, Y. L., BERRIOSHAMMOND, M., WESTHOLM, F. A., MURPHY, C., EULITZ, M., WETZEL, R., SOLOMON, A., SCHIFFER, M. & STEVENS, F. J. 1995b. Recombinant immunoglobulin variable domains generated from synthetic genes provide a system for in vitro characterisation of light-chain amyloid proteins. *Protein Science*, 4, 421-432.
- STEWARD, R. E., ARMEN, R. S. & DAGGETT, V. 2008. Different disease-causing mutations in transthyretin trigger the same conformational conversion. *Protein Engineering Design and Selection*, 21, 187-195.
- STREETS, A. M., SOURIGUES, Y., KOPITO, R. R., MELKI, R. & QUAKE, S. R. 2013. Simultaneous Measurement of Amyloid Fibril Formation by Dynamic Light Scattering and Fluorescence Reveals Complex Aggregation Kinetics. *PLoS ONE*, 8, e54541.
- SUNDE, M. & BLAKE, C. C. F. 1998. From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. *Quarterly Reviews of Biophysics*, 31, 1-39.

- SUNDE, M., SERPELL, L. C., BARTLAM, M., FRASER, P. E., PEPYS, M. B. & BLAKE, C. C. F. 1997. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of Molecular Biology*, 273, 729-739.
- TAKEGOSHI, K., NAKAMURA, S. & TERAOKA, T. 2001. <sup>13</sup>C-<sup>1</sup>H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chemical Physics Letters*, 344, 631-637.
- TANAKA, M., CHIEN, P., NABER, N., COOKE, R. & WEISSMAN, J. S. 2004. Conformational variations in an infectious protein determine prion strain differences. *Nature*, 428, 323-328.
- TARTAGLIA, G. G. & VENDRUSCOLO, M. 2008. The Zyggregator method for predicting protein aggregation propensities. *Chemical Society Reviews*, 37, 1395-1401.
- TOMASELLI, S., ESPOSITO, V., VANGONE, P., VAN NULAND, N. A. J., BONVIN, A. M. J. J., GUERRINI, R., TANCREDI, T., TEMUSSI, P. A. & PICONE, D. 2006. The  $\alpha$ -to- $\beta$  Conformational Transition of Alzheimer's A $\beta$ -(1-42) Peptide in Aqueous Media is Reversible: A Step by Step Conformational Analysis Suggests the Location of  $\beta$  Conformation Seeding. *Chembiochem*, 7, 257-267.
- TOMITA, S., KIRINO, Y. & SUZUKI, T. 1998. Cleavage of Alzheimer's Amyloid Precursor Protein (APP) by Secretases Occurs after O-Glycosylation of APP in the Protein Secretory Pathway. *Journal of Biological Chemistry*, 273, 6277-6284.
- TORIZAWA, T., SHIMIZU, M., TAOKA, M., MIYANO, H. & KAINOSHO, M. 2004. Efficient production of isotopically labeled proteins by cell-free synthesis: A practical protocol. *Journal of Biomolecular NMR*, 30, 311-325.
- TROVATO, A., CHITI, F., MARITAN, A. & SENO, F. 2006. Insight into the Structure of Amyloid Fibrils from the Analysis of Globular Proteins. *PLoS Comput Biol*, 2, e170.
- TROVATO, A., SENO, F. & TOSATTO, S. C. E. 2007. The PASTA server for protein aggregation prediction. *Protein Eng. Des. Sel.*, 521-523.
- TU, L.-H. & RALEIGH, D. P. 2013. Role of Aromatic Interactions in Amyloid Formation by Islet Amyloid Polypeptide. *Biochemistry*, 52, 333-342.
- TYCKO, R. 2006a. Characterization of Amyloid Structures at the Molecular Level by Solid State Nuclear Magnetic Resonance Spectroscopy.
- TYCKO, R. 2006b. Solid-state NMR as a probe of amyloid structure. *Protein and Peptide Letters*, 13, 229-234.
- TYCKO, R. 2011. Solid-State NMR Studies of Amyloid Fibril Structure. In: LEONE, S. R., CREMER, P. S., GROVES, J. T. & JOHNSON, M. A. (eds.) *Annual Review of Physical Chemistry*, Vol 62. Palo Alto: Annual Reviews.
- TYCKO, R., SCIARRETTA, K. L., ORGEL, J. P. R. O. & MEREDITH, S. C. 2009. Evidence for novel  $\beta$ -sheet structures in Iowa mutant  $\beta$ -amyloid fibrils. *Biochemistry*, 48, 6072-6084.
- ULRICH, E. L., AKUTSU, H., DORELEIJERS, J. F., HARANO, Y., IOANNIDIS, Y. E., LIN, J., LIVNY, M., MADING, S., MAZIUK, D., MILLER, Z., NAKATANI, E., SCHULTE, C. F., TOLMIE, D. E., WENGER, R. K., YAO, H. & MARKLEY, J. L. 2008. BioMagResBank. *Nucleic Acids Research*, 36, D402-D408.
- VAN MELCKEBEKE, H. L. N., WASMER, C., LANGE, A., AB, E., LOQUET, A., BÖCKMANN, A. & MEIER, B. H. 2010. Atomic-Resolution Three-



- Dimensional Structure of HET-s(218–289) Amyloid Fibrils by Solid-State NMR Spectroscopy. *Journal of the American Chemical Society*, 132, 13765-13775.
- VAN NOSTRAND, W. E., MELCHOR, J. P., CHO, H. S., GREENBERG, S. M. & REBECK, G. W. 2001. Pathogenic Effects of D23N Iowa Mutant Amyloid  $\beta$ -Protein. *Journal of Biological Chemistry*, 276, 32860-32866.
- VAN STOKKUM, I. H. M., SPOELDER, H.J.W., BLOEMENDAL, M., VAN GRONDELLE, R., AND GROEN, F.C.A. 1990. Estimation of protein secondary structure and error analysis from CD spectra. *Analytical Biochemistry*, 191, 110-118.
- VASSAR, P. S. & CULLING, C. F. A. 1959. FLUORESCENT STAINS, WITH SPECIAL REFERENCE TO AMYLOID AND CONNECTIVE TISSUES. *Archives of Pathology*, 68, 487-498.
- VIRCHOW, R. 1857. Neue beobachtungen uber amyloid degeneration. *Virchows Archiv fur pathologische Anatomie und Physiologie und fur klinische Medizin*, 11, 188-189.
- VIVIAN, J. T. & CALLIS, P. R. 2001. Mechanisms of Tryptophan Fluorescence Shifts in Proteins. *Biophysical Journal*, 80, 2093-2109.
- VRIEND, G. 1990. WHAT IF: A molecular modeling and drug design program. *Journal of Molecular Graphics*, 8, 52-56.
- WALLACE, B. A. 2009. Protein characterisation by synchrotron radiation circular dichroism spectroscopy. *Quarterly Reviews of Biophysics*, 42, 317-370.
- WALSH, D. M., THULIN, E., MINOGUE, A. M., GUSTAVSSON, N., PANG, E., TEPLow, D. B. & LINSE, S. 2009. A facile method for expression and purification of the Alzheimer's disease-associated amyloid  $\beta$ -peptide. *FEBS Journal*, 276, 1266-1281.
- WANG, J., DICKSON, D. W., TROJANOWSKI, J. Q., LEE V. M. 1999. The levels of soluble versus insoluble brain Abeta distinguish Alzheimer's disease from normal and pathologic aging. *Exp Neurol*, 158, 328-337.
- WANG, Y. & JARDETZKY, O. 2002. Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Science*, 11, 852-861.
- WASMER, C., LANGE, A., VAN MELCKEBEKE, H., SIEMER, A. B., RIEK, R. & MEIER, B. H. 2008. Amyloid Fibrils of the HET-s(218–289) Prion Form a  $\beta$  Solenoid with a Triangular Hydrophobic Core. *Science*, 319, 1523-1526.
- WAUGH, D. S. 2005. Making the most of affinity tags. *Trends in Biotechnology*, 23, 316-320.
- WECHALEKAR, A. D., HAWKINS, P. N. & GILLMORE, J. D. 2008. Perspectives in treatment of AL amyloidosis. *British Journal of Haematology*, 140, 365-377.
- WESTERMARK, G. T. & WESTERMARK, P. 2011. Localized amyloids important in diseases outside the brain – lessons from the islets of Langerhans and the thoracic aorta. *FEBS Journal*, 278, 3918-3929.
- WESTERMARK, P., SLETTEN, K., JOHANSSON, B. & CORNWELL, G. G. 1990. Fibril in senile systemic amyloidosis is derived from normal transthyretin. *Proceedings of the National Academy of Sciences*, 87, 2843-2845.

- WHITMORE, L. & WALLACE, B. A. 2008. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers*, 89, 392-400.
- WILLIAMS, T. L. & SERPELL, L. C. 2011. Membrane and surface interactions of Alzheimer's A $\beta$  peptide – insights into the mechanism of cytotoxicity. *FEBS Journal*, 278, 3905-3917.
- WILLIAMSON, J. A., LORIA, J. P. & MIRANKER, A. D. 2009. Helix Stabilization Precedes Aqueous and Bilayer-Catalyzed Fiber Formation in Islet Amyloid Polypeptide. *Journal of Molecular Biology*, 393, 383-396.
- WISHART, D. & SYKES, B. 1994. The <sup>13</sup>C Chemical-Shift Index: A simple method for the identification of protein secondary structure using <sup>13</sup>C chemical-shift data. *Journal of Biomolecular NMR*, 4, 171-180.
- XUE, W.-F., HOMANS, S. W. & RADFORD, S. E. 2008. Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. *Proceedings of the National Academy of Sciences*, 105, 8926-8931.
- YAMAGUCHI, K., TAKAHASHI, S., KAWAI, T., NAIKI, H. & GOTO, Y. 2005. Seeding-dependent propagation and maturation of amyloid fibril conformation. *Journal of Molecular Biology*, 352, 952-960.
- YANAMANDRA, K., GRUDEN, M. A., CASAITE, V., MESKYS, R., FORSGREN, L. & MOROZOVA-ROCHE, L. A. 2011. alpha-Synuclein Reactive Antibodies as Diagnostic Biomarkers in Blood Sera of Parkinson's Disease Patients. *PLoS ONE*, 6.
- YANG, Z., LASKER, K., SCHNEIDMAN-DUHOVNY, D., WEBB, B., HUANG, C. C., PETTERSEN, E. F., GODDARD, T. D., MENG, E. C., SALI, A. & FERRIN, T. E. 2012. UCSF Chimera, MODELLER, and IMP: An integrated modeling system. *Journal of Structural Biology*, 179, 269-278.
- YAU, J. & SHARPE, S. 2012. Structures of amyloid fibrils formed by the prion protein derived peptides PrP(244–249) and PrP(245–250). *Journal of Structural Biology*, 180, 290-302.
- ZANIER, K., NOMINÉ, Y., CHARBONNIER, S., RUHLMANN, C., SCHULTZ, P., SCHWEIZER, J. & TRAVÉ, G. 2007. Formation of well-defined soluble aggregates upon fusion to MBP is a generic property of E6 proteins from various human papillomavirus species. *Protein Expression and Purification*, 51, 59-70.
- ZHAO, L. N., LONG, H. W., MU, Y. G. & CHEW, L. Y. 2012. The Toxicity of Amyloid beta Oligomers. *International Journal of Molecular Sciences*, 13, 7303-7327.
- ZHU, M., SOUILLAC, P. O., IONESCU-ZANETTI, C., CARTER, S. A. & FINK, A. L. 2002. Surface-catalyzed amyloid fibril formation. *Journal of Biological Chemistry*, 277, 50914-50922.

## Appendix 1

### Cell growth media and buffers

#### *Luria Bertani (LB)*

Component	Final amount in 1L (g)
Tryptone	10
Yeast Extract	5
NaCl	10

#### *Terrific Broth (TB)*

Solution A – pH 7.4 and autoclaved

Component	Final amount in 900 mL
Tryptone	12 g
Yeast Extract	24 g
Glycerol	4 mL

Solution B – pH 7.4 and filter sterilised and added to solution A.

Component	Final concentration in 100 mL (M)
Na <sub>2</sub> HPO <sub>4</sub>	0.17
KH <sub>2</sub> PO <sub>4</sub>	0.72

#### *Minimal media*

Solution A – pH 7.2 and autoclaved

Component	Final concentration in 1L (mM)
Na <sub>2</sub> HPO <sub>4</sub>	88
KH <sub>2</sub> PO <sub>4</sub>	55

Solution B – filter sterilised and added to solution A.

For isotopically labelled samples  $^{13}\text{C}$ -glucose and  $^{15}\text{NH}_4\text{Cl}$  are used instead.

<b>Component</b>	<b>Final concentration in 1L (mM)</b>
<b>Glucose</b>	22
<b>NH<sub>4</sub>Cl</b>	19
<b>MgSO<sub>4</sub></b>	1
<b>CaCl<sub>2</sub>·H<sub>2</sub>O</b>	0.1
<b>Kanamycin/Ampicillin</b>	0.3
<b>Thymine HCl</b>	0.03

*Super optimal broth (SOB)*

<b>Component</b>	<b>Final concentration in 1L</b>
<b>Tryptone</b>	20 g
<b>Yeast Extract</b>	1 g
<b>NaCl</b>	100 mM
<b>KCl</b>	2.5 mM

*Super optimal broth with catabolite repression (SOC)*

<b>Component</b>	<b>Final concentration in 1L</b>
<b>Tryptone</b>	20 g
<b>Yeast Extract</b>	1 g
<b>NaCl</b>	100 mM
<b>KCl</b>	2.5 mM
<b>Glucose</b>	20 mM

*Luria Betani –Agar*

<b>Component</b>	<b>Final amount in 0.5 L (g)</b>
<b>Tryptone</b>	5
<b>Yeast Extract</b>	2.5
<b>NaCl</b>	5
<b>Agar</b>	7.5
<b>Kanamycin/ampicillin</b>	0.03

*CMB80 buffer*

<b>Component</b>	<b>Final concentration in 1L (mM)</b>
<b>KOAc</b>	10
<b>CaCl<sub>2</sub></b>	80
<b>MnCl<sub>2</sub></b>	20
<b>MgCl<sub>2</sub></b>	10
<b>Glycerol</b>	10 % v/v

**SDS-PAGE, Tris-tricine gel and Western blot buffers**

*10 x SDS-PAGE Running buffer*

<b>Component</b>	<b>Final concentration in 1L</b>
<b>Tris HCl</b>	0.25 M
<b>Glycine</b>	1.9 M
<b>Sodium dodecyl sulphate (SDS)</b>	35 mM (1 % v/v)

*SDS-PAGE Sample buffer*

Component	Final concentration in 10 mL
Tris HCl	50 mM
Glycerol	10 % v/v
SDS	70 mM
Dithiothreitol	0.01 mM
Bromophenol blue	1.5 mM

*SDS-PAGE gel composition*

Component	Stacking gel 4% (ml)	Resolving gel 15 % (ml)
30 % Acrylamide	1.3	5
Tris base 1.5 M pH 8.8	2.5	-
Tris base 0.5 M pH 6.8	-	2.5
ddH <sub>2</sub> O	6.1	
SDS	0.1	0.1
TEMED	7.5 µl	7.5 µl
Ammonium persulphate (APS) (10%)	75 µl	75 µl

*Tris-Tricine gel buffer – pH 8.45*

Component	Final concentration in 1L
Tris base	3 M
Sodium dodecyl sulphate (SDS)	10 mM (0.3 % v/v)

*Tris-Tricine SDS-PAGE gel composition*

<b>Component</b>	<b>Stacking gel 4% (ml)</b>	<b>Resolving gel 15% (ml)</b>
<b>30 % Acrylamide</b>	1.3	5
<b>Tris gel buffer</b>	2.5	3.3
<b>ddH<sub>2</sub>O</b>	6.2	0.7
<b>Glycerol</b>	-	1
<b>TEMED</b>	7.5 µl	7.5 µl
<b>Ammonium persulphate (APS) (10%)</b>	75 µl	75 µl

*5 x Inner cathode buffer – pH 8.25*

<b>Component</b>	<b>Final concentration in 1L</b>
<b>Tris base</b>	0.5 M
<b>Tricine</b>	0.5 M
<b>SDS</b>	0.5 %

*5 x Outer cathode buffer – pH 8.9*

<b>Component</b>	<b>Final concentration in 1L</b>
<b>Tris base</b>	1 M

*Tris-tricine sample buffer*

<b>Component</b>	<b>Final concentration in 1L</b>
<b>Tris base</b>	0.5 M
<b>Tricine</b>	0.5 M
<b>SDS</b>	0.5 %

*Coomassie stain*

<b>Component</b>	<b>Final concentration in 1L</b>
<b>Coomassie Blue R250</b>	0.5 % w/v
<b>Acetic acid</b>	10 % v/v
<b>Methanol</b>	40 % v/v
<b>dd H<sub>2</sub>O</b>	50 % v/v

*Coomassie destain solution*

<b>Component</b>	<b>Final concentration in 1L (% v/v)</b>
<b>Acetic acid</b>	10
<b>Methanol</b>	40
<b>dd H<sub>2</sub>O</b>	50

*10 x Western blot transfer buffer*

<b>Component</b>	<b>Final concentration in 1L</b>
<b>Tris HCl</b>	0.25 M
<b>Glycine</b>	1.9 M
<b>Sodium dodecyl sulphate (SDS)</b>	7 mM (0.2 % w/v)

*1 x Western blot transfer buffer*

<b>Component</b>	<b>Final concentration in 1L (% v/v)</b>
<b>10 x Transfer buffer</b>	10
<b>Methanol</b>	20
<b>dd H<sub>2</sub>O</b>	70



*10 x Phosphate buffered saline (PBS) – pH 7.4*

<b>Component</b>	<b>Final concentration in 1L (M)</b>
<b>NaCl</b>	1.36
<b>KCl</b>	0.026
<b>Na<sub>2</sub>HPO<sub>4</sub></b>	0.1
<b>KH<sub>2</sub>PO<sub>4</sub></b>	0.017

*10 x Tris-acetate-EDTA buffer*

<b>Component</b>	<b>Final concentration in 1L</b>
<b>Tris base</b>	48.4 g
<b>Glacial acetic acid</b>	11.4 ml
<b>EDTA</b>	3.7 g

**Medin purification buffers**

*Amylose binding buffer - pH 7.4*

<b>Component</b>	<b>Final concentration in 1L (mM)</b>
<b>Tris-HCl</b>	50
<b>NaCl</b>	150
<b>EDTA</b>	1

*Amylose elution buffer – pH 7.4*

<b>Component</b>	<b>Final concentration in 1L (mM)</b>
<b>Tris-HCl</b>	50
<b>NaCl</b>	150
<b>EDTA</b>	1
<b>Maltose</b>	50

*His-Trap A – pH 7.4*

Component	Final concentration in 1L (mM)
NaPO <sub>4</sub>	20
NaCl	500

*His-Trap B – pH 7.4*

Component	Final concentration in 1L (mM)
NaPO <sub>4</sub>	20
NaCl	500
Imidazole	500

*Aggregation buffer – pH 7.4*

Component	Final concentration in 1L (mM)
NaPO <sub>4</sub>	20
NaCl	150

**SMA purification buffers**

*IE buffer A – pH 8*

Component	Final concentration in 1L (mM)
Tris-HCl	50
NaCl	10
EDTA	1

*IE buffer B – pH 8*

Component	Final concentration in 1L (mM)
Tris-HCl	50
NaCl	500
EDTA	1

*IE buffer C–pH 5.6*

<b>Component</b>	<b>Final concentration in 1L (mM)</b>
NaAc	50
NaCl	10
EDTA	1

*IE buffer D–pH 5.6*

<b>Component</b>	<b>Final concentration in 1L (mM)</b>
NaAc	50
NaCl	500
EDTA	1

*Gel filtration buffer –pH 8*

<b>Component</b>	<b>Final concentration in 1L (mM)</b>
Tris-HCl	50
NaCl	150
EDTA	1

**SMA aggregation buffers**

*pH 2*

<b>Component</b>	<b>Final concentration in 1L (mM)</b>
HCl	20

*pH 5.6*

<b>Component</b>	<b>Final concentration in 1L (mM)</b>
NaAc	50
NaCl	100

*pH 7.4*

<b>Component</b>	<b>Final concentration in 1L (mM)</b>
<b>NaPhos</b>	20
<b>NaCl</b>	150

## Appendix 2

### Site directed ligase independant mutagenesis (SLIM) primers

#### D25N

Forward long:

GGT TCC TCC ATC GAA GTT ACC GGT ATC ATC ACC

Forward short:

C GGT ATC ATC ACC CAG GGT G

Reverse long:

GTA ACT TCG ATG GAG GAA CCC AGG TCA AC

Reverse short:

CAG GTC AAC CTG CAG CCA CTG

#### K30I

Forward long:

CAG GTT AAC CTG GGT TCC TCC AAA GAA GTT ACC G

Short forward:

C AAA GAA GTT ACC GGT ATC ATC ACC

Reverse long:

GAG GAA CCC AGG TTA ACC TGC AGC CAC TG

Reverse short:

CAG CCA CTG GTC GTT ACC GTA G

### Sequencing data

Sequencing was performed by GATC Biotech (London, U.K.). Following amplification in XL1 blue *E.Coli* cells, plasmids were purified using a mini-plasmid prep kit (Sigma Aldrich) and prepared to a concentration of approximately 80 ng/μl.

Single read sequencing was performed using T7 forward and reverse primers.

#### D25N

Template Sequencing MGSSHHHHHGGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLME 60  
MGSSHHHHHGGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLME 60  
\*\*\*\*\*

Template Sequencing AFAKRQKGEMDSLRFlyDGIRIQADQTPEDLDMEDNDIIEAHREQIGGRLDKQGNFNAWV 120  
AFAKRQKGEMDSLRFlyDGIRIQADQTPEDLDMEDNDIIEAHREQIGGRLDKQGNFNAWV 120  
\*\*\*\*\*

Template Sequencing AGSYGNDQWLQVNLGSSKEVTGIIITQGARNFGSVQFVA 158  
AGSYGNDQWLQVNLGSSKEVTGIIITQGARNFGSVQFVA 158  
\*\*\*\*\*

## K30I

Template Sequencing MGSSHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLME 60  
MGSSHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLME 60  
\*\*\*\*\*

Template Sequencing AFAKRQKEMDSLRFlyDGIRIQADQTPEDLDMEDNDIEAHREQIGGRLDKQGNFAWV 120  
AFAKRQKEMDSLRFlyDGIRIQADQTPEDLDMEDNDIEAHREQIGGRLDKQGNFAWV 120  
\*\*\*\*\*

Template Sequencing AGSYGNDQWLQVDLGSSIEVTGIITQGARNFGSVQFVA 158  
AGSYGNDQWLQVDLGSSIEVTGIITQGARNFGSVQFVA 158  
\*\*\*\*\*

## SMA

Sequencing SMA MGSSHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLME 60  
MGSSHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLME 60  
\*\*\*\*\*

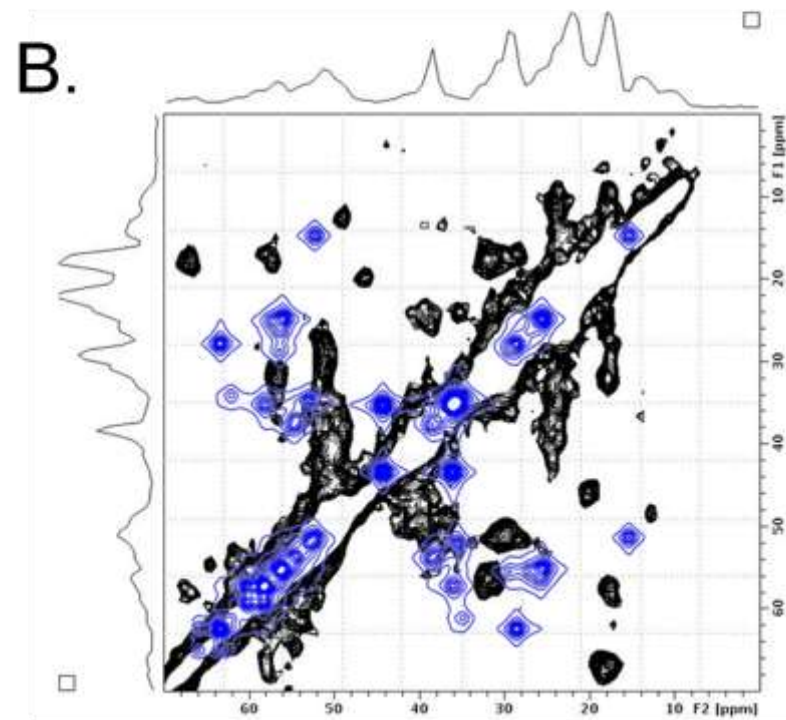
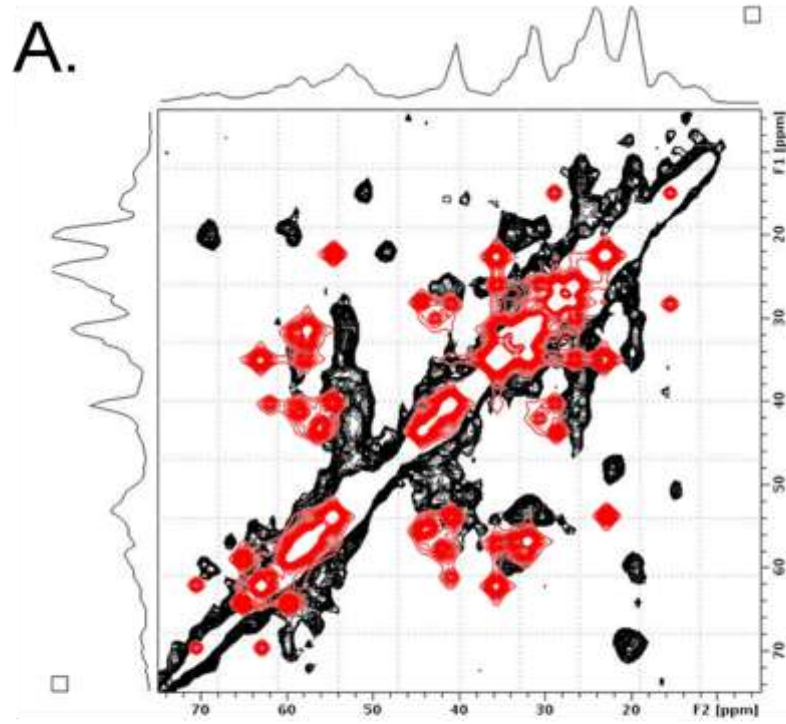
Sequencing SMA AFAKRQKEMDSLRFlyDGIRIQADQTPEDLDMEDNDIEAHREQIGGADIVMTQSPDSL 120  
AFAKRQKEMDSLRFlyDGIRIQADQTPEDLDMEDNDIEAHREQIGG-DIVMTQSPDSL 119  
\*\*\*\*\*

Sequencing SMA AVSLGERATINCKSSQSVLYSSNNRNYLAWYQKLGQPPKLLIYWASTRESGVPDRFSGS 180  
AVSLGERATINCKSSQSVLYSSNNRNYLAWYQKLGQPPKLLIYWASTRESGVPDRFSGS 179  
\*\*\*\*\*

Sequencing SMA GSGTDFTLTISSLQAEDVAVYYCHQYYSHPQTFGQGTKLELKR 223  
GSGTDFTLTISSLQAEDVAVYYCHQYYSHPQTFGQGTKLELKR 222  
\*\*\*\*\*

### Appendix 3

Medin DARR spectra overlaid with simulated spectra for random coil (A) and  $\alpha$ -helix (B) conformations.



## Appendix 4

### Alignment file:

```
>P1;2beg
structureX:2beg:17      :D :42  : :Abeta:: 1.90: 0.19
LVFFAEDVGSNKGAIIGLMVGGVVIA*
```

```
>P1;1med
sequence:1med:20      : :50   : :medin:: 2.00:-1.00
DQWLQVDLGSSK-EVTGIITQGARNF*
```

### Input file:

```
# Homology modeling by the automodel class
from modeller import *          # Load standard Modeller
classes
from modeller.automodel import * # Load the automodel class

log.verbose() # request verbose output
env = environ() # create a new MODELLER environment to build
this model in

# directories for input atom files
env.io.atom_files_directory = './:../atom_files'

# Read in HETATM records from template PDBs
env.io.hetatm = False

a = automodel(env,
              alnfile = 'med.ali', # alignment filename
              knowns   = '2beg',   # codes of the
templates
              sequence = '1med')   # code of the
target
a.starting_model= 1 # index of the first model
a.ending_model  = 6 # index of the last model
# (determines how many
models to calculate)
a.make() # do the actual homology
modeling
```