A Chemical, Structural and Biophysical Exploration into the Nanoscale Properties of Amyloid Fibrils

David Beal
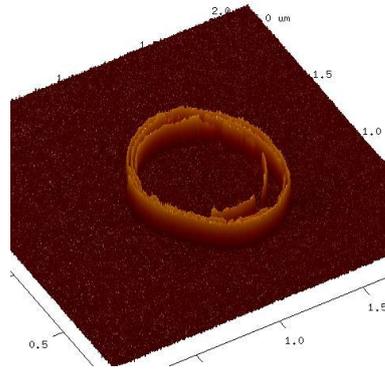
Thesis submitted in partial fulfilment of the requirements for the degree of PhD in Biochemistry

243 pages

University of Kent

Declaration

No part of this thesis has been submitted in support of any other application for a degree

or qualification of the University of Kent or any other University or institution of learning.

Dave Beal

September 2015

Acknowledgements

Table of Contents

Table of Figures

Abbreviations

ADC	Antibody drug conjugate
AFM	Atomic force micrscopy
APS	Ammonium persulphate
BCN	(1α,8α,9β)-Bicyclo[6.1.0]non-4-yne
CuAAC	Copper catalysed azide alkyne cycloaddition
TM-AFM	Tapping mode atomic force microscopy
APP	Amyloid precursor protein
BSE	Bovine spongiform encephalopathy
CDF	Cumulative density function
CJD	Creutzfeld-Jakob disease
CD	Circular dichroism
DSB	Denaturing sizing buffer
Da	Dalton
DCM	Dichloromethane
DOSY	Diffusion ordered spectroscopy
DMF	N,N-Dimethylformamide
DNA	Deoxyribonucleic acid
DMSO	Dimethylsulfoxide
DSB	Denaturing sizing buffer
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
DTT	Dithiothreitol
E.coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EPL	Expressed Protein Ligation
FITC	Fluorescein isothiocyanate
GST	Glutathione S-transferase
IDC	Intrinsically disordered protein
kDa	Kilo daltons
L_p	Persistence length
MAS	Magic angle spinning
MAP	Methionine aminopeptidase
MeCN	Acetonitrile
mQ	Milli-Q
LB	Lysogeny broth
LC-MS	Liquid chromatography-mass spectrometry
NMR	Nuclear magnetic resonance
Ni-NTA	Nickel-Nitrilotriacetic acid
MLB	Melki lysis buffer
MLBG	Melki lysis buffer denaturing buffer
MLBGE	Melki lysis buffer elution
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NHS	N-Hydroxysuccinimide
mAu	Milli absorbance unit
PBS	Phosphate buffered saline
PMSF	Phenylmethanesulphonyl fluoride
ssNMR	Solid-state nuclear magnetic resonance
SEB	α-Synuclein ion exchange entry buffer
SEM	Standard Error Mean
SEIB	α-Synuclein ion exchange elution buffer
SFFB	Sup35 fibril forming buffer

SPAAC	Strain promoted azide alkyne cycloaddition	
SyFFB	α-Synuclein fibril forming buffer	
SIA	Succinimidyl iodoacetate	
SLB	α-Synuclein lysis buffer	
TCEP	Tris(2-carboxyethyl)phosphine	
TFE	Trifluoroethanol	
TFBenol	4,4,4-Trifluorobut-2-enol	
PolyQ	Polyglutamine	
PCR	Polymerase chain reaction	
PTM	Post-translational modification	
PPI	Protein-protein interaction	
Rpm	Revolutions per minute	
rDNA	Recombinant deoxyribonucleic acid technology	
SDD-AGE	Semi-denaturing detergent agarose gel electrophoresis	
SDS	Sodium dodecylsulphate	
SDS-PAGE	Sodium dodecylsulphate polyacrylamide gel electrophoresis	
TAE	Tris base, acetic acid and EDTA buffer	
TEMED	Tetramethylethylenediamine	
Tris-HCl	Tris(hydroxymethyl)aminomethane hydrochloride salt	
Tris base	Tris(hydroxymethyl)aminomethane	
TEM	Transmission electron microscopy	
ThT	Thioflavin T	
TMS	Tetramethylsilane	
UPR	Unfolded protein response	
USA	Ureidosuccinate	

Abstract

Amyloid structures have been identified as a key pathological marker in a range of conditions such as Alzheimer's and Parkinson's disease. A combination of their disease relevance and potential nanomaterial applications has made understanding amyloid formation and the source of its toxicity of great interest. The fragmentation of amyloid fibrils into smaller particles is important for the kinetics of their formation and how they elicit toxicity. This thesis sheds light on how fibrils formed from different proteins fragment, and compares their resistance to mechanical stress, in order to decipher if fibril fragmentation mechanism is fibril dependent or a property shared by amyloid fibrils in general. To accomplish these goals, this thesis is divided into three sections.

The first part of this thesis develops methods for recombinant production and purification of Sup35NMC, Sup35NM and α-synuclein proteins. *In vitro* polymerisation conditions for these proteins and the commercial proteins lysozyme, β-lactoglobulin and insulin B chain has enabled, via AFM imaging methods and MatLab tracing, the calculation of persistence length from far larger populations than has previously been observed.

The second part investigates the resistance to mechanical stress of α-synuclein, β-lactoglobulin and lysozyme fibrils using a fragmentation assay, which enables the determination and comparison of the fibrils' resistance to mechanical stress. This work has demonstrated that polymer rigidity (L_p) is inherently linked to the resistance of the tested amyloid fibrils to mechanical stress.

The third part of the thesis successfully implements a chemical labelling strategy using 'click' chemistry to allow a modification of fibril functionality. This has been applied to fluorescence labelling strategies and ^{19}F-labelling for novel methods of determining the population size of amyloid fibrils by ^{19}F DOSY NMR.

1

General Introduction

1.1   History and Origins of Amyloid

In 1854 the pioneer pathologist Rudolph Virchow showed evidence of starch like accumulations in abnormal brain tissue using iodine/sulphuric acid staining, and termed them amyloid through this association with starch and its Latin/Greek root Amylum/Amylon (Sipe and Cohen, 2000). Subsequent research identified these 'plaques' as proteinaceous in nature and, in fact, the product of the aggregation of misfolded proteins. In the majority of cases protein misfolding and aggregation leads to unstructured, amorphous deposits but the formation of amyloid is far more ordered, developing into long needles or fibrils, with dimensions ranging from 2-10nm in width and a heterogeneous mixture of lengths from nanometres to many micrometres (Fitzpatrick et al., 2013). The core, structured region of amyloid fibres, contains a cross-beta sheet architecture, the presence of which is now the structural definition of amyloid. The confirmation of the cross-beta sheet is usually by biophysical methods like X-ray diffraction (Bonar et al., 1969) and the binding of cross-beta sheet specific dyes such as Thioflavin T (Biancalana and Koide, 2010; Khurana et al., 2005; Krebs et al., 2005) and Congo Red.

Table 1.1   Observed Amyloidogenic Proteins.*

Protein	Type	Disease/Role
Amyloid-β peptide	Disease	Alzheimer's disease
α-Synuclein	Disease	Parkinson's disease
Huntingtin	Disease	Huntington's disease
β2-microglobulin	Disease	Haemodialysis-related amyloidosis
Lysozyme	Disease	Lysozyme amyloidosis
Insulin	Disease	Injection localized amyloidosis
Curli	Functional	Role in colonisation of inert surfaces and aids in Biofilm formation
Sup35	Functional	Acts as an epigenetic switch generating new phenotypes
Chaplins	Functional	Lowers water tension allowing formation of aerial hyphae

*Examples of the peptides/proteins associated with Amyloid formation and their biological significance whether a relationship to a disease or a biological function of the material. Adapted from (Chiti and Dobson, 2006; Knowles et al., 2014).

The phenomenon of amyloid formation is not limited to a single protein/peptide sequence but a trait of many different and unrelated protein sequences. Examples include amyloid associated with disease and functional biological amyloid materials (Table 1.1). Whilst the identified human proteome currently consists of 17294 proteins (Kim et al., 2014) the number of proteins shown to form amyloid fibrils, across all species, is currently only 40 (Chiti and Dobson, 2006). Whilst this list of proteins, to date, is relatively small it has been postulated, due to the presence of smaller amyloiogenic sequences in far more proteins, that if subjected to the appropriate conditions all proteins could form amyloid structures (Dobson, 1999).

The physical properties such as strength that make amyloid a useful material for organisms also makes them potentially useful as a nanomaterial (Maji, 2011). When considering their potential as a nanomaterial, their involvement in numerous debilitating illnesses, and the difficulty in obtaining information about even the earliest points in the assembly, it is of no surprise that there is a vast amount of research carried out in the area of amyloid. Understanding the underlying processes of amyloid formation is therefore imperative to curing amyloid associated disease and developing the next generation of nanomaterials.

1.2   The Mechanism of Fibril Formation

Proteins are the workhorses in all cells and organisms. Many processes in the cell are under the strict control/regulation of numerous pathways, which are built up of numerous highly selective protein-protein interactions (PPI) including the formation of multimeric complexes and catalysing post-translational modifications (PTM) such as phosphorylation and acetylation.

The process of Translation, the ribosome-mediated synthesis of polypeptides from mRNA, is a core biological process in which genetic information is converted into proteins. The folding of the synthesised polypeptide (primary structure) into the secondary and tertiary

structures are what gives the protein the appropriate three dimensional shape to function

(Dobson, 2003) (Fig. 1.1).
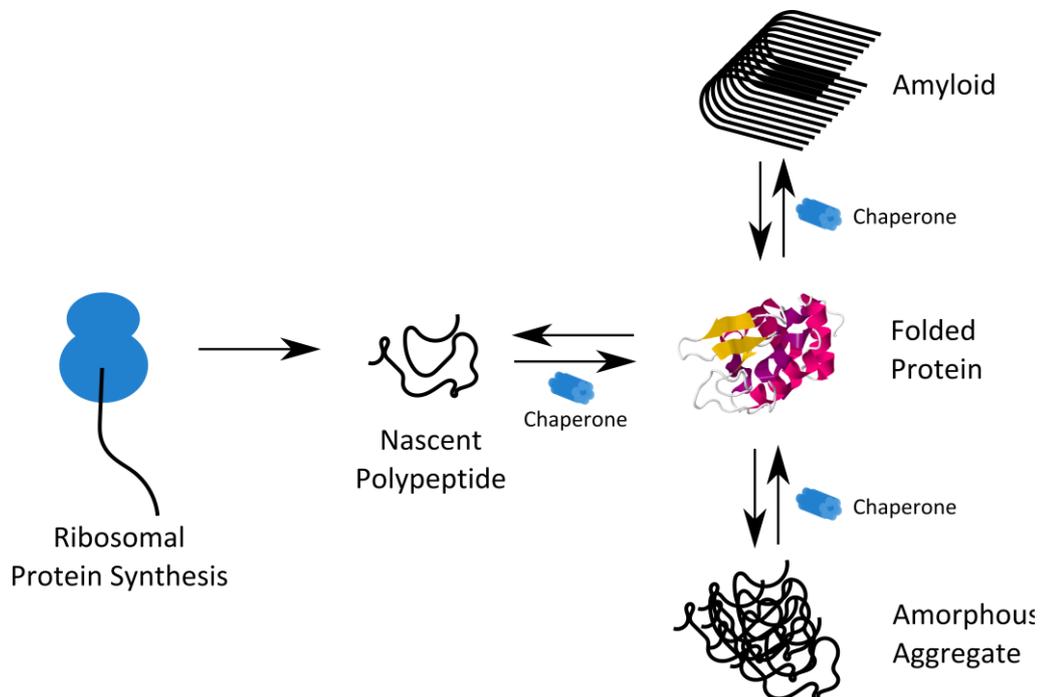
Figure 1.1   Abberrant Protein Folding Can Lead to Aggregation.
The ribosomal synthesis of proteins generates a nascent peptide chain which is folded by molecular chaperones (see blue cylinders) into native folded proteins. The misfolding of proteins can give rise to amorphous aggregates as well as the generation of amyloid fibrils. With Lysozyme as a folded protein example.

The shape and activation state of a protein can also be altered by the modification of

amino acid sidechains in a process called Post-Translational Modification e.g.

phosphorylation and acetylation (Seo and Lee, 2004). A single three dimensional image of

a protein only represents one conformation, usually the lowest energy, but this is not an

accurate representation of a protein molecule in solution. A protein molecule is dynamic in

its shape, constantly changing due to its bonds stretching and bending, and moving

between multiple low energy conformations (Frauenfelder et al., 1991). By considering

each different species and their energy it is possible to create an energy landscape showing

folding intermediates and aggregates (Fig. 1.2).

The 'protein misfolding hypothesis' describes the formation of amyloid whereby the aberrant folding of a normally soluble protein to give an unfolded or partially folded species interacts with another similarly misshapen protein, the addition of further misfolded monomers to this dimer and subsequent species forms the fibril (Soto, 2001), see the right-hand side of figure 1.2.
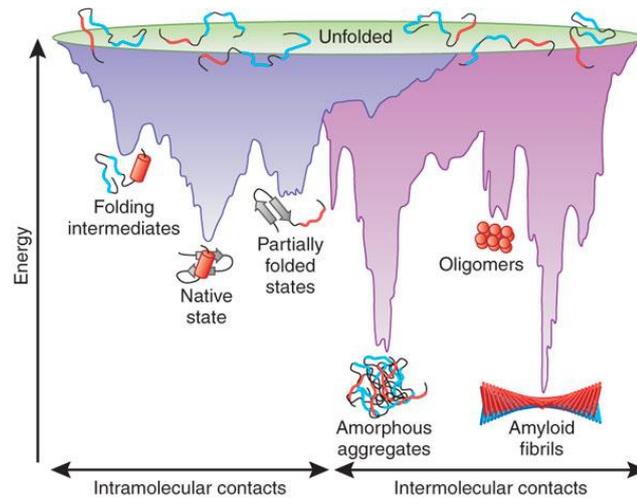
Figure 1.2   Energy Landscape of Amyloid Formation.
An energy level diagram showing the energy of species per monomer involved in the proper folding of a protein monomer (left) leading to the aggregation of monomeric proteins in to fibrillar aggregates (right). Taken from (Hartl and Hayer-Hartl, 2009)

When the appropriate conditions are encountered the misfolding of proteins can occur spontaneously but the propensity of this process can be modified by two main factors: primary sequence alteration and post-translational modifications.

1.2.1   Primary Sequence Alterations

A single change to the nucleotide sequence which codes for a protein can result in a change in the amino acid sequence of the protein. These mutations could have a negligible effect on the protein but conversely could also seriously affect the function of the protein by removing catalytic residues or changing the overall shape of the protein. A single change to the amino acid sequence has the potential to alter the energy barriers for the interconversion between protein conformations and therefore inhibiting or promoting amyloid formation. This is seen for example with mutations to α-synuclein, aggregates of

which have been shown to form the major component of Lewy bodies, the pathogenic

marker associated with the onset of Parkinson's disease (Dauer and Przedborski, 2003).
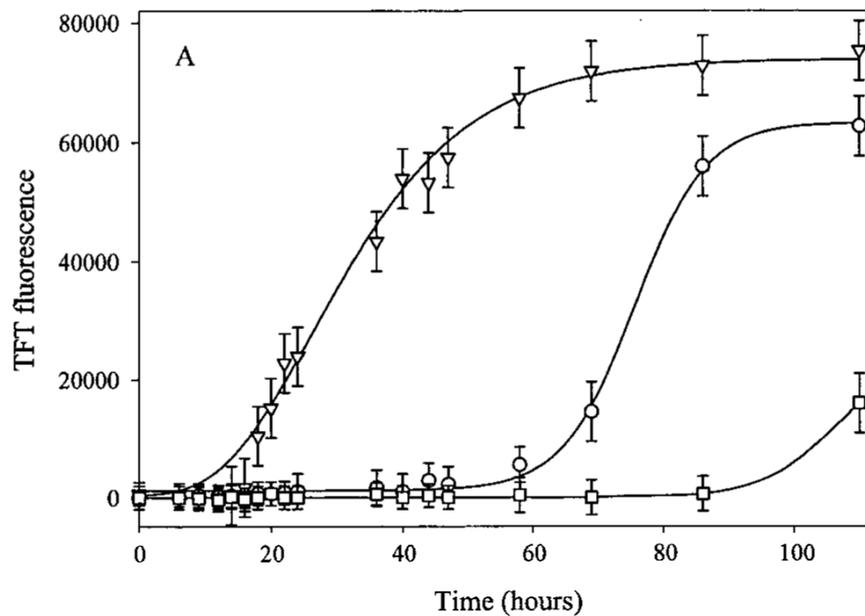
Figure 1.3   Protein Sequence Affects Fibril Formation/Kinetics.
Amyloid growth curve monitored by Thioflavin T fluorescence showing the kinetics of fibril formation for the formation of amyloid fibrils of WT α-Synuclein (circles), A30P α-Synuclein (squares) and A53T α-Synuclein (inverted triangles). Adapted from Li et al. (2001) (Li et al., 2001)

There are two main types of Parkinson's disease: sporadic, which is more common in older

patients and involves the aggregation of WT α-synuclein, and inherited, where symptoms

typically occur at a much younger age and involve mutations in the α-synuclein protein

such as A30P and A53T. The α-synuclein mutant A53T has been shown *in vitro* to have

much faster fibril formation kinetics than the wild-type monomer, which potentially

explains the early onset of disease observed with this mutation (Fig. 1.3, triangles).

Although the A30P mutant is also associated with the early onset inherited form of the

disease, in vitro assessment of the fibril formation kinetics shows that it is actually slower

than the wild-type monomer (Fig. 1.3, A30P - squares, WT - circles). The earlier onset of

disease associated with this slower polymerising mutant can be rationalised as it is

proposed to stabilise toxic prefibrillar oligomers (Conway et al., 1998; Li et al., 2001).
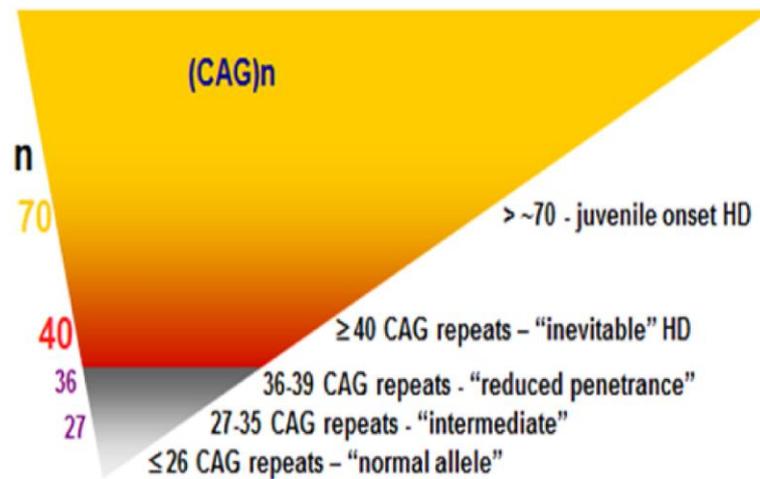
Figure 1.4   Genetic Expansion Can Increase Amyloidogenicity.
Graph showing a comparison of the number of repeating units in non-Huntington's disease patients. Progression from the bottom of the graph, healthy individuals to the highest part of the graph with juvenile onset Huntington's disease is followed by an increase in the number CAG repeats. The number of repeats required to show Huntington's symptoms is 37. Adapted from http://medicaliaorg.ning.com/group/neurosciences/forum/topics/parkinson-s-disease-and-other-movement-disorders and (Fahn and Cohen, 2007).

The propensity to form amyloid aggregates can also be influenced by the expansion of

regions of repeating amino acids in a protein sequence via the incorporation of

trinucleotide repeats into the gene encoding the protein of interest The expansion of

trinucleotide repeat regions has been associated with several disease states, trinucleotide

repeat disorders, although the disease phenotype is not necessarily due to the translation

of the expanded gene into a protein, e.g. fragile X syndrome (Bagni and Tassone, 2012). In

Huntington's disease the repeat region falls in a protein coding region, where a repeating

CAG is translated into a polyglutamine sequence (polyQ). The expansion of the Huntingtin

protein with a polyQ sequence is thought to be a causative process in progression of the

neurological disorder Huntington's disease (Scherzinger et al., 1997). The disease

phenotype associated with Huntington's is closely linked to the size of the repeat region,

with thirty seven repeats being the lower limit to disease onset and the aggregation of

monomers into amyloid fibrils (Duyao et al., 1993) (Fig. 1.4).

1.2.2   Post Translational Modification

The process of translation and folding gives native protein conformations but their ability

to undertake cellular processes is modulated by post-translational modifications. Common

post-translational modifications utilised in the cell are phosphoryation, acetylation and

glycosylation. It is the addition or removal of these groups which can activate or deactivate

proteins, allowing the control of cellular processes.
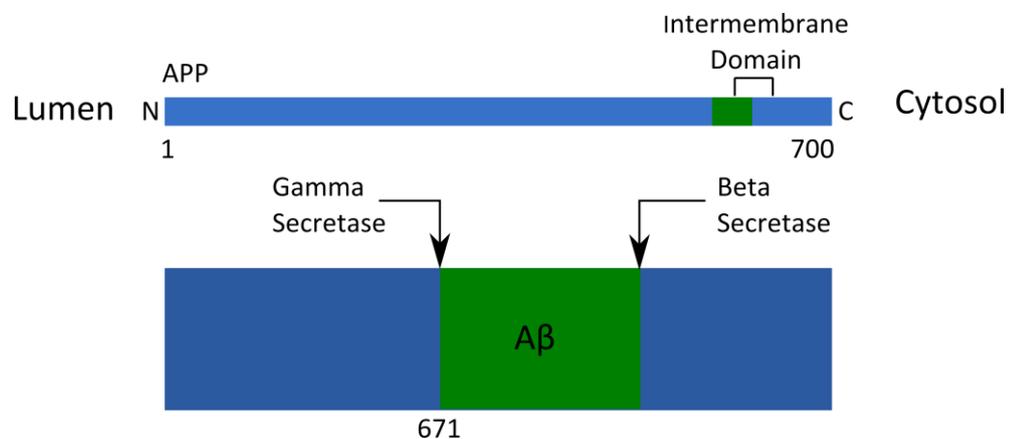
Figure 1.5   APP Processing Generates Amyloidogenic Peptides.
Top – A diagram showing full length amyloid precursor protein (APP) and the intermembrane
domain where secretase activity is observed. Bottom – A zoom in on the intermembrane
domain which contains the sequence for the peptide Aβ which is liberated by action o f γ-
Secretase on residue 671 of the protein sequence and the non selective proteolysis by beta-
secretase which gives a range of peptides from 40-42 residues long.

The addition of these groups to the protein also has the potential to alter the energy

barriers to conformational change and therefore, potentially, amyloid formation as

observed with the phosphorylation of S87 of α-synuclein (Paleologou et al., 2010).

Post-translational modifications do not only involve the addition of groups to proteins, but

also the hydrolysis of the peptide chain into smaller fragments by hydrolases. The specific,

enzyme mediated, cleavage of a peptide has the potential to liberate a smaller fragment

with higher propensity to adopt an amyloidogenic conformation.   One of the key

biomarkers of Alzheimer's disease is the presence of amyloid deposits derived from peptides originating from the protein Amyloid Precursor Protein (APP). The processing of APP by β and γ secretases gives rise to peptides ranging from 40-42 residues in length. It is these small peptides which aggregate to form amyloid fibrils and are associated with Alzheimer's disease (Aguzzi and O'Connor, 2010) (Fig. 1.5).

1.3   Structure of Amyloid Fibrils

Throughout history, the need to cure or relieve the symptoms of human diseases has driven many studies of biological systems. The modern era of major advances in biology/molecular biology has been catalysed by the accurate characterisation of the structure and function of biomolecules. The development of x-ray diffraction, where a crystalline material is placed in an x-ray beam and the scatter pattern of the beam is collected and converted into an atomic resolution 3D structure, has been key to this advance. This is exemplified by work of Dorothy Hodgkin, from 1945 onwards, examining the structure of small biomolecules such as cholesterol and insulin to the work of Rosalind Franklin, from 1950 onwards, on the macromolecular structure of DNA (Jaskolski et al., 2014).

The field of x-ray crystallography relies on the deconvolution of the diffraction pattern obtained when a large crystal of the molecule is placed in an X-ray beam (Smyth and Martin, 2000). Amyloid fibrils are large ordered but heterogeneous polymers, which limits their analysis by high resolution x-ray diffraction methods. By placing non-crystalline materials into the x-ray beam and analysing the scattering spectrum obtained, it is still possible to gain structural insight into the material analysed. In fact, cross-beta structures give a specific x-ray fibre diffraction pattern. This cross-β pattern gives rise to a 4.7-8Å diffraction on the meridian due to the distance between of the strands of the cross-beta sheet and 10-12Å diffraction on the equator which is due to the distance between the

sheets. Due to this consistent x-ray diffraction it is an important method for confirming the

presence of the cross-beta sheet core of amyloid fibrils (Morris and Serpell, 2012) (Fig. 1.6).
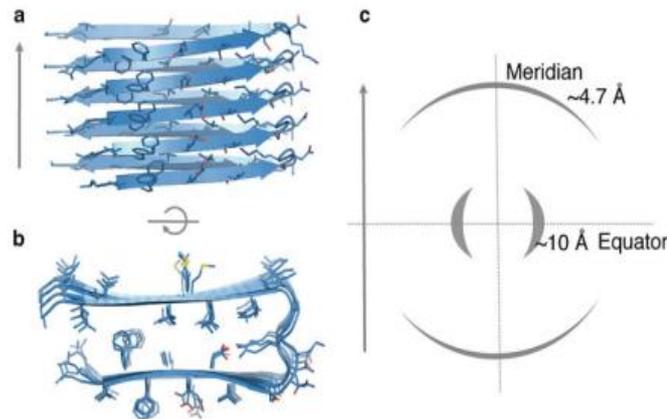
Figure 1.6   Cross β-sheet Architecture of Amyloid Fibrils.
A model of the cross beta-sheet architecture of Aβ(1-42) peptides from above (a) and through
the fibril (b). The β-strands are packed perpendicular to the fibre axis separated by 4.7-8Å.
The β-sheets run parallel to the fibre axis with a distance of 10Å separating them. These
values can be accessed from the Meridian and Equator of the X-ray diffraction pattern (c).
Adapted from (Morris and Serpell, 2012)

Nuclear Magnetic Resonance (NMR) spectroscopy has been a key technique for the

identification and investigation of small molecules for many years. Its versatility and ability

to differentiate between atoms due to their different electronic environments as well as

their connectivity and spatial orientation means it is highly suited to many biological

applications including the structural characterisation of biomolecules (Marion, 2013). NMR

is typically a solution based technique that can detect molecules upto 30kDa in size. The

slower tumbling of larger materials means that in the time frame of the NMR experiment

the visualisation of average peaks for each ^{1}H species is impossible. NMR requires the use

of NMR active isotopes, having a spin quantum number of ±½, of which ^{1}H, ^{13}C, ^{15}N and ^{31}P

are examples of isotopes commonly found in biomolecules. When these nuclei are aligned

with a magnetic field their chemical environment is probed by irradiation with radio

frequency electromagnetic radiation causing the nuclei spin state to change. The different

wavelengths of radiowaves absorbed gives information on the electronic environment of the nuclei. As amyloid fibrils are very large polymeric materials the use of solution NMR to understand their structure and dynamics is limited. Recent advances in the application of magic angle spinning (MAS) solid state NMR (ssNMR) techniques has however opened up (Bai et al., 2010). In these instruments the sample, in the form of a dry powder containing an internal reference, 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), is placed into a rotor and spun allowing for average NMR spectra of the solid to be captured (Schuetz et al., 2010; Shi and Ladizhansky, 2012). New ssNMR approaches allow for structural information to be gleamed from solid materials such as amyloid (Heise, 2008) (Fig. 1.7).
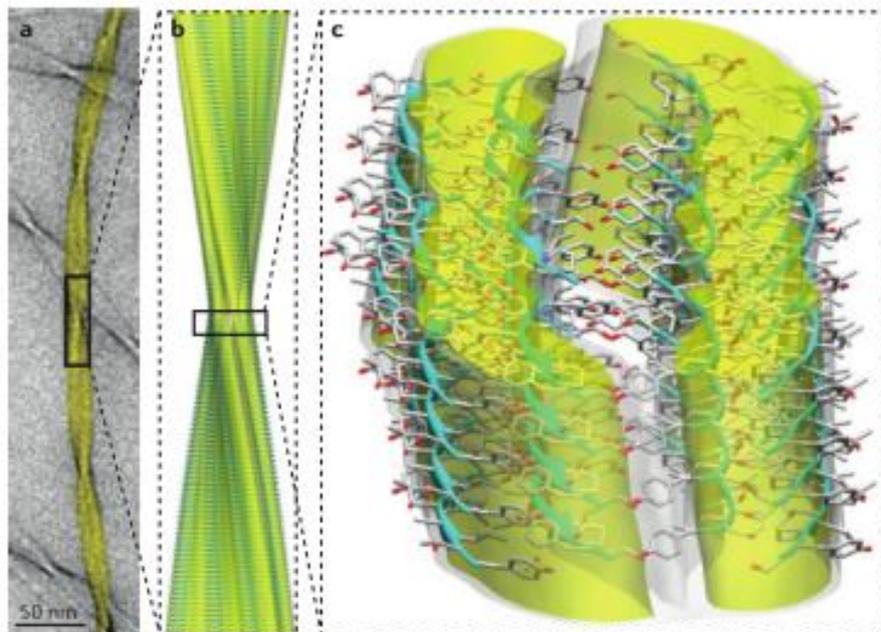
Figure 1.7   Amyloid Structural Modelling by Combined ssNMR and cryoEM.
This high resolution model was generated by combining data acquired through cryo-electron microscopy (a) and solid-state NMR (b) the fibril surfaces assigned by solid state NMR (c) of an 11 amino acid peptide from the protein Transthyretin. Adapted from (Fitzpatrick et al., 2013)
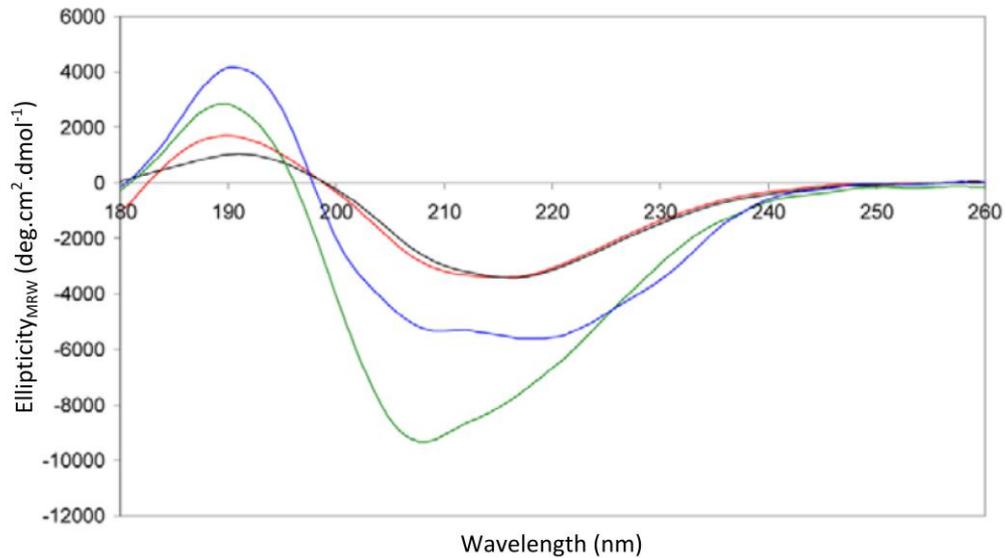
Figure 1.8   Circular Dichroism Shows Conformational Changes on Conversion from Monomer to Polymer.

Circular Dichroism spectra showing the conversion of Myostatin Precursor Protein on conversion of monomer (green) to functional dimer (blue) to prefibrillar aggregates formed overnight (red) to a mixture of fibrils and prefibrillar species after a week of incubation (black). Adapted from (Starck and Sutherland-Smith, 2010)

Proteins are polypeptide chains that consist of multiple amide bonds, aromatic species and disulphide bonds, all of which absorb in the UV region, 180nm-240nm, of the electromagnetic spectrum. Circular Dichroism (CD) uses left and right circularly polarised light to probe this spectral region, giving information on different types of secondary structure in a sample. CD is therefore particularly useful for proteins (Greenfield, 2006) and is effective for monitoring amyloid formation, due to its ability to follow the structural changes associated with cross beta sheet formation when monomers convert to fibrils (Fig. 1.8).

Infrared spectroscopy is a vibrational spectroscopy which uses infrared light to excite bonds in molecules, causing them to stretch and bend. Functional groups in proteins give rise to different absorbances and can be used to monitor changes in conformational states. For example, the stretch of the carbon oxygen double bond in an amide, typically 1630-1680cm^{-1}, is commonly used (Barth, 2007).

1.4   Imaging of Amyloid Fibrils

The elongation of fibrils occurs by addition of monomers to the ends of fibrils (Pinotsi et al., 2014; Scheibel et al., 2001) but lateral growth of a fibril is achieved by bundling of these individual strands into thicker fibres. These fibrils start to display a further level of structural organisation as they entwine themselves together to give twisted structures. For each fibril formed, it is possible for them to laterally associate in different ways giving rise to different morphologies of fibril, which is of potential interest when considering the biological interactions of amyloid fibrils.

Amyloid fibrils are typically 2-10nm in width and have a heterogeneous mixture of lengths typically ranging from nanometres to many micrometres. Their size has traditionally limited their visualisation by optical microscopy but recent advances in super resolution microscopy is making this more feasible (Pinotsi et al., 2014). It is, however, possible to visualise these structures using microscopy techniques such as transmission electron microscopy (TEM) and atomic force microscopy (AFM). These techniques allow for discrimination of particles in terms of their lateral assembly and provide information about the different fibril morphology populations contained in a sample. Imaging using TEM requires negative staining of the fibrils with a heavy metal ion, typically uranium from uranyl acetate. The stain coats the outside of the fibril and stops the electron beam from passing through the sample which gives rise to the observed contrast, enabling the measurement of length and investigation of morphological structure of the fibres (Fig. 1.9, Left). TM-AFM is a scanning probe microscopy technique that differs from other microscopy techniques in that is does not use the absorption or emission of energy from the electromagnetic spectrum to visualise materials, instead images are constructed through the physical interaction of a probe with a sample. The mechanism by which this microscope works allows not just imaging of a sample but investigation of the nanomechanical properties associated with the material such as the Young's modulus as

well as the forces which hold material together (Adamcik and Mezzenga, 2012; Sweers et

al., 2011).
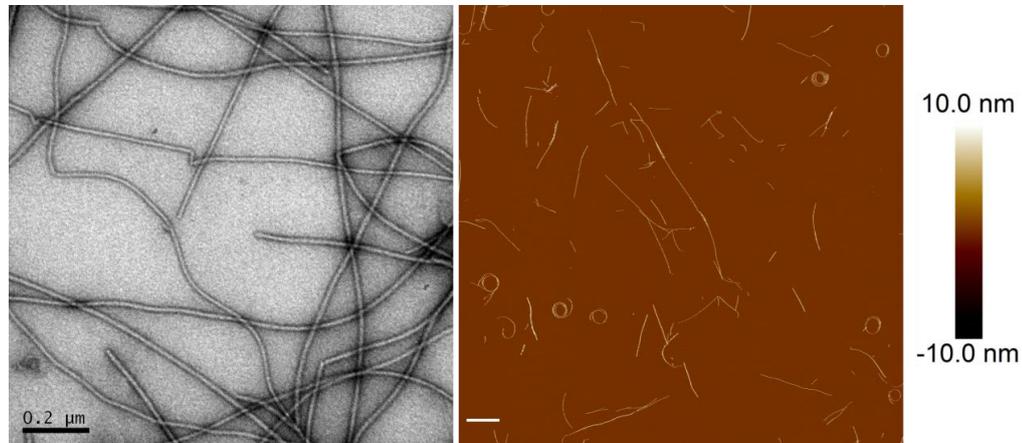
Figure 1.9   Fibrillar Structure of Amyloid Imaged using Microscopy.
Amyloid fibrils derived from α–synuclein.   Fibrils imaged using Transmission Electron
Microscopy with Uranyl Acetate staining (left, 200nm scale bar) and Atomic Force Microscopy
(right) with 20x20μm imaged at a resolution of 2048 x 2048 pixels (1000nm scale bar) and a
height scale bar on the far right. The AFM height colour scale is set between -10 and 10nm.
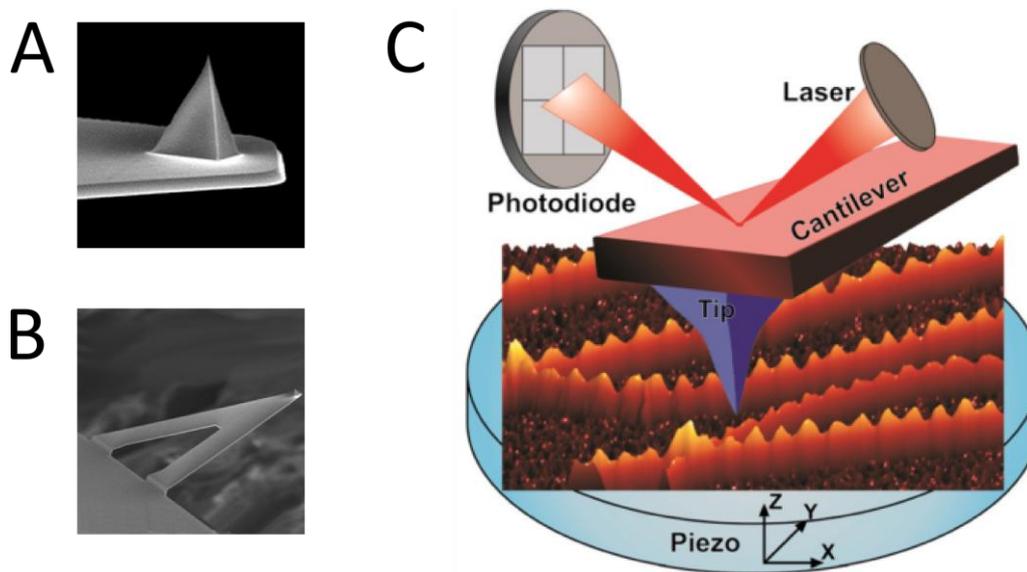
Figure 1.10   The workings of Atomic Force Microscopy.
A – A scanning electron micrograph of the tip of a Bruker scan asyst probe with a height, 2.5-
8μm and nominal tip radius – 2nm. http://www.brukerafmprobes.com/p-3726-scanasyst-
air.aspx.   B – An SEM of a bruker scanasyst cantilever made of silicon nitride
http://www.brukerafmprobes.com/p-3726-scanasyst-air.aspx.   C – The determination of
amyloid particles height and length using Atomic Force Microscopy (Adamcik and Mezzenga,
2012).

When a probe (Fig. 1.10, A) attached to a cantilever (Fig. 1.10, B) is oscillated by a piezo electric current and then lowered onto a surface a tapping motion is generated. For imaging purposes, the probe is moved over the surface of interest and any change in height is identified by the deflection of a laser aligned onto a sensor via reflection off of the cantilever (Fig. 1.10, C). The probe generates a height measurement per pixel, which when placed together produces a topographic map (Fig. 1.09, Right) that allows for the assessment of the lengths, heights and morphology of individual particles.

1.5   Intermolecular Interactions

The conversion of monomers to ordered amyloid aggregates via cross-beta sheet formation results in large polymers which are held together by intermolecular forces such as Van der Waals interactions and hydrogen bonding. One of the major properties associated with amyloid is their resistance to disaggregation by sodium dodecyl sulfonate (SDS), conditions which would normally interrupt most non-covalent intermolecular interactions. This effect is potentially due to the cumulative effect of numerous intermolecular interactions and the inability of SDS to penetrate the cross-beta sheet and interrupt the interaction.

There is also potential for covalent bonds to form between two strands. For example copper has been shown to catalyse formation of dityrosine crosslinks between Aβ monomers suggesting that the stability could be aided by covalent interactions (Al-Hilaly et al., 2013; Atwood et al., 2004; Kok et al., 2013) (Fig. 1.11).
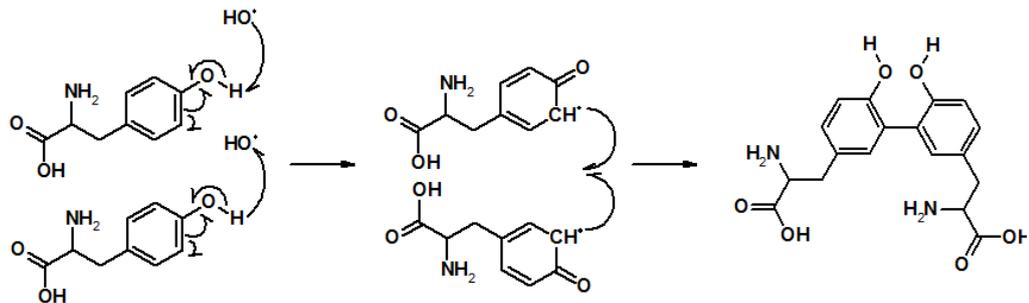
Figure 1.11   Reaction Mechanism for Radical Mediated Formation of Dityrosine Crosslinks.

The formation of a dityrosine moiety (right) from tyrosine monomers (left) follows a radical mechanism in which two carbon centred radicals (middle) combine in a termination reaction.

The SDS resistance of amyloid fibrils has led to the development of electrophoresis based methods for characterising aggregate size. Semi-Denaturing Detergent Agarose Gel Electrophoresis (SDD-AGE) has been used to show and compare the different sizes of amyloid aggregates from *in vitro* and *in vivo* samples (Fig. 1.12).
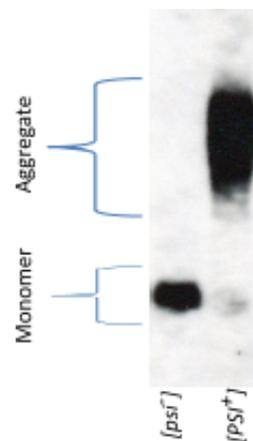
Figure 1.12   An example of an Immunoblotted SDD AGE.

An SDD AGE gel showing the different sizes of Sup35 protein in [*PSI*$^+$] and [*psi*$^-$] cells, larger aggregates do not move travel through the gel as far as the monomeric protein and form a diffuse band due to the range of sizes contained with a population of amyloid species. Image supplied by Ricardo Marchante.

1.6   Dye Binding Properties

In 1854, Rudolph Virchow used non specific staining to show the presence of deposits in tissue samples using iodine and sulphuric acid. Thus one of the first and defining characteristic properties of amyloid is the ability of the cross-beta sheet structure to bind certain dye molecules.
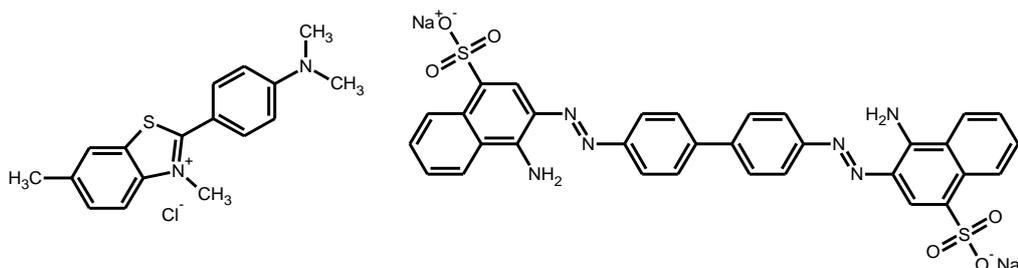
Figure 1.13   Chemical Structures of Amyloid Binding Dyes.
The chemical structures of the dyes commonly used in the analysis and identification of amyloid fibrils. Left) Thioflavin T (ThT) Right) Congo Red

The main dye molecules used in this context are the highly π-conjugated dyes Thioflavin T (ThT) and Congo Red (Fig. 1.13), which both give a change in the absorbance/emission profile when in the presence of amyloid fibrils. Although the binding mode of these dyes to the amyloid is not resolved it is assumed that the dyes bind in the grooves of the cross beta sheet structure locking of the molecules into one conformation giving rise to the change in their absorption/emission characteristics (Biancalana and Koide, 2010; Khurana et al., 2005; Krebs et al., 2005).

These amyloid binding dyes have been useful in the identification of amyloid plaques in histological samples. The change in wavelength of the spectroscopic profile of these dyes, in particular ThT, which exhibits a stokes shift from 450nm to 480nm (Naiki et al., 1989), on binding to amyloid fibrils also offers great advantages for measuring the kinetics of fibril formation *in vitro*, where an increase in fluorescence is observed when fibrils are formed.

1.7   Resistance to Proteolysis

The control of protein levels in the cell is regulated by their synthesis and degradation into

their constitutive amino acids by proteases which recognise specific regions of proteins and

hydrolyse appropriate amide bonds. When a ligand is bound to a protein (holo form), it has

been shown to have increased stability to denaturation and proteolysis than its apo-form

(Lomenick et al., 2011; Pace and McGrath, 1980).
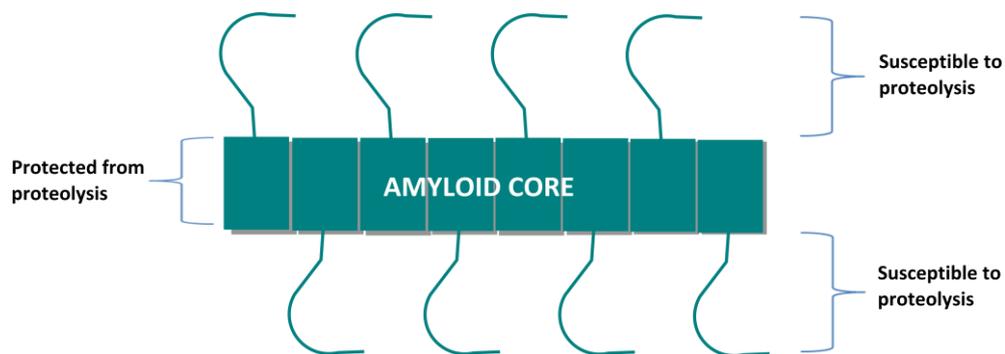
Figure 1.14   Amyloid Core is Protected From Proteolysis.
The amyloid core (repeating rectangles) is protected from proteolysis but the exposed
polypeptide chains are still susceptible to proteolysis.

Similarly, when a protein adopts the amyloid state the cross-beta sheet of the structure is

protected from proteolysis but regions that are exposed can still be accessed by proteases

and therefore digested (Fig. 1.14).   Thus limited proteolysis, followed by mass

spectrometric analysis is able to determine which regions are protected and therefore part

of the amyloid core (Frare et al., 2006; Kheterpal et al., 2001).

1.8   The Kinetics of Fibril Formation

The process of amyloid fibril formation follows a nucleated growth mechanism and is

described simplistically by the cartoon in Fig. 1.15.   The process initiates with a

conformational change in the protein that exposes surfaces that are normally buried, which

then interact with other similarly misfolded protein molecules to give a dimer in a

reversible process (Knowles et al., 2014). The requirement for conformational change prior

to amyloid formation is shown for β_2-microglobulin, where a shift in pH from pH7 to below

pH5 initiaties the conversion of monomeric β$_2$-microglobulin to its amyloid form, as determined by an increase in Thioflavin T fluorescence and the generation of long needle like structures. The cause of this pH induced fibril formation was shown, by ANS binding, NMR and CD, to be conformational changes resulting in partially unfolded β$_2$-microglobulin which is then susceptible to polymerisation (McParland et al., 2000). Further reversible and unfavourable interactions of the transient dimeric species with misfolded monomers give rise to small oligomeric species. The reversibility of the process continues until a nucleus is formed where subsequent favourable addition of monomers to the ends of the species results in fibril elongation and generation of the fibrillar species observed *in vivo* and *in vitro* (Fig. 1.15).
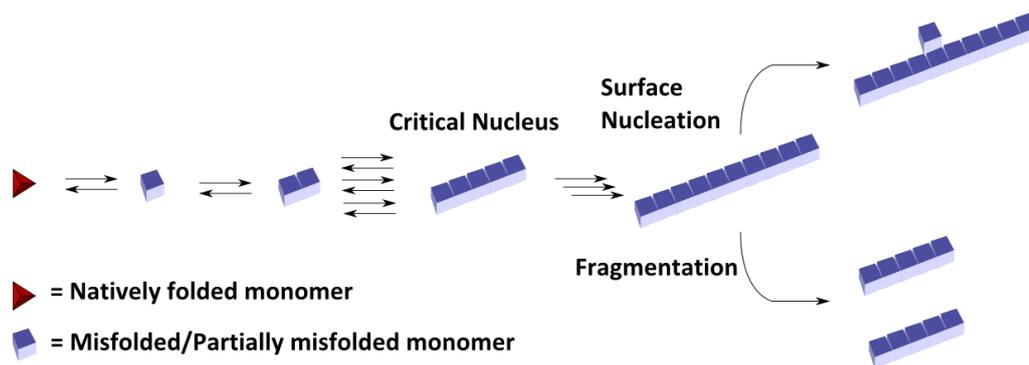
Figure 1.15   Reaction Scheme Showing Amyloid Fibre Formation.
This depiction represents the conversion of properly folded monomers (red pyramid) to misfolded monomer (blue cube) and then their subsequent oligomerisation to a critical nucleus and elongation to fibrils. This process is then followed by secondary effects such as fragmentation and surface nucleation.

The key step of a nucleated growth mechanism is the generation of a nucleus onto which subsequent monomer addition follows to give fibrillar species. The rate of nucleus formation can be affected by many different factors such as monomer concentration and temperature.
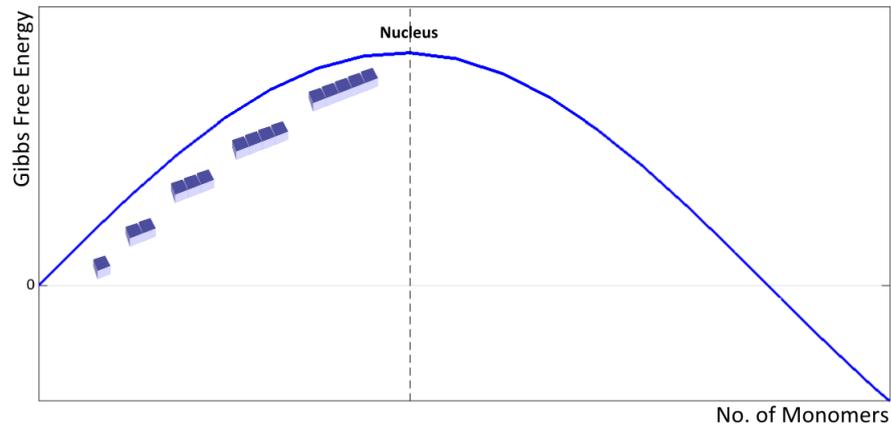
Figure 1.16   Energetic Considerations of Critical Nucleus Formation.
Energy level diagram showing the increase in Gibbs free energy associated with the reversible
generation of the critical nucleus followed by the decrease in Gibbs free energy as the process
essentially becomes irreversible.

The change in Gibb's free energy (ΔG^o) of a system takes into account enthalpic and

entropic considerations giving a prediction of how likely a process is to occur. A positive

ΔG^o is equivalent to an unfavourable process where as a negative ΔG^o is evidence of a

favourable process.
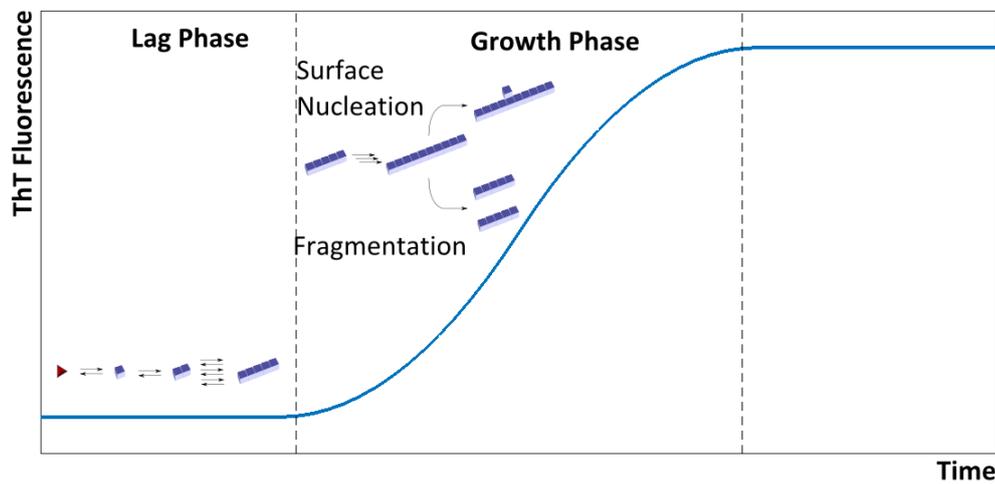
Figure 1.17   The Kinetics of Fibril Formation
This is a depiction of the standard sigmoidal Thioflavin T fluorescence curve associated with
the formation of amyloid fibrils. A lag phase dominated by the formation of the critical
nucleus is followed by exponential growth where fibrils growth is mainly observed.

The process of nucleation in terms of Gibbs free energy change shows the initial reversible processes of dimerisation and oligomerisation gives rise to an increase in ΔG^o to a point whereby ΔG^o then starts to decrease again, towards lower energy species. The higher energy species formed in this process are transient and short lived when compared to the lower energy species and therefore of lower concentration. It is this free energy maximum that describes the nucleus of the fibril formation under the conditions used (Fig. 1.16) (Knowles et al., 2014; Xue et al., 2008).

The characterisation of fibril formation kinetics has been aided by the use of the profluorescent dye Thioflavin T. Binding of ThT to the cross beta sheet architecture of the amyloid fibrils gives a change in the absorption/emission profile of the material, thereby allowing fibril formation processes to be followed (Biancalana and Koide, 2010; Khurana et al., 2005; Krebs et al., 2005). By monitoring fluorescence periodically over the course of fibril formation, in the presence of ThT, a sigmoidal curve is observed as fibrils are formed.

It is possible to assign the dominant reaction processes shown in Fig. 1.15 to each region of the sigmoidal curve (Fig. 1.17). At the start of fibril formation the initial lag phase is associated mainly with monomer conformational interconversion and interaction up until the generation of the critical nuclei (Xue et al., 2008). The subsequent exponential growth phase is dominated by fibril growth and an increase in the proportion of beta-sheet in the sample. The lag phase is a stochastic process which is dependent on the probability of a first critical nucleus forming. A side effect of the stochasticity of nucleation is that two fibril formations carried out using identical conditions may not have equal lag phase lengths. The exponential growth of the system following the lag phase is not only due to the elongation of nuclei into fibrils by monomer addition, but also a series of secondary processes which may occur once fibril elongation has started (Lorenzen et al., 2012). It must be noted that nucleation still occurs but is no longer the primary mechanism of

increasing the proportion of cross beta sheet in the sample. Currently there are two main secondary effects that are observed to occur: fibril fragmentation and aggregate surface nucleation. Fibril fragmentation is the process where a fibril breaks into two smaller particles, and affects fibril formation by increasing the number of ends available for fibril extension (Fig. 1.15) (Xue and Radford, 2013). Surface nucleation is the process where a fibril interacts laterally with a monomer, causing a conformational change in the monomer and essentially catalysing the generation of new nuclei (Fig. 1.15) (Foderà et al., 2008; Kashchiev et al., 2013). The exponential phase of fibril growth occurs until the monomer is consumed to a level where no further increase in amyloid content of a sample and plateauing of the growth.

1.9   Amyloid Toxicity

Since its discovery in 1854 amyloid has been associated with the progression of numerous diseases, most notably neurodegenerative, and age related diseases such as Alzheimer's, Parkinson's and Huntington's disease. The common structural component, the cross-beta sheet is independent of amino acid sequence and leads to an assumption that the toxicity associated with diseases occurs via a common mechanism. Consequently by developing an inhibitor of fibril formation a panacea for amyloid associated disease could be realised. Despite this assumption and a vast amount of research on the biological effect of amyloid and their disease association there are still questions relating to how and why toxicity is observed. Identifying the toxic species that are associated with amyloid related diseases is currently one of the most important questions with respect to desigining therapeutic interventions of these diseases. The current contenders for the toxic species are fibrils, prefibrillar oligomers and fibril fragments.

There are several current hypotheses on the mechanism of amyloid toxicity including cell membrane pore formation, through interaction with fibrillar materials, or loss of function

of the aggregating species. The debate surrounding the identity of the toxic amyloid species is highlighted by the disease synonymous with neurodegeneration due to amyloid formation, Alzheimer's disease. The generation of fibrils from amyloid beta had been thought to be the causative reason for cytotoxicity observed in Alzheimer's disease. For some time the disagreement between observed fibril load and toxicity (Ferreira et al., 2007) has led to the postulation that large fibrils are relatively innocuous. In fact, the generation of small "oligomeric" species, on pathway to fibrils, and their internalisation or interaction with cell surface receptors are thought to cause cell death and the onset of neurodegeneration (Kayed and Lasagna-Reeves, 2013; Kayed et al., 2003).
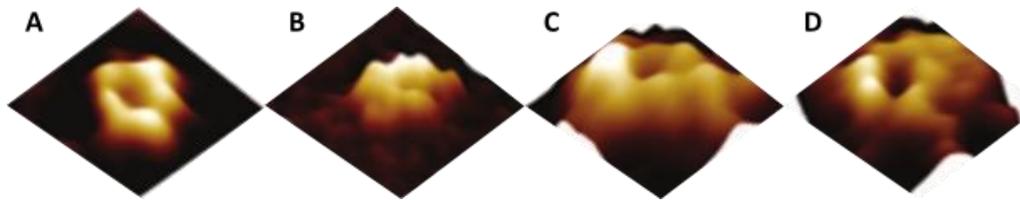
Figure 1.18   Oligomer Particles on the Pathway to Amyloid can Cause Membrane Pore Formation.
Atomic Force Microscope images of different amyloidogenic proteins forming pores on lipid bilayers of DOPC. Aggregating proteins: A – Aβ(1-40), B – α-synuclein, C – amylin and D – serum amyloid A (SAA). Adapted from (Quist et al., 2005).

Small prefibrillar oligomers have also shown a propensity to form circular pore like structures on cell membranes. These structures, like ion channels, may allow the diffusion of small molecules and ions into and out of cells or cell compartments and thus negatively affecting cellular homeostasis (Fig. 1.18) (Quist et al., 2005). However this still does not explain why the onset of toxicity is delayed as these prefibrillar structures would have been expected to exist all through life.

Anti-Alzheimer's drug discovery has been attempted through inhibition of secretases, inhibition of fibril formation, and removal of plaques/plaque intermediates with antibodies (Aguzzi and O'Connor, 2010). These research have yet to lead to a viable treatment, adding

weight to the assumption that fibrils are not the sole cause of toxicity. Thus the identity and mechanisms of amyloid-associated toxicity in disease is still an open question.

Another potential mechanism of cytotoxicity revolves around the sequestering of necessary proteins to amyloid aggregates, which means that they are subsequently unable to take part in their usual cellular processes, leading to a disease phenotype. The cellular response to misfolded/aggregated proteins is to resolubilise or degrade them. As well as their role in the folding of proteins into their native state, chaperones also have disaggregation function (Duennwald et al., 2012; Shorter and Lindquist, 2004), which allows proteins that have been sequestered into amyloid fibrils to be extracted and either refolded correctly or degraded (Dobson, 2003; Hartl and Hayer-Hartl, 2002). This process forms part of the Unfolded Protein Response pathway (UPR). The major component of the UPR is to degrade misfolded or damaged proteins into their constituent amino acids, releasing them into the cytosol were they can be reused, thereby maintaining homeostasis. The UPR works via the attachment of ubiquitin to proteins, which targets them for degradation by the proteosome (Lecker et al., 2006). If the UPR cannot rectify the protein misfolding issues of the cell then its final response is to signal for apoptosis of the cell (Walter and Ron, 2011). The ability of a cell or organism to cope with protein misfolding is inextricably linked to age which also fits with the pathology of amyloid disease progression and its associated 'incubation time'.

Although the formation of amyloid deposits is a specific process involving the polymerisation of monomers into fibrils, it is possible for other proteins to be incorporated into the aggregates. Once these interactors have been removed from the cells' accessible proteome, the cells mechanism for dealing with protein misfolding may be overwhelmed (Olzscha et al., 2011).

1.10   Prions

In the 1800s, the work of Gregor Mendel, the father of modern genetics, described the guidelines for the inheritance of traits from one organism to its offspring. The theory of Mendelian inheritance is governed by 3 laws, Mendel's laws, and explains the transfer of dominant and recessive characteristics from parents to offspring. Since then, DNA has been shown to be the molecule which carries the information of inheritance.

Many infectious agents, such as viruses, work by hijacking a cells own transcriptional/translational machinery to propagate and therefore require the transmission of DNA from one cell to another. In 1982, Stanley Prusiner suggested, controversially, that the source of toxicity/neurodegeneration observed in the spongiform encephalopathy Scrapie was due to a protein only mechanism. This mechanism requires the transmission of a proteinaceous particle from one cell to another (Prusiner, 1982). The term prion was coined by Prusiner and is a portmanteau of the words **Pro**teinaceous and **In**fectious. Subsequent work by Prusiner has since shown that the causative prion proteinaceous material consists of amyloid aggregates (Prusiner et al., 1983).

After the discovery that a protein only mechanism is responsible for the transmission of scrapie, it was soon confirmed that other spongiform encephalopathies, such as bovine spongiform encephalopathy (BSE), Creutzfeld-Jackson Disease (CJD) and Kuru, followed this process.

1.10.1   Yeast Prions

In 1994 Reed Wickner noted that the yeast protein Ure2p exhibited the characteristics of a prion. In fact its conversion to a prion state, [*URE3*], offered an advantage to the organism when placed in an environment with poor nitrogen sources (Wickner, 1994). Ure2p is a protein of the yeast *Saccharomyces cerevisiae* that regulates the catabolism of nitrogen sources via binding of the transcription factor Gln3p GATA. In the presence of good nitrogen sources such as ammonia and glutamine, Ure2p prevents the utilisation of poor

nitrogen sources such as ureidosuccinate (USA) by preventing their uptake into the cell by

repressing the transcription of permeases (Lian et al., 2006; Xu et al., 1995). Ure2p binds to

Gln3p GATA and stops its movement to the nucleus therefore repressing transcription. The

conversion from [ure3] to [*URE3*] was shown to follow the conversion of Ure2p from

soluble monomer to amyloid (Wickner et al., 2004). When Ure2p is monomeric it blocks

Glnp3 GATA from entering the nucleus but in its amyloid state the interaction is inhibited

and the Gln3p GATA is free to move to the nucleus and allows permease transcription. The

net result of this epigenetic regulation is the ability for the organism to utilise poor nitrogen

sources such as ureidosuccinate (USA) under diminished nutrient conditions.
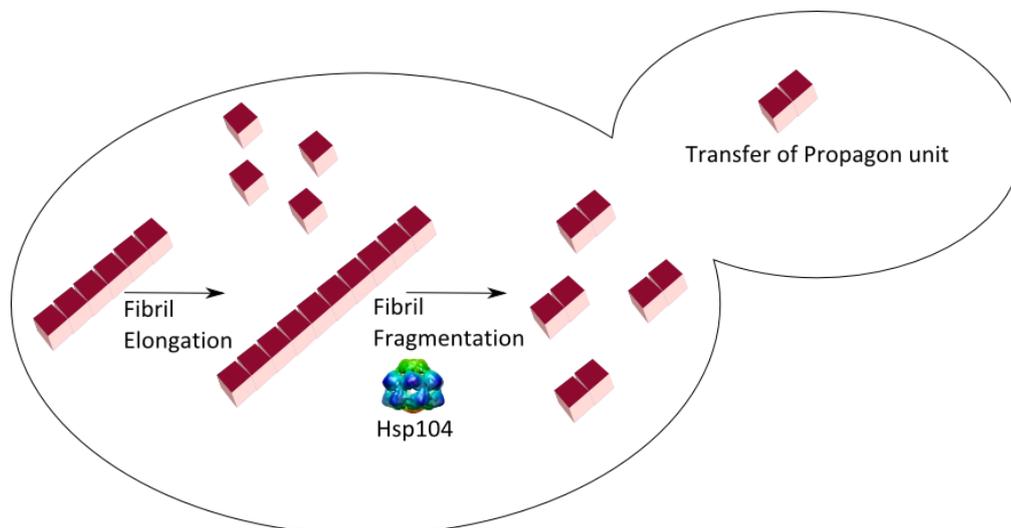
Figure 1.19   Prion Transmission Through Amyloid Fragmentation.
In *Saccharomyces Cerevisae* the chaperone (Hsp104) mediated fragmentation of Sup35 fibrils
generates propagon particles. The transfer of these particles from mother to daughter is
required for transfer of the *[PSI*[+]*]* phenotype.

The work showing the prion like properties of [*URE3*] also drew parallels with the [*PSI*[+]]

phenotype in *Saccharomyces cerevisae* and postulated that it also was a prion of the yeast

protein Sup35 (Wickner, 1994)*.*

The [*PSI*[+]] phenotype in *Saccharomyces cerevisae* was reported by Brian Cox in 1965 as

evidence of a non-Mendelian trait (Cox, 1965). The observation of prions, in particular

[*PSI*[+]], in yeast, a highly tractable model organism, has allowed for great advances in understanding the processes involved in prion initiation and transmission. Using the yeast model system, it has been shown that transmission of the [*PSI*[+]] phenotype from mother to progeny is dependent on the formation of small 'Propagon' particles by the action of chaperone proteins such as Hsp104 (Fig. 1.19) (Byrne et al., 2009; Cox et al., 2003).

The link of prion phenotypes with the aggregation of proteins into amyloid deposits also offers an explanation for the existence of strains observed in prion systems due to their propensity to form different morphologies of fibrils. The morphological variation observed in this instance could be due to differences in assembly and lateral association of the fibrils. Sup35 is a 76kDa protein with 3 different domains: the N-domain is required for the formation of the amyloid state, the M-domain contains polar solubilising residues and the C-domain is responsible for the translation termination activity of the protein. One of the potential mechanisms of amyloid associated disease is the loss of activity of the aggregated proteins so a loss of translation termination would be expected on formation of Sup35 fibrils. The C domain of Sup35 is outside of the fibril core and exposed to solvent (Baxa et al., 2011) and can therefore potentially interact with other translation termination factors. Recent work has demonstrated that fibrils can indeed still take part in translation termination but the size of the aggregate has a significant role in the ability of the C-domain to carry out this process (Pezza et al., 2014).

The work in yeast has shown that contrary to the aggressive, life-threatening phenomenon observed in human prion diseases, the yeast equivalents are useful epigenetic switches enabling the conversion between phenotypes (True and Lindquist, 2000; True et al., 2004).

1.11   Functional Amyloids

The conversion of native protein to amyloid state is not necessarily detrimental to cells. As highlighted in table 1 there are several proteins that enable cells and whole organisms to perform functions which allow them to survive and spread.

Spidersilk is a multifunctional material used by spiders to suspend themselves in webs, catch prey and protect their offspring and requires an incredibly strong material. Spidroin, one of the components of dragline silk, undergoes a pH sensitive conversion from monomer to fibrillar form (Askarieh et al., 2010). The silk formed is incredibly strong, thought to be in the region of Kevlar, and is very flexible (Xu and Lewis, 1990), properties which allow it to fulfil its function.

1.12   Nanomaterials

Materials science is a field which develops and analyses novel materials with properties that are potentially useful, and has underpinned many of the technological advances in the past 100 years, e.g. polymers. The origin of nanomaterials can be traced back to the work of Sumio Iijima, in 1991, on the identification and characterisation of carbon nanotubes (Fig. 1.20) (Iijima, 1991).

The ability to generate materials with different properties on the nanoscale (<100nm) has the potential for impacting on many areas of research such as drug delivery and diagnostics, and is an area of great current interest. The size and relative surface area of nanomaterials make them different to traditional materials and also allows them to be used in conjunction with other materials, as a coating, to enhance or add properties such as conductance and water resistance.

In the macroscale, the determination of the strength of a material, its ability to withstand an applied load without failure, is an important characteristic of the substance. There are many different types of strength, i.e. yield strength and tensile strength, which is indicative

of the force in Newtons per metre (Nm^{-1}) at which a material fails to return to its unstressed state when the force is removed. The ability of a material to return to its original state after deformation is known as elasticity, and is described by the Young's Modulus (Elastic Modulus of a material, E). The characterisation of a material's properties on the nanoscale requires a different set of techniques to those required on the macroscale. Electron Microscopy and Atomic Force Microscopy have been utilised to great effect for analysing the structure of nanomaterials, and AFM also offers the ability to directly measure physical properties of nanomaterials through the application of force (Adamcik and Mezzenga, 2012).
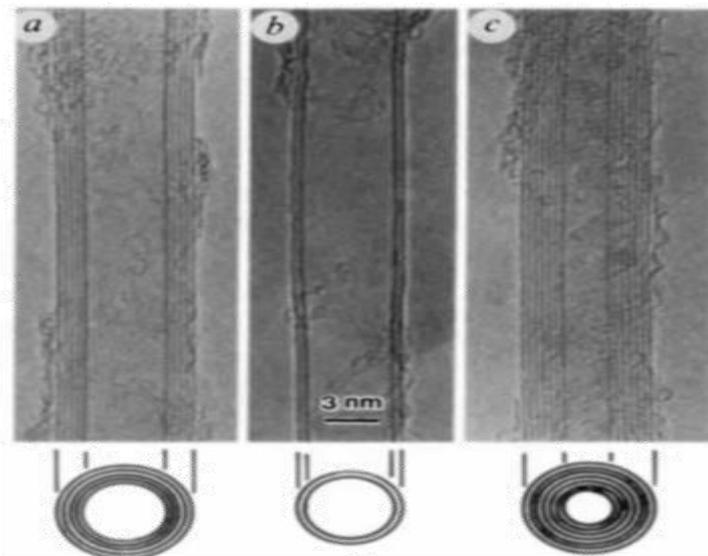
Figure 1.20   First Generation Nanomaterials.
Electron micrographs of graphitic carbon microtubules, a- tube of 5 graphitic sheets, b- tube of 2 graphitic sheets and c- tube of seven graphitic sheets. Adapted from (Iijima, 1991)

Nanomaterials derived from biological materials have the potential to offer many advantages over other materials including cost, readily available starting materials and the potential of using host organisms to generate the materials but it is their strength (Fig. 1.21) and relative ease of assembly that make amyloid an interesting starting point for bespoke nanomaterials (Knowles and Buehler, 2011).

The ability to make changes to the structures and properties of the products with relative ease by genetic modification of the organisms means that this field has a large overlap with the emerging fields of Chemical Biology and Synthetic Biology and has led to the formation of modified fibrils with different functionality (Knowles and Buehler, 2011). For example amyloid designs have allowed for the capture of carbon dioxide (Li et al., 2014), the addition of electrical conducting groups (Scheibel et al., 2003) and retroviral transduction (Li et al., 2014).
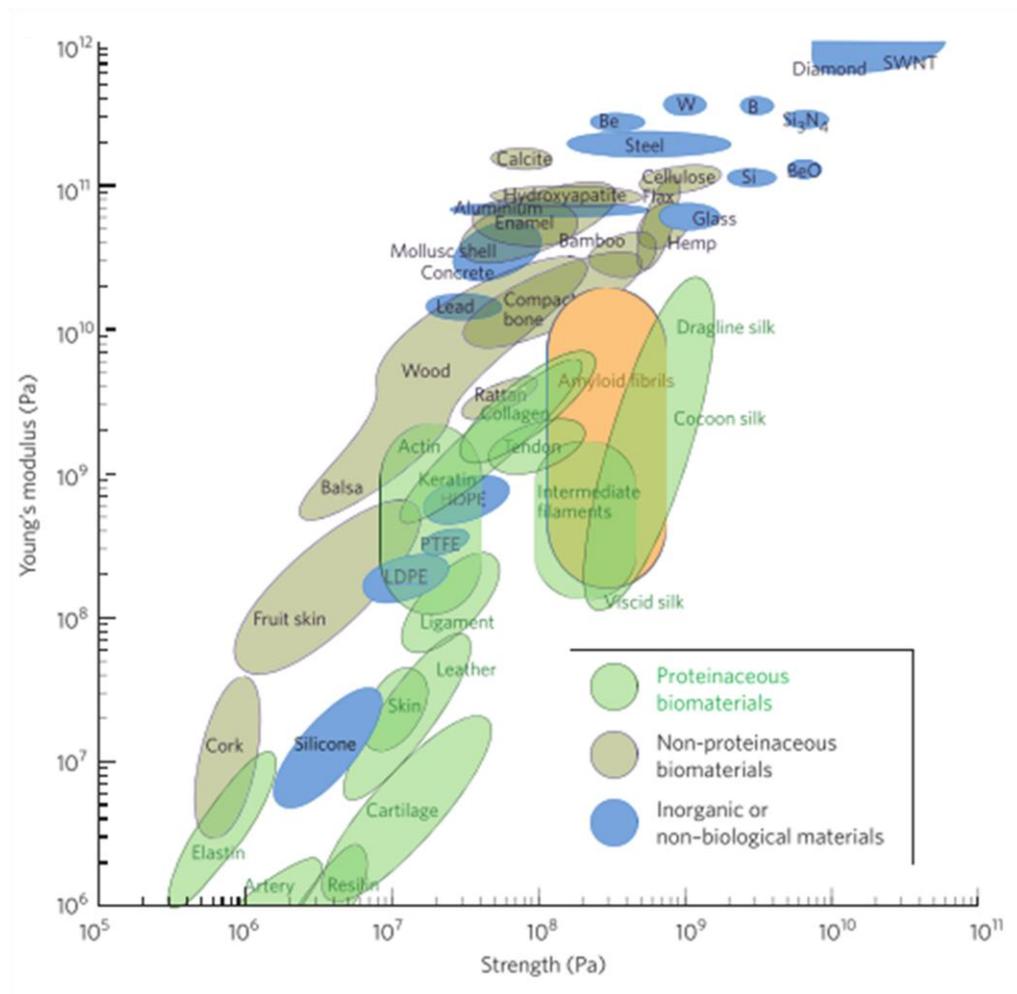
Figure 1.21   Amyloid has Exceptional Physical Properties.
A graph comparing the different physical properties (Strength vs Young's Modulus) of known materials varying from biological polymers to plastics and modern manufacturing materials, amyloid fibrils are one of the strongest nanomaterials. Adapted from (Knowles and Buehler, 2011)

1.13   Aims of thesis

The overall goals of the project were to find generic structural and mechanical properties of fibrils and how this affects their stability towards fragmentation. This study has important implications for understanding the response to breakage of different fibril types and how this then impacts on disease, infectivity and the properties as a nanomaterial.

To achieve these goals the project was separated into three main aims:

The aim of the first part was to develop quantitative and reproducible methods for the preparation of amyloid fibrils *in vitro* either from commercial protein sources or via the bacterial expression and purification of protein monomers, and to develop AFM imaging methods for quantitative imaging of these amyloid systems.

The second part aimed to investigate the resistance of fibrils to mechanical stress *in vitro* by monitoring the size distributions of amyloid fibrils subjected to fragmentation by stirring, and AFM imaging and subsequent MATLAB analysis to quantify particle size distributions and generate a model of fibril fragmentation.

The third part of the project was to develop novel methods of determining the population size of amyloid fibrils utilising chemical labelling strategies and DOSY NMR.

Materials and Methods

2.1   Materials

All media was sterilised in a Prestige Medical Desktop autoclave at 121°C, 15lb/in^2 for 11

mins. All *E.coli* liquid cultures were incubated at 37°C and shaken at 180rpm and solid

cultures were incubated overnight at 37°C overnight or for 48 hours at room temperature.

Table 2.1   Lysogeny broth composition

Substance	g/100ml
Tryptone	1
Yeast extract	0.5
NaCl	1
Agar (for solid media)	2

For media requiring antibiotic selection an appropriate antibiotic was added to the media

post sterilisation. For this thesis only ampicillin resistant strains were used therefore media

was supplemented with 100mg/ml ampicillin to give a 0.1% w/v concentration.

Table 2.2   *E.coli* Strains Utilised

Role	Name	Competence type	Genotype
Expression	BL21[DE3] pLysS	Chemical	F- ompT hsdSB (rB-mB-) gal dcm (DE3) pLysS (CamR)
Plasmid propagation	Top10	Chemical	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lacZ*ΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G

Table 2.3   Plasmids Used

Insert	Vector	Restriction sites for insert	Source
Sup35NMC	pET15b	*Nde*I/*Bam*HI	R. Melki (Krzewska and Melki, 2006)
Sup35NM	pET15b	*Nde*I/*Bam*HI	This Study
Sup35NM-Cys	pET15b	*Nde*I/*Bam*HI	This Study
Sup35NM-LCys	pET15b	*Nde*I/*Bam*HI	This Study
Sup35NM-LLCys	pET15b	*Nde*I/*Bam*HI	This Study
α-Synuclein	pRSETB	*Nde*I /HindIII	D. Smith (Cappai et al., 2005a)

Table 2.4   Primers Used for Sequencing and PCR

Name	Sequence	Restriction Enzyme	Source	Role
Sup35 forward	GGAATTCCATATGTCGGATTCAAACCAAGGC	Nde1	This work	PCR
Sup35 NM reverse	CGGGGATCCCGATGATGT,ATCGTTAACAACTTCGTCATCCAC	BamH1	This work	PCR
Sup35 NM-LCys rev	CGCGCGGGATCCCGCTATTAGCAGCCGGCTGCCGCGCGCTGCC GCCGCTGCCATCGTTAACAACTTCGTCATCCAC	BamH1	This work	PCR
Sup35.1326	CCACACATAGCCATATCAAC	-	This work	Sequencing
Sup35_f_1456	CGACCGTCACATCAATGCTCC	-	This work	Sequencing
T7p	CGACCGTCACATCAATGCTCC	-	Commercial	Sequencing
T7Term	CTAGTTATTGCTCAGCGG	-	Commercial	Sequencing

Table 2.5   Antibodies used for Western Blotting

Name	Epitope	Organism Raised	Source	Dilution
MT50	Sup35 MC	Rabbit	University of Kent	1 in 5000
Anti-rabbit antibody FITC conjugated	Rabbit IgG	Goat	Sigma, F9887	1 in 2000

Table 2.6   Chemicals

Chemical	Cat.no.	Supplier	Relative Molecular Mass	Density
Propargylamine	P50900	Sigma-Aldrich	55.08	0.86
Fluorescein Isothiocyanate Isomer 1	F4247	Sigma-Aldrich	389.382	
3,3,3-Trifluoropropionyl Chloride	JI6800	Manchester Organics	146.5	1.00
11-Azido-3,6,9-trioxaundecan-1-amine	17758	Sigma Aldrich	218.25	1.10
N-Succinimidyl iodoacetate (SIA)	22349	Thermofisher	283.02	
4,4,4-Trifluorobutylamine	C19847	Manchester Organics	127	0.93
N-[(1R,8S,9S)-Bicyclo[6.1.0]non-4-yn-9-ylmethyloxycarbonyl]-1,8-diamino-3,6-dioxaoactane	745073	Sigma Aldrich	324.42	
Dichloromethane		Fluka	84.93	1.33
DMF		Sigma	73.09	0.944
Diethyl ether		Fluka	74.12	0.713
Diisopropylethylamine		Sigma	129.24	0.742
Pyridine		Fluka	79.1	0.982
Sodium Ascorbate	A7631	Sigma	198.11	
Tris(3-hydroxypropyltriazolylmethyl)amine	762342	Sigma	434.5	
4,4,4-Trifluorobut-2-en-1ol	007510	Fluorochem	126.08	
1-Ethynyl-3,5,-bis(trifluoromethyl)benzene	630241	Sigma	238.13	1.346
NHS-PEG$_{12}$-Azide	26131	Thermo Scientific	740.79	
Copper (II) Sulphate pentahydrate	209198	Sigma	249.69	

2.2   DNA Agarose Gel

Table 2.7   Composition of 50 x TAE Buffer

Compound	Amount
Tris base	242g
0.5M EDTA	100ml
Glacial acetic acid	57.1ml
mQ Water	Upto 1L

Tris base was dissolved in 600ml of mQ H_2O and treated with glacial acetic acid (57.1ml) and 0.5M EDTA (100ml). The resulting solution was then diluted to 1L by the addition of mQ water.

Table 2.8   Composition of a Typical 1% Agarose gel

Material	Amount
1 x TAE buffer	50ml
Agarose	0.5g
Ethidium bromide	1µl

2.2.1   Electrophoresis Conditions

The agarose (Melford, MB1200) was dissolved in 1 x TAE buffer by microwaving on full power for a total of 80 seconds punctuated by careful swirling to mix the solution. The gel was allowed to cool slightly before addition of 1µl of 10mg/ml ethidium bromide in H_2O (Sigma Aldrich, E1510) per 50ml of solution. Ethidium bromide was replaced with 5µl of 10000 x DMSO concentrate of SYBR® Safe (Invitrogen, S33102) for gel imaging only. A lane comb was then added and the gel was allowed to cool and set solid. The gel was aligned so that the lane comb is at the anode of the gel tank and then covered with 1 x TAE buffer.

DNA samples were treated with 5 x sample loading buffer (Bioline, BIO-37045) and loaded into the gel, along with 3µl of Bioline Hyperladder™ 1kb (Bioline, BIO-33053) before running at 80V until loading dye passes the mid-point of the gel.

SYBR® Safe stained gels were imaged using a Fujifilm FLA-5100 scanner and Image Reader FLA-5000 with FITC absorption and emission characteristics whilst ethidium bromide stained gels were visualised on a UVP transilluminator for band excision.

2.3   Agarose Gel Extraction

The appropriate bands were excised from the gel using a clean, sharp scalpel and placed in eppendorf tubes of known mass. The amount of gel in the tubes was calculated and the

DNA extracted from the gel following the manufacturer's instructions, using either Qiagen's QIAquick Gel Purification Kit or Thermofisher's GeneJET Gel Extraction Kit.

2.4 DNA Quantification

The concentration of DNA samples was measured using an Eppendorf Biophotometer or a BMG Labtech LVis plate in conjunction with a BMG labtech Spectrostar Nano or FluorStar Omega platereader, by monitoring the absorbance of the sample at 260nm and correcting against a blank sample. The extinction coefficient of double stranded DNA is $0.02(\mu g/ml)^{-1}cm^{-1}$ therefore according to the Beer Lambert Law ($A=\varepsilon cl$) an absorbance reading of 1 is equal to 50 µg/ml.

2.5 Preparation of Competent *E.coli* cells

An *E.coli* strain (Top10 or BL21 [DE3] pLysS) (50µl) was inoculated into LB (Top10 - without antibiotic and pLysS with chloramphenicol) (4ml) and incubated at 37°C overnight.

This overnight culture was then diluted 1 in 50 into 100ml of LB (antibiotic as above) and incubated at 37°C until the OD600 was between 0.4 and 0.45. The culture was then placed on ice for 10 mins, all subsequent manipulations were carried out on ice, in a cold room and using a pre-chilled centrifuge. The culture was harvested by centrifugation at 3000rpm for 8 minutes at 4°C and the supernatant discarded. The cell pellet was resuspended (by swirling and <u>not</u> pipetting) in 100mM sterile calcium chloride (25ml) and incubated on ice for 10 mins. The resulting suspension was clarified by centrifugation at 3000rpm for 8 minutes at 4°C and the supernatant discarded. The cell pellet was resuspended in sterile 100mM calcium chloride, 30% (v/v) glycerol (2ml) and then divided into 100µl aliquots in prechilled 1.5ml eppendorf tubes and then stored at -80°C.

2.6   *E.coli* Transformation

A 100µl aliquot of competent cells (Top10 or BL21[DE3] pLysS) were removed from -80°C storage and placed on ice to thaw. 1µl of circular plasmid DNA (~100ng/µl) was added to the sample and the cells were incubated on ice for 30 minutes before subjecting to heat shock at 42°C, in a water bath, for 60 seconds. The heat shocked cells were immediately placed back on ice and then diluted with 1ml of LB (without antibiotic) and incubated at 37°C for 1 hour. 100µl of this transformation mixture was then spread onto an antibiotic selective LB plate, antibiotic choice is dependent on plasmid transformed, and incubated at 37°C overnight or room temperature over the weekend.

2.7   Polymerase Chain Reaction (PCR) general method

PCR reactions were carried out using a Roche High Fidelty PCR kit (Roche, Cat. no. 11732641001) according to the manufacturer's instructions using bespoke forward and reverse primers (Table 2.5) and a parent DNA sample (Table 2.3) with the following compositions (Table 2.9).

Table 2.9   **Typical PCR Reaction Composition**

Material	Concentration	Volume
Parent DNA (plasmid)	>50ng/µl	2µl
Primer 1	100pmol/µl	1µl
Primer 2	100pmol/µl	1µl
dNTP mixture	See Table 2.10	5µl
PCR buffer	10 x concentrate	5µl
Taq Polymerase	1U/µl	1µl
H$_2$O		35µl

Table 2.10   **Deoxynucleotide Triphosphate (dNTP) mixture**

Material	Volume
Deoxyadenosine Triphosphate (dATP) 100mM	2µl
Deoxythymidine Triphosphate (dTTP) 100mM	2µl
Deoxyguanosine Triphosphate (dGTP) 100mM	2µl
Deoxycytidine Triphosphate (dCTP) 100mM	2µl
Double distilled water	98µl

The above reagents were combined, with the enzyme added to the mixture last. The samples were then heated using a Techne TC-3000 thermocycler with these basic parameters (Table 2.11), but the melting temperature of the primer guiding the choice of annealing temperature used.

Table 2.11   Typical Thermocycler Parameters

Program method	Temperature (^{o}C)	Time
Lid temperature	105	
Initial denaturation	94	3 minutes
Annealing	55	45 seconds
Extension	72	3 minutes
Cycle (30 times)		
Final extension	72	5 minutes

The samples were then purified following the PCR purification protocol.

2.8   PCR Purification

Purification of PCR reactions was achieved using either Qiagen's QIAquick PCR purification kit or Thermofisher's GeneJET PCR purification kit, following the manufacturer's instructions.

2.9   DNA Digestion using Restriction Enzymes

DNA digestions were carried out using Promega enzymes and following the manufacturer's guidelines, briefly the DNA was diluted with mQH$_2$O and a buffer appropriate for the enzymes to be used. According to the Promega Restriction Enzyme Tool the appropriate buffer for the combination of restriction enzymes *NdeI*/*BamHI* is the Promega buffer D (Table 2.12).

Table 2.12   Table Showing the Promega Buffer Compatibility of BamHI and NdeI*

Enzyme	Buffer Supplied	A	B	C	D	Multi-Core	E	H	Incubation Temp
BamHI	E	75-100%*	75-100%	75-100%	50-75%	75-100%	100%	50-75%	37°C
NdeI	D	<10%	<10%	25-50%	100%	25-50%			37°C

*This table shows the appropriate choice of Promega buffer for the digestion of DNA with the Promega BamHI and NdeI restriction enzymes. Adapted from the Promega Restriction Enzyme tool.

The composition of the reaction mixtures were the same as described by the manufacturer's instructions (Table 2.13). The enzymes, stored on ice at all times, were then added to a solution of the DNA and buffer, mixed and centrifuged. The resulting solutions were then incubated at 37°C for 2 hours.

Table 2.13   Typical Restriction Enzyme Reaction Composition

Component	Concentration	Volume
DNA sample	>50ng/µl	15µl
Restriction Enzyme 1	10U/µl	1µl
Restriction Enzyme 2	10U/µl	1µl
Buffer	10 x concentrate	4µl
mQH$_2$O		20µl
Total volume		41µl

The digestion samples were then purified either by preparative agarose gel and extraction of the DNA from the gel or by PCR purification (depending on the size of the DNA sample).

2.10   DNA ligation

Ligation of digested DNA fragments together was achieved using Roche's Rapid DNA Ligation kit (Roche, 11635379001) according to the manufacturer's instructions. Briefly, insert and vector were combined in a 10:1 ratio following equation 1.

Equation 1   Equation for calculating the amount of insert required for ligation reactions.

(Amount of vector (ng) x Size of insert (kb))/ Size of vector (kb) x Ratio of insert/Vector = ng of insert

This was then diluted with water and DNA dilution buffer added, mixed and then centrifuged. Ligation buffer was added to this solution followed by T4 DNA ligase and the reaction mixed, centrifuged and incubated at room temperature for 2 hours.

The standard reaction composition, where X comes from equation 1 and the DNA concentration and Y is the amount required to make the total volume 20.5µl is shown in Table 2.14.

Table 2.14   Typical Restriction Enzyme Reaction composition

Material	Concentration	Volume
Vector	As above	2µl
Insert	As above	Xµl
T4 Ligation Buffer	2 x concentrate	10µl
DNA dilution buffer	5 x concentrate	2µl
T4 DNA Ligase	5U/µl	0.5µl
mQH$_2$O		Yµl
Total volume		20.5µl

10µl of this solution was then transformed into Top10 competent cells according to the transformation protocol (2.6). 100µl of the transformation culture was spread onto antibiotic LB plates and the residual 900µl of the transformation mixture was pelleted at 3000rpm for 5 mins, 800µl of the supernatant removed and the cell pellet was gently resuspended in the remaining solution and spread onto a separate plate to make sure transformants were observed.

2.11   Plasmid isolation and purification (miniprep)

Isolation and purification of plasmid DNA from *E.coli* cultures was achieved using either Qiagen's QIAprep® Spin MiniPrep kit or Thermofisher's GeneJET Plasmid MiniPrep kit, following the manufacturer's instructions.

2.12   DNA sequencing

DNA sequencing was carried out with primers described in the Table 2.6 using the service

provided by Beckmann Coulter Genomics and the sequence analysis was carried out using

the ApE – A Plasmid Editor software.

2.13   Sodium Dodecylsulphate Polyacrylamide Gel Electrophoresis - SDS PAGE

2.13.1   General SDS PAGE Reagents

Table 2.15   4x Sample Loading Buffer Composition

Material	Amount
0.5M Tris-HCl, pH 6.8	1ml
10% w/v Sodium Dodecylsulphate (SDS)	1.6ml
50% v/v Glycerol	5ml
Bromophenol blue	A few grains
2-Mercaptoethanol (added just before use)	5% v/v volume

Table 2.16   Coomassie Stain Composition

Material	Amount
Coomassie Brilliant Blue (R or G)	0.4g
Methanol	100ml
Glacial Acetic Acid	20ml
H_2O	280ml

Table 2.17   Commassie Destain Composition

Material	Amount
Methanol	500ml
Glacial Acetic Acid	200ml
H_2O	500ml

Table 2.18   General SDS PAGE materials

Material
40% w/v Ammonium persulphate (stock solution divided into 100µl aliquots and stored at -20°C)
30% Acrylamide solution (29:1 acrylamide:bisacrylamide) – from Biorad, 161-0156
40% Acrylamide solution (37.5:1 acrylamide:bisacrylamide) – from Biorad, 161-0148
70% v/v Ethanol solution
Tetramethylethylenediamine (TEMED) – from Sigma Aldrich, T8133
2-Mercaptoethanol – Sigma Aldrich, M6250
Bromophenol blue – Fisher Chemicals, B/4630/46
Coomassie Brilliant blue R250 – BioRad, 161-0400
70% v/v Glycerol

2.13.2   Tris-Glycine gel

Table 2.19   Stacking Gel Buffer Composition

Material
1M Tris
0.4% w/v sodium dodecylsulphate (SDS, from 10% w/v stock)
pH 6.8

Table 2.20   Resolving Gel Buffer Composition

Material
1.5M Tris
0.4% w/v sodium dodecylsulphate (SDS, from 10% w/v stock)
pH 8.8

Table 2.21   Composition of 10 x Tris-Gly Running buffer

Material
0.25M Tris-base
1.92M glycine
1.5% w/v SDS

Table 2.22   Resolving gel composition*

Material	7.5%	10%	12.5%	15%	
30 % Acrylamide	2.7ml	3.6ml	4.5ml	5.4ml	
H_2O	5.4ml	4.5ml	3.6ml	2.7ml	
Resolving gel buffer	2.7ml	2.7ml	2.7ml	2.7ml	
40% w/v APS	40µl	40µl	40µl	40µl	ADDED LAST
TEMED	5 µl	5 µl	5 µl	5 µl	ADDED LAST

*The choice of % gel is dependent on molecular weight of protein analysed. APS and TEMED added simultaneously.

Table 2.23   Stacking Gel composition

Material		
30 % Acrylamide	1ml	
H_2O	4.2ml	
Stacking gel buffer	1.75ml	
40% w/v APS	40µl	ADDED LAST
TEMED	5 µl	ADDED LAST

2.13.2.1   Tris-Glycine Electrophoresis conditions

The gels were prepared and run using either a Biorad Protean 3 system or Invitrogen Novex Mini Cell system (with empty Novex cartridges). The resolving gel was placed in to the cassette first, leaving enough space for the lane comb, and 70% w/v ethanol placed on top of the gel solution. When the gel had set the ethanol solution was disposed of and any residue was removed using filter paper. The stacking gel was then placed on top of the resolving gel and the lane comb was inserted into the solution and the gel was left to set.

The samples were treated with 4 x sample loading dye, treated with 5% v/v 2-mercaptoethanol added just prior to use, to give an overall 1 x concentration. The samples were heated to 95°C for 5 mins and then centrifuged at 2000rpm for 30 seconds.

The gel was placed into the gel tank and the internal chamber filled with 1 x Tris-Gly Running buffer and the external chamber was filled with 1 x Tris-Gly Running buffer so that the electrode was covered. The gel was loaded with 3µl of PageRuler Prestain Ladder (Thermofisher, 26616) and treated samples prior to connecting to a power pack and

running at 180V until the Bromophenol blue from the sample loading buffer was about to exit the gel.

The gel was removed from the cartridge and either scanned for fluorescent bands (FITC) using a Fujifilm FLA-5100 scanner and Image Reader FLA-5000 with the appropriate filters selected or stained with Coomassie stain for 1 hour and then destained with coomassie destain until the background blue colour was sufficiently removed to give good sample band definition. The gel was then scanned using a flat bed scanner, either an Epson Perfection V33 or Epson Perfection 3200 photo in conjunction with Epson Scan software.

2.13.3   Tris-Tricine gel

Method from http://wolfson.huji.ac.il/purification/Protocols/Tricine.html

Table 2.24   Gel Buffer Composition

Material
3M Tris-HCl
0.3% w/v SDS (added after solution pH changed)
pH8.45

Table 2.25   Composition of Tris-Tricine Running Buffers

10 x Cathode Buffer	10 x Anode Buffer
1M Tris-base	2.1M Tris base
1M tricine	pH8.9
pH 8.25	
1% w/v SDS (added after solution pH changed)	

Table 2.26   Tris-Tricine Gel Composition

Solutions	Resolving gel	Stacking gel	
40% Acrylamide solution	5ml	1.05ml	
Gel Buffer	5ml	2.5ml	
70% v/v Glycerol	2ml	0ml	
H$_2$O	3.1ml	6.7ml	
40% w/v APS	16.5µl	20µl	ADDED LAST[a]
TEMED	6.6µl	8µl	ADDED LAST[a]

[a]APS and TEMED added simultaneously.

2.13.3.1   Electrophoresis conditions

The gels were prepared using an Invitrogen Novex Mini Cell system with empty novex cartridges. The gel was prepared and run according to the same method as the tris-glycine gel (2.13.2.1) except that the inner gel compartment was filled with 1 x Cathode buffer, the outer compartment was filled to above the electrode with 1 x Anode buffer and the gel was run at 4°C.

2.14   Western Blot

2.14.1   Buffers

Table 2.27   Composition of Transfer Buffer

Material	Amount	
Glycine	2.9g	
Tris base	5.8g	
SDS	0.37g	
Methanol	200ml	Added just before use
mQ H$_2$O	800ml	

Table 2.28   Composition of Western Blotting Buffers

10 x Tris Buffered Saline (TBS)	TBS-T	TBS-M
500mM Tris	1 x TBS	1 x TBS
1.5M NaCl	0.05% w/v Tween-20	5% w/v Marvel skimmed milk powder
pH 7.8		

2.14.2   Transfer Conditions

An unstained acrylamide gel was acquired according to the SDS PAGE protocol 2.13 and the gel removed from the cassette.

Hybond ECL (Enhanced Chemiluminesence) membrane (GE Healthcare, RPN303D) and blotting paper (2 pieces) were cut to the same size as the acrylamide gel (8.25cm x 5cm). The ECL membrane and acrylamide gel were incubated together in transfer buffer for 20 minutes.

A piece of blotting paper was soaked in transfer buffer and placed onto the metal plate of the western transfer apparatus and rolled to remove air bubbles. The ECL membrane was placed on top of the blotting paper, the acrylamide gel was placed on top of this and the stack completed by the addition of the last piece of blotting paper, soaked in transfer buffer, to the top. The assembly was completed, after removal of air bubbles from the stack, by placing the metal top plate on the western transfer apparatus. Transfer from the acrylamide gel to the ECL membrane was achieved by applying a current of 10V to the stack for 40 minutes.

The membrane was removed from the stack and 'blocked' with TBS-M for 10 minutes before incubating with a primary antibody, diluted appropriately, Table 2.5, in 15ml of TBS-M, for 2 hours at room temperature or overnight/over the weekend at 4°C. The primary antibody was then removed and the membrane was washed with fresh TBS-M before incubating with an appropriate secondary antibody, diluted in 15ml of TBS-M, for 1 hour at room temperature. The antibody solution was removed and the membrane was washed with TBS-T for five minutes. The wash was repeated a further two times before visualisation.

FITC labelled secondary antibody visualisation was achieved using a Fujifilm FLA-5100 scanner and Image Reader FLA-5000 with the appropriate filters selected.

2.15   Protein Concentration determination

Protein concentration was measured using an Eppendorf Biophotometer or BMG Labtech

LVis plate (S/N 680-0194) in conjunction with a BMG labtech platereader (either FLUOstar

Omega or SPECTROstar[Nano]) by monitoring the absorbance of the sample at 280nm and

correcting against a blank sample.

The protein concentration is linked to the absorbance of the sample through the Beer

Lambert Law (A=εcl) where the extinction coefficient of the protein is dependent on the

amino acid residues contained in the protein and can be predicted (Table 2.29).

The concentration derived, mg/ml, can then be converted to molarity using the molecular

weight of the protein.

Table 2.29   Extinction Coefficients and Molecular Weights of Proteins Used in This Thesis

Protein	$A^{0.1\%}_{280nm}$	Molecular Weight	Data Source
Sup35NMC	0.945	78714.2	Expasy ProtParam[a]
Sup35NM	0.971	30681.8	Expasy ProtParam[a]
Sup35NM-LCys	0.949	31388.5	Expasy ProtParam[a]
α-Synuclein	0.412	14460.1	Expasy ProtParam[a]
β-Lactoglobulin	0.96	18276	Sigma MSDS / (Croguennec et al., 2004)
Lysozyme	2.64	14307	Sigma MSDS
Insulin B Chain	0.876	3495.89	Sigma MSDS / Expasy ProtParam[a]

[a] http://web.expasy.org/protparam/

2.16   Protein Expression

A single colony from a freshly streaked ampicillin selective LB plate was inoculated into

30ml of ampicillin selective LB and incubated at 37°C overnight.

Table 2.30   Protein Expresssion Reagents

Material
1M Isopropyl-β-D-1-thiogalactopyranoside (IPTG)
Liquid LB medium (see Table 2.1)
Ampicillin (see Table 2.1)

20ml of the overnight culture was removed and centrifuged at 3000rpm for 10 minutes and the supernatant was removed before the cell pellet was resuspended in 5ml of fresh LB medium. This suspension was then used to inoculate 1L of LB with ampicillin medium which was then incubated at 37°C until the OD600 of the culture was ~0.6. Protein expression in the culture was then initiated by addition of 1ml of 1M IPTG, 1mM working concentration, to the culture and then incubated at 37°C for 4 hours. The cells were harvested by centrifugation at 4000rpm for 10 minutes and then either taken on to purification or snap frozen and stored at -80°C until purification.

2.17   Protein Purification

2.17.1   Buffers and Reagents

Table 2.31   **Reagents required for Protein Purification**

Material
100mM Phenylmethanesulphonyl fluoride in propan-2-ol
200mM Nickel (II) chloride (NiCl$_2$)

Table 2.32   **Composition of Sup35 Purification Buffers***

Melki Lysis Buffer (MLB)	Melki Lysis Buffer Denaturing (MLBG)	Denaturing Sizing Buffer (DSB)
20mM Tris-HCl	20mM Tris-HCl	20mM Tris-HCl
1M NaCl	1M NaCl	0.5M NaCl
20mM imidazole	20mM imidazole	6M guanidine HCl
100mM PMSF (added just prior to usage)	6M guanidine HCl	
pH8	pH 8	pH 8

*For eluting from the Ni-NTA resin the concentration of imidazole in the buffers was increased to 0.5M.

Table 2.33   **Composition of α-Synuclein Purification Buffers**

α -Synuclein Lysis Buffer (SLB)	α-Synuclein Ion Exchange Entry Buffer (SEB)	α -Synuclein Ion Exchange Elution Buffer (SElB)
20mM Tris	20mM Tris	20mM Tris
100mM NaCl		1M NaCl
pH 7.5	pH 7.5	pH 7.5

2.17.2 Nickel Affinity Resin (Ni-NTA)

The Ni-NTA resin was prepared by washing GE Healthcare Chelating Sepharose™ Fast Flow resin, 17-0571-01, with 2 column volumes of mQH$_2$O, 1 column volume of 200mM NiCl$_2$, 1 column volume of mQH$_2$O and finally 2 column volumes of lysis buffer.

2.17.3 Sup35 Full length

2L of *E.coli* cell pellet containing Sup35NMC was resuspended in 20ml of MLB, cooled to 4°C, and lysed by sonication at 18µm amplitude for four 30 second bursts with 30 second intervals between each burst. The lysate was then clarified by centrifugation at 15000rpm for 30 minutes at 4°C and the supernatant was incubated with 1ml of Ni-NTA resin at 4°C for 30 minutes. The resin was collected by centrifugation at 4000rpm for 5 minutes, the supernatant placed to one side, and the resin washed with 10ml of MLB. The resin was then washed with 15ml of MLBG prior to elution of the protein with 1ml of MLBGE.

5ml of material from the Ni-NTA purification was further purified by size exclusion chromatography using a HiLoad™ 16/600 Superdex™ 200pg, from GE Healthcare (28-9893-35), using DSB on an ÄktaPrime Plus system at a flow rate of 0.3ml/minute. The appropriate fractions were combined and frozen at -80°C.

2.17.4 Sup35NM

2L of *E.coli* cell pellet containing Sup35NM was resuspended in 20ml of MLBG and lysed by sonication at 18µm amplitude for six 30 second bursts with 30 second intervals between each burst. The lysate was then clarified by centrifugation at 15000rpm for 30 minutes and the supernatant was incubated with 1ml of Ni-NTA resin at room temperature for at least 2 hours. The resin was collected by centrifugation at 4000rpm for 5 minutes, the supernatant placed to one side, and the resin washed with 15ml of MLBG. The protein was eluted from the Ni-NTA resin by washing the resin with 1ml of MLBGE.

5ml of material from the Ni-NTA purification was further purified by size exclusion chromatography using a GE Healthcare HiLoad™ 16/600 Superdex™ 200pg, from GE

Healthcare (28-9893-35), using DSB on an ÄktaPrime Plus system at a flow rate of 0.3ml/minute. The appropriate fractions were combined and frozen at -80°C.

2.17.5   Sup35NM-LCys

2L of Sup35NM cell pellet were resuspended in 20ml of MLBG + 2.5mM 2-Mercaptoethanol and lysed by sonication at 18μm amplitude for six 30 second bursts with 30 second intervals between each burst. The lysate was then clarified by centrifugation at 15000rpm for 30 minutes and the supernatant was incubated with 1ml of Ni-NTA resin at room temperature for at least 2 hours. The resin was collected by centrifugation at 4000rpm for 5 minutes, the supernatant placed to one side, and washed with 15ml of MLBG. The protein was eluted from the Ni-NTA resin by washing the resin with 1ml of MLBGE.

5ml of the sample from the Ni-NTA purification was defrosted and 2M dithiothreitol (DTT) added to a final concentration of 50mM and incubated at RT for 1 hour. This material was further purified by size exclusion chromatography using a GE Healthcare HiLoad^{TM} 16/600 Superdex^{TM} 200pg, from GE Healthcare (28-9893-35), using DSB on an ÄktaPrime Plus system at a flow rate of 0.3ml/minute. The appropriate fractions were combined and frozen at -80°C.

2.17.6   α-Synuclein

1L of *E.coli* cell pellet containing α-synuclein was resuspended in 20ml of SLB + Roche Complete Protease Inhibitor tablets (Roche, 11697498001) and sonicated on ice at 18μm amplitude for 10 second bursts with 10 second intervals for a total of 180 seconds. The lysate was centrifuged at 13000rpm for 30 minutes at 4°C and the supernatant collected and acidified to pH3.5 using 1M HCl. The suspension was then rolled at room temperature for 30 minutes and the precipitate collected by centrifugation at 13000rpm for 10 minutes. The supernatant was then neutralised, using 1M NaOH, and then treated with ammonium sulphate, $(NH_4)_2SO_4$, to give a 30% solution, as determined using saturation tables

(https://www.sigmaaldrich.com/content/dam/sigma-

aldrich/docs/Sigma/Bulletin/a5479bul.pdf), and then incubated at 4°C for 20 minutes.

The suspension was then clarified by centrifugation at 13000rpm for 20 minutes and the

supernatant treated with further $(NH_4)_2SO_4$ to give a 50% solution overall (using the same

table) and then rolled for 30 minutes at 4°C. The resulting precipitate was collected by

centrifugation at 13000rpm for 30 minutes at 4°C and then dissolved in 2.5ml SEB and the

excess ammonium sulphate was removed by buffer exchange into SEB using a PD10

column, GE Healthcare 17-0851-01, according to the manufacturer's instructions.

The resulting material was then purified by ion exchange chromatography on an ÄktaPrime

Plus system with a strong anion exchange HiTrap^{TM} Q FF 5ml cartridge (GE Healthcare, 17-

5156-01) and the standard ion exchange method with a linear gradient from SEB to SElB.

The desired protein fractions were collected and assessed by SDS PAGE analysis.

The sample was then snap frozen and stored at -80°C or transferred directly into fibril

forming buffer using a PD10 column according to the manufacturer's instructions.

2.18   *In vitro* Polymerisation

2.18.1   Materials

Table 2.34   Composition of Fibril forming buffers

Sup35 fibril forming buffer (SFFB)	α -Synuclein fibril forming buffer (SyFFB)	Lysozyme, β-lactoglobulin and insulin B chain fibril forming buffer
20mM Sodium phosphate 50mM NaCl	25mM Sodium phosphate	10mM HCl
pH 7.4	pH 7.5	pH2

All buffers were filtered using a 0.2µm syringe filter prior to use in polymerisation reactions.

Table 2.35   Commercial Proteins Utilised for *in vitro* Polymerisation

Protein	Purity	Supplier	Catalogue No.
Lysozyme from chicken egg white	≥90%	Sigma Aldrich	L6876
β-Lactoglobulin B, from bovine milk	≥90%	Sigma Aldrich	L8005
Insulin Chain B Oxidised, from bovine pancreas	≥80%	Sigma Aldrich	I6383

2.18.2   Sup35 Full length, Sup35NM, Sup35NM-LCys Polymerisation

2.5ml of pure protein in DSB was exchanged into 3.5ml of SFFB using a PD10 column according to the manufacturer's instructions. The material was then diluted to the appropriate concentration (10µM, 5µM etc) and incubated at 37°C without agitation. Post buffer exchange, using a PD10 column, all manipulations of the Sup35 solutions were with low bind tips (VWR 53503-294 (with Spidersilk^{TM} surface)), and Protein LoBind 1.5ml tubes (Eppendorf, 022431081).

2.18.3   α-Synuclein

2.5ml of pure protein in SElB was buffer exchanged into SyFFB using a PD10 column according to the manufacturer's instructions. The sample was treated with 0.0002% w/v sodium azide, syringe filtered using a 0.2µm filter and incubated at a concentration >200µM and at 37°C with orbital agitation (160rpm) for 2 weeks.

2.18.4   Lysozyme (pH2)

Lysozyme was dissolved in 10mM HCl to give a protein concentration of ~20mg/ml and then syringe filtered through a 0.2µm membrane. The resulting solution was diluted to give a concentration of at least 699µM (10mg/ml), heated to 60°C for 144 hours in a heat block and then allowed to cool to room temperature.

2.18.5   β-Lactoglobulin (pH2)

β-Lactoglobulin was dissolved in 10mM HCl to give a protein concentration of ~20mg/ml and then rolled for 1 hour. The solution was filtered through a 0.2μm membrane and the resulting solution was diluted to give at least 300μM (5.5mg/ml), and then heated to 90°C for 5 hours in a heat block and then allowed to cool to room temperature.

2.18.6   Insulin chain B (pH2)

Insulin chain B was dissolved in 10mM HCl to give a protein concentration of ~10mg/ml and then syringe filtered through a 0.2μm membrane. The resulting solution was diluted to give a concentration of at least 500μM (1.7mg/ml), and then heated to 60°C for 48 hours in a heat block and then allowed to cool to room temperature.

2.19   Fibril Fragmentation

500μl of a 120μM fibril suspension (according to initial monomer concentration) was put into a clear HPLC vials (HPLS, Chromacol, V944/950) with a stirring bar (cylindrical 8 x 3mm, VWR, 442-4520) and placed onto the centre of an IKA color squid stirrer plate (IKA, 0003671000), held in position by double sided tape. The sample was then stirred at 1000rpm under the cover of an aluminium foil covered plastic box to prevent the interaction of light with the sample. At appropriate time points aliquots of the fragmentation were removed, diluted appropriately and adsorbed onto freshly cleaved mica surfaces.

The surfaces were then imaged according 2.33.

2.20   Residual Monomer Assay

100μl of 120μM unfragmented fibrils and of the fibril fragmentation were centrifuged at 75000rpm for 15 minutes using a Beckman Coulter Optima Max Ultracentrifuge. The top 10μl of the supernatant from each was removed and then treated with 4 x SDS PAGE sample loading dye and heated to 95°C for 5 minutes. These samples were then loaded

onto a tris-tricine or 15% tris-glycine SDS PAGE along with a serial dilution of the monomeric protein from 60μM to 1.875μM. The gel was then run according to the SDS PAGE protocol, 2.2.13, with Coomassie staining.

The stained gel was then scanned on a flat bed scanner, either an Epson Perfection V33 or Epson Perfection 3200 photo in conjunction with Epson Scan software. The resulting image was then analysed using the densitometry component of the Aida Image Analyzer V4.15 software.

The fibril and fragmented fibril samples were then compared with the serial dilution and each other to show any change in the amounts of monomers not incorporated into aggregates.

2.21   Fibril Formation Kinetics – Thioflavin T (ThT) Assay

Monomeric protein solutions were transferred into fibril forming buffer using a PD10 column according to the manufacturer's instructions. The samples were diluted to the appropriate concentration, typically 10μM, 5μM and 2.5μM, with fibril forming buffer. 1μl of 1mM ThT solution was added to 100μl of protein monomer solution. The final volume is 100μl per well using a black 96-well plate (BRANDplate, Brand GMBH, 781608). Each plate contains reference samples of fibril forming buffer and ThT with no protein. The plates are sealed using a transparent film (X-Clear Advanced Polyolefin StarSeal, Star Lab, E2796-9795) and placed into a BMG labtech platereader (Omega) with filters for 440nm and 480nm (BMG 440BP10 and BMG 480BP10 respectively). The samples were mixed with double orbital shaking and the fluorescence measured every 30 minutes by excitation at 440nm and monitoring the emission of the sample at 480nm.

2.22 SDD-AGE of *In vitro* Derived Polymers

2.22.1 Reagents

Table 2.36   Composition of SDD AGE Buffers

3 x SDD AGE Sample Buffer	Buffer G	Laemmli Buffer
180mM Tris-HCl	20mM Tris Base	20mM Tris-base
15% v/v glycerol	200mM glycine	200mM glycine
6% w/v SDS		0.1% w/v SDS
0.15% w/v bromophenol blue		
pH 6.8		

Table 2.37   Composition of SDS-Agarose Gel

Material
1.5% w/v agarose
0.1% w/v SDS
All dissolved in Buffer G

2.22.2 Procedure

The agarose was suspended in Buffer G and then microwaved on full power for 20 second bursts, until the agarose is dissolved. A stirrer bar is added and the mixture is stirred, whilst 20% w/v SDS was added to give a final concentration of 0.1%. The gel was poured into an empty DNA gel holder, with a lane comb inserted. The gel was allowed to set and then orientated so that the gel is in line with the electrodes with the comb at the negative end (cathode) of the gel. The gel was covered with Laemmli buffer and incubated for 30 minutes.

The samples were prepared by treatment with 3 x SDD AGE sample buffer and then incubated for 5 minutes prior to loading into the gel lanes.

The gel was then connected to a powerpack and run at 125V until the bromophenol blue from the sample buffer had travelled 7cm.

The gel was removed and imaged using a flat bed scanner, either an Epson Perfection V33 or Epson Perfection 3200 photo in conjunction with Epson Scan software, if fluorescently labelled fibrils were used.

2.23 Protein Ligation through Amide formation
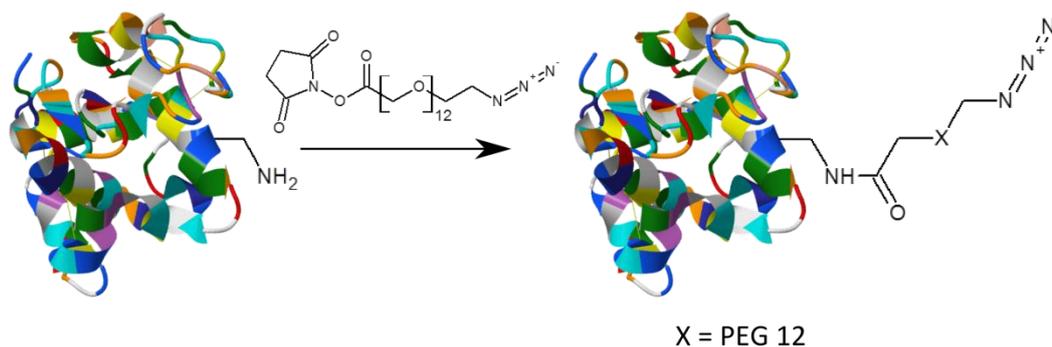
X = PEG 12

Hen egg lysozyme, >500µM, in pH 7.8 phosphate buffered saline (PBS) was cooled to 4°C and mixed by rolling. To this cooled solution 1.1 molar equivalents of a 10mM DMSO stock of NHS PEG$_{12}$ N$_{3}$ was added. The samples were mixed by rolling for 2 hours at 4°C and then quenched with 200µl of 1M tris(hydroxymethyl)aminomethane and rolled for 30 minutes.

The residual reagents were removed by buffer exchange into pH 8, 20mM sodium phosphate using a PD10 column according to the manufacturer's instructions.

2.24 Protein Ligation By Iodoacetamide displacement
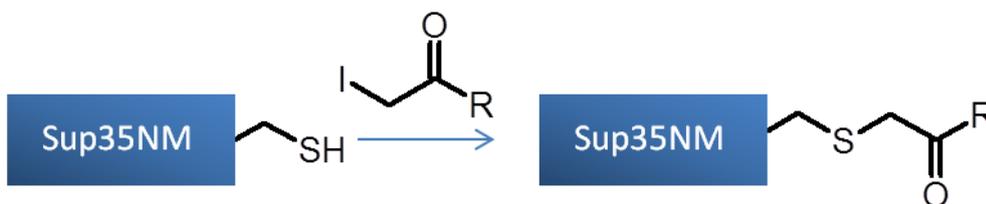
A denatured solution of a cysteine containing mutant (20mM phosphate, 50mM NaCl, 6M guanidine, pH 7.5) was defrosted from -80°C and then treated with a stock solution of an

Iodoacetamide reagent (5 molar equivalents in DMSO) and then incubated in a flask, protected from light by silver foil, for 2 hours.

The sample was then purified by buffer exchange into fibril forming buffer using a PD10 following manufacturer's instructions or stored at -80°C until purification/further reaction in a vial protected from light by silver foil.

2.25   Protein Ligation using CuAAC chemistry
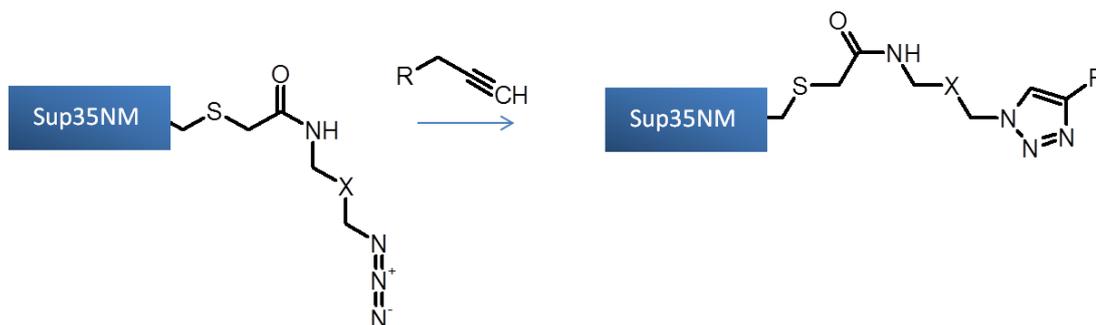
This method is applicable to both azide modified lysozyme and Sup35 although if the material is at pH2 the pH should be raised to at least pH5 for the duration of the experiment and then re-acidified at its conclusion. The presence of denaturant (guanidine) also seems to slow the rate of the conjugation.

2.25.1   General method for CuAAC reaction

The tris(3-hydroxypropyltriazolylmethyl)amine was added to a solution of the azide containing monomer and alkyne. The CuSO$_4$ was then added to this solution with the sodium ascorbate being added last. The reactions were then incubated with mixing, on a roller, for 2 hours at room temperature.

The excess reagents could be removed by PD10 purification according to the manufacturer's instructions or the reaction quenched by lowering the pH of the solution to pH2.

Table 2.38   General CuAAC Reaction Composition

Material	Molar equivalent
Azide containing monomer	1
Alkyne	5
CuSO$_4$.5H$_2$O	5
Tris(3-hydroxypropyltriazolylmethyl)amine	20
Sodium ascorbate	100

2.26   Protein Ligation Using SPAAC Chemistry
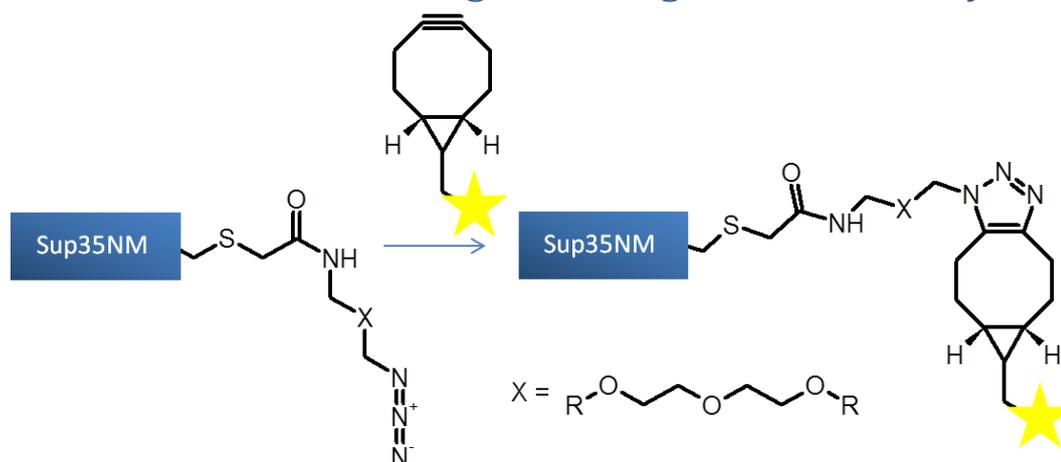
2.26.1   General method for SPAAC reaction

The azide containing monomer was treated with a stock solution of the BCN labelled

material and incubated at room temperature for 4 hours.

Excess reagents could then be removed by buffer exchange using a PD10 column according

to manufacturer's instructions.

Table 2.39   General SPAAC Reaction Composition

Material	Molar Equivalents
Azide containing monomer	1
Bicyclononyl (BCN)	10

2.27   Synthesis of FITC-Alkyne

A FITC alkyne conjugate was synthesised following the method of Xu et al (Xu and Ye,

2011).

Mol. Form.   = $C_{21}H_{11}NO_5S$
Form. Weight = 389.38074

Mol. Form.   = C_3H_5N
Form. Weight = 55.0785

Mol. Form.   = $C_{24}H_{16}N_2O_5S$
Form. Weight = 444.45924
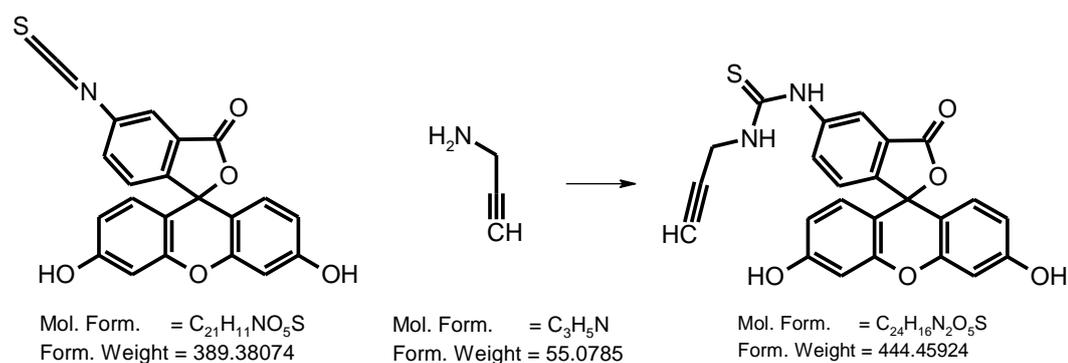
Table 2.40   Reagent table for the synthesis of FITC-Alkyne

Reagent	Molar equivalent	Density (g/ml)	Amount	mmoles
FITC	1	-	5mg	0.0128
Progargylamine	55	0.86	45µl	0.703
DMF	Solvent		55µl	

The FITC (5mg/0.01mmoles) was dissolved in DMF (55µl) and treated with propargylamine

(45µl, 0.703mmoles) and stirred overnight at room temperature in the dark.   To this

solution 950µl of 1:1 MeCN:H_2O were added to give a 12.8mM stock solution of the alkyne

labelled dye (assuming complete and successful reaction of the FITC).

2.28   Synthesis of N-(2-{2-[2-(2-azidoethoxy)ethoxy]ethoxy}ethyl)-2-iodoacetamide – Iodoacetamide PEG N$_3$

Mol. Weight = 218.25   Mol. Form.   = C$_6$H$_6$INO$_4$ Mol. Weight = 386.19
Mol. Form.   = C$_8$H$_{18}$N$_4$O$_3$ Mol. Weight = 283.02   Mol. Form.   = C$_{10}$H$_{19}$IN$_4$O$_4$
Density = 1.1g/L
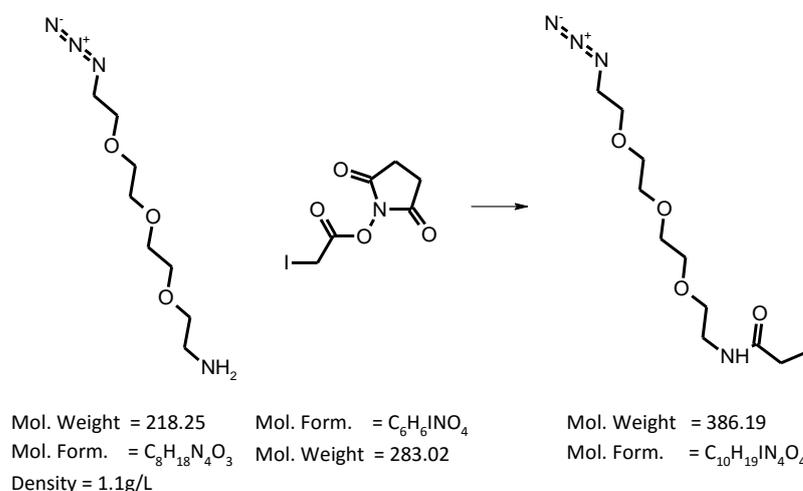
Table 2.41   Reagent table for the synthesis of Iodoacetamide PEG N$_3$

Reagent	Molar equivalent	Density (g/ml)	Amount	mmoles
11-Azido-3,6,9-trioxaundecan-1-amine	1.05	1.10	1.16mg	0.0053
SIA	1	-	1.94mg	0.00505
DCM	Solvent		200µl	

The   amine   (1.16mg/0.0053mmoles)   and   succinimidyl   iodoacetate   (SIA)
(1.94mg/0.00505mmoles) were dissolved separately in (50µl and 100µl respectively)
dichloromethane (DCM) and cooled to 4°C. The solutions were mixed and incubated on ice
for 5 minutes before rolling at room temperature for 2 hours.

The DCM was evaporated under a stream of N$_{2(G)}$ and then redissolved in DMSO to give a
stock solution (~20mM).

LC-MS – shows one peak   MS (ES): (MH$^+$) 387.0511

2.29   Synthesis of 2-iodo-N-(4,4,4-trifluorobutyl)acetamide – Iodoacetamide CF₃

Molecular Weight   = 127.1082296   Molecular Formula   = C₆H₆INO₄   Molecular Weight   = 295.0414396
Molecular Formula   = C₄H₈F₃N   Molecular Weight   = 283.02061   Molecular Formula = C₆H₉F₃INO
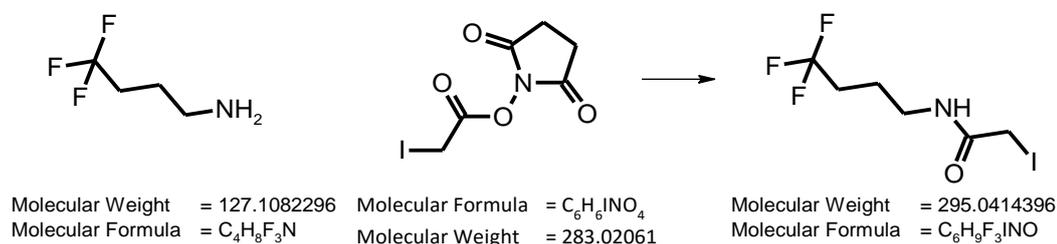
Table 2.42   Reagent table for the synthesis of Iodoacetamide CF3

Reagent	Molar equivalent	Density (g/ml)	Amount	mmoles
4,4,4-Trifluorobutylamine	1.05	0.93	1.18mg	0.0092
SIA	1	-	3.4mg	0.0088
DCM	Solvent		200µl	

The amine (1.18mg/0.0092mmoles) and succinimidyl iodoacetate (SIA) (3.4mg/0.0088mmoles) were dissolved separately in (2 x 100µl) DCM and cooled to 4°C. The solutions were mixed and incubated on ice for 5 minutes before rolling at room temperature for 2 hours.

The DCM was evaporated under a stream of N₂(g) and then redissolved in DMSO to give a stock solution (~20mM).

LC-MS shows one peak MS (ES): (MH⁺) 295.9665

2.30   Synthesis of 3,3,3-trifluoro-N-(prop-2-yn-1-yl)propanamide - CF₃-Alkyne

Molecular Weight   = 146.4955896   Molecular Weight = 55.0785   Molecular Weight   = 165.1131496
Molecular Formula   = C₃H₂ClF₃O   Molecular Formula = C₃H₅N   Molecular Formula = C₆H₆F₃NO
Density = 1.00   Density = 0.86
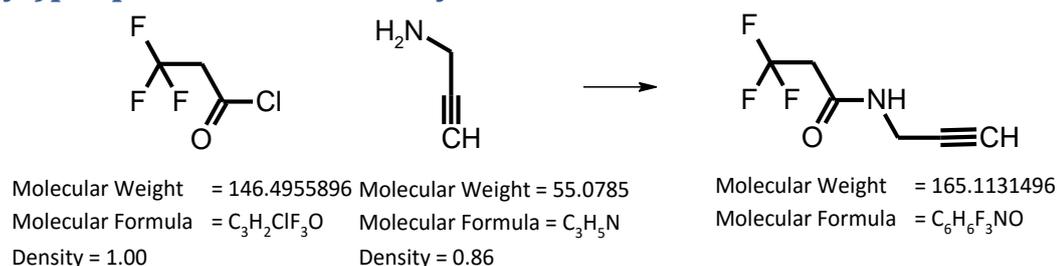
Table 2.43   **Reagent table for the synthesis of CF$_3$-Alkyne**

Reagent	Molar equivalent	Density (g/ml)	Amount	mmoles
3,3,3-Trifluoropropionyl Chloride	1	1	146µl	0.001
Propargylamine	2.5	0.86	160µl	0.0025
Diethyl ether	Solvent		2ml	

Propargylamine (160µl/0.0025mmoles) was dissolved in diethyl ether (2ml) and cooled to 4$^\text{o}$C. The acid chloride (146µl/0.001mmole) was then added to this solution in three portions and the reaction incubated at 4$^\text{o}$C for 10 minutes before mixing at room temperature for 20 minutes.

The reaction was diluted with diethyl ether (3ml) and then washed with 700µl 1M HCl solution (x2) and then a saturated NaCl solution (x2). The diethyl ether was then concentrated under a stream of $N_{2(g)}$ to give a gummy solid. The material was then dissolved in DMSO to give a 100mM stock solution.

LC-MS shows one major component with some small impurities MS (ES): (MNa$^+$)188.0321

2.31   Synthesis of N-[(1R,8S,9s)-Bicyclo[6.1.0]non-4-yn-9-ylmethyloxycarbonyl]-1,8-diamino-3,6-dioxaoctane FITC conjugate - BCN-FITC

Mol. Weight = 324.415
Mol. Form. = $C_{17}H_{28}N_2O_4$

Mol. Form. = $C_{21}H_{11}NO_5S$
Form. Weight = 389.38

Mol. Weight = 713.796
Mol. Form. = $C_{38}H_{39}N_3O_9S$
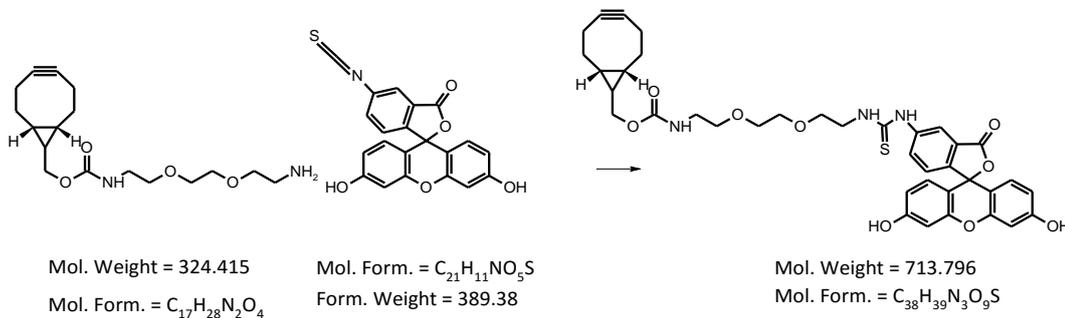
Table 2.44 **Reagent table for the synthesis of BCN-FITC**

Reagent	Molar equivalent	Density (g/ml)	Amount	mmoles
N-[(1R,8S,9S)-Bicyclo[6.1.0]non-4-yn-9-ylmethyloxycarbonyl]-1,8-diamino-3,6-dioxaoctane	1	-	2.1mg	0.0065
FITC	1.05	-	2.66mg	0.00683
DCM	Solvent		100µl	
DMF	Solvent		50µl	

The BCN amine (2.1mg/0.0065mmoles) was dissolved in DCM (100µl) and the FITC (2.66mg/ 0.00683mmoles) was dissolved in DMF (50µl). The two solutions were cooled to 4°C, combined and then incubated at 4°C for 5 minutes before mixing, by rolling, at room temperature for 2 hours.

The DCM was removed by evaporation under a stream of $N_{2(g)}$ and the resulting DMF solution was diluted with DMSO to give an ~60mM stock solution.

LC-MS shows one major component with some small impurities MS (ES): (MH^{+}) 714.2371

2.32 Synthesis of N-[(1R,8S,9s)-Bicyclo[6.1.0]non-4-yn-9-ylmethyloxycarbonyl]-N′-(1,1,1)trifluoropropionyl-1,8-diamino-3,6-dioxaoctane – BCN-CF$_3$

Mol. Weight = 324.415   Mol. Weight = 46.496   Mol. Weight = 434.45
Mol. Form.   = $C_{17}H_{28}N_2O_4$ Mol. Form.   = $C_3H_2ClF_3O$ Mol. Form.   = $C_{20}H_{29}F_3N_2O_5$

Table 2.45   **Reagent table for the synthesis of BCN-CF$_3$**

Reagent	Molar equivalent	Density (g/ml)	Amount	mmoles
N-[(1R,8S,9S)-Bicyclo[6.1.0]non-4-yn-9-ylmethyloxycarbonyl]-1,8-diamino-3,6-dioxaoactane	1	-	2mg	0.00615
3,3,3-Trifluoropropionyl Chloride	1.6	1.00	1.44mg	0.00984
Pyridine	40	0.982	19.8µl	0.246
DCM	Solvent		190µl	

The BCN amine (2mg/0.00615mmoles) was dissolved in 100µl DCM, cooled to 4°C and then treated with a 50µl DCM solution containing pyridine (19.8µl/0.245mmoles). The acid chloride (1.44mg/0.00984mmoles) was dissolved in 40µl DCM, cooled to 4°C and then added to the amine solution slowly. The reaction was then incubated at 4°C for 5 minutes and then mixed at room temperature for 2 hours.

The material was then diluted with 300µl DCM and washed with 300µl pH2 HCl three times. The organics were then evaporated under a stream of N$_{2(g)}$ and the resulting material was dissolved in DMSO to give a concentration of about 20mM.

LC-MS very small peak which looks like two peaks MS (ES): (MH$^+$) 435

2.33   Electron Microscopy

2.33.1   Surface Deposition

Fibril samples were diluted appropriately, typically 1:4 with double distilled H$_2$O, 20µl was then placed onto a 400 mesh carbon/formavar coated copper surface and incubated for 30 seconds. The excess solution from the surface was removed by a pipette attached to a water vacuum pump. The surface was then incubated with 20µl of 2% w/v uranyl acetate solution for 1 minute. The excess uranyl acetate was then removed by vacuum and the

sample allowed to air dry for 10 minutes before storing in a grid box until use. The vacuum manifold was washed with copious amounts of H_2O to remove uranium residues.

2.33.2   Imaging

Surfaces were imaged using a Jeol JEM 1230 transmission electron microscope, at an accelerating voltage of 80kV, with Gaitan Multiscan camera Model 791 and Digital Micrograph software. Images were captured at multiple magnifications with and without scale bars.

2.34   Atomic Force Microscopy (AFM)

2.34.1   Deposition

A mica disc (Agar scientific, F7013) was attached to a steel surface (Agar scientific,) using double sided tape. The top layer of mica was removed using tape.

20µl of a fibril sample diluted appropriately in fibril forming buffer was incubated on the cleaved mica surface for an appropriate length of time (Table 2.46). The excess sample was then washed away with an appropriate volume of mQH_2O (Table 2.46). The surface was then dried under a stream of $N_{2(g)}$ (Standard) or by removing residual water with a filter paper placed on the extreme edge of the mica surface held vertically and then the drying was finished with a very gentle stream of $N_{2(g)}$ (Gentle).

Table 2.46   The different deposition conditions for AFM imaging

Protein	Deposition concentration	Incubation time	Wash Volume	$N_{2(g)}$ dry conditions
Sup35 Full length	10µM	5 minutes	1000µl	Standard
Sup35 NM	10µM	5 minutes	1000µl	Standard
Sup35NM – LCys	10µM	5 minutes	1000µl	Standard
Insulin B chain	3µM	5 minutes	1000µl	Standard
α-Synuclein	0.48µM	5 minutes	1000µl	Gentle
β-Lactoglobulin	0.6µM	5 minutes	1000µl	Standard
Lysozyme	3µM	5 minutes	1000µl	Standard

2.34.2   Imaging

Prepared samples were imaged using a Bruker Multimode AFM with a Nanoscope V controller. Samples were imaged using a ScanAsyst probe (A silicon nitride triangular tip with tip height = 2.5-8µm, nominal tip radius = 2nm, nominal spring constant 0.4N/m and nominal resonant frequency 70kHz) and the ScanAsyst experiment in the Nanoscope 8.15 software.

Typically images were captured at a resolution of 9.77nm per pixels scanned (typically 10µm x 10µm at 1024 x 1024 pixels or 20µm x 20µm at 2048 x 2048 pixels). All images are processed using the Nanoscope Analysis software. The image baseline was flattened using 3rd order correction, the streaks removed from the fibrils by lowering the z threshold % and the data then saved as a PNG file and a processed image file with a file extension of .001 for recognition by the tracing software.

2.34.3   MatLab Analysis

The processed image file, with a file extension of .001, was opened in the fibre tracing MatLab application and Auto Trace Polymers was selected. The apparent width, search Phi and z cutoff values were inputted into the software and the tracing was initiated (Table 2.47). The completed traces were saved. The optimum values of each parameter were determined by trial and error modification of each value until the most accurate trace was achieved.

Table 2.47   Table showing the values used for the tracing of the fibrils samples

Protein	Apparent Width (pixels)	Search Phi (°)	Z cutoff (nm)
Lysozyme	3	50	1
α-Synuclein	4	40	1.5
β-Lactoglobulin	2	60	1

The resulting image file was saved and the fibril data extracted using the PolProp app (Beal, Appendix 1).

The persistence lengths were calculated from the traced polymers using the CalcLp app (Xue, unpublished). The method of bias correction was calculated by taking the combined fragmentation data and inputting into the LcBiasCorr app (Xue, unpublished). This correction method was then used in conjunction with the same data using the PlotLcDistr app (Xue, unpublished) to give the corrected data.

2.35   Liquid Chromatography - Mass Spectrometry (LC-MS) – Electrospray

LC-MS of samples was achieved using an Agilent series 1100 HPLC system controlled by Bruker Compass Hystar 3.2 software. The instrument was primed with appropriate buffers prior to the appropriate column being attached. The samples were placed in a HPLC vial (HPLS, Chromacol, V944/950) which where necessary (particularly protein samples) contained an insert (Verex Insert, 5mm diameter, 175µl, conical bottom, with spring bottom, Phenomenex, ARO-4521-13) allowing for the use of low volumes.

2.35.1   Protein

Table 2.48   Protein electrospray LC-MS conditions

Column	Phenomenex, 00B-4167-130, Jupiter, 5µm, C4, 300Å, 50 x 2.00mm
Buffer A	95.95% v/v H_2O, 0.05% v/v trifluoroacetic acid
Buffer B	4.955% v/v H_2O, 95% v/v acetonitrile, 0.045% v/v trifluoroacetic acid

The column was equilibrated with Buffer A at a flow rate of 0.2ml/min and a column temperature of $30^{o}C$. A sample table was prepared using the Hystar software. The generic protein desalt method (desalt.m) was used as the chromatography method, the generic protein mass spec method (peptide_lc_nomsms.m) was used for the mass spectrum capture and the generic protein processing method (protein.m) was used for processing.

Sample analysis was carried out using the Bruker Compass DataAnalysis 4.1 software. A datafile was opened, and an average of the MS data under a peak was selected giving a

charge envelope which was smoothed using a Savitsky-Golay function.   The resulting

charge envelope was then deconvoluted to give the deconvoluted molecular weight.

2.35.2   Small molecule

Table 2.49   Protein electrospray LC-MS conditions

Column	Phenomenex, 00B-4387-130, Synergi, 2.5µm, Hydro-PR 100Å, 50 x 2.00mm 2µm
Buffer A	95.95% v/v H_2O, 0.05% w/v ammonium formate
Buffer B	4.955% v/v H_2O, 95% v/v methanol, 0.045% w/v ammonium formate

The column was equilibrated with Buffer A at a flow rate of 0.2ml/min and a column

temperature of $30^{o}C$. A sample table was prepared using the Hystar software. The generic

small molecule method (Small molecule.m) was used as the chromatography method and

the generic small molecule mass spec method (Small_start.m) was used for the mass

spectrum capture.

Sample analysis was carried out using the Bruker Compass DataAnalysis 4.1 software. A

datafile was opened, and an average of the MS data under a peak was selected giving and

smoothed using a Savitsky-Golay function. No charge deconvolution is required for small

molecules.

2.36   Nuclear Magnetic Resonance spectroscopy (NMR)

2.36.1   Sample preparation

Samples are diluted to an appropriate concentration with fibril forming buffer, volume of

500µl, D_2O (60µl) was also added to samples as well, dioxane (varied amount depending on

the sample concentration) and a fluorinated reference sample (either trifluoroethanol or

4,4,4-trifluorobut-2-en-1-ol) (varied amount depending on the sample concentration).

2.36.2 Acquisition conditions

All spectra were acquired at 25°C using a 4-channel, 14.1 T (600 MHz 1H) Bruker Avance III NMR spectrometer equipped with a 5mm QCI-F cryoprobe and Bruker TopSpin software by Michelle Rowe. The parameters and experiments used for each experiment are shown in Table 2.50.

Table 2.50   NMR Parameters for Spectra Acquisition

Sample	Pulse Seq.	Data-points	No. Of Scans	Spectral Width (ppm)	Offset (ppm)	Acquisition time (s)
Lysozyme Monomer (19F)	stegp1s1d	32768	512	80.5136	-95	0.3604480
Lysozyme Fibrils with TFBenol (19F)	stegp1s1d	8192	4096	80.5136	-95	0.0901120
Lysozyme Fibrils with TFE (19F)	stegp1s1d	8192	4096	80.5136	-95	0.0901120
Lysozyme Monomer (1H)	stebpgp1s191d	32768	128	7211.539	4.700	2.2719147
Lysozyme Fibrils with TFBenol (1H)	stebpgp1s191d	32768	128	7211.539	4.700	2.2719147
Unlabelled Fibrils (1H)	stebpgp1s191d	32768	128	7211.539	4.700	2.2719147
Sup35NM 19F Labelled (1H)	stebpgp1s191d	32768	4096	7211.539	4.700	2.2719147

2.36.3 Data Analysis

The captured spectra were processed using an exponential window function with varying levels of line broadening, phased using the automated 0th order correction algorithm (apk0), then manually baseline corrected using the Bruker Topspin 3.2 software and exported as a textfile. The data was then imported into MatLab and concatenated to give a data set containing the NMR spectra of the sample with gradient strength of between 5% and 95% maximum. A single peak is selected from the data and the area under the appropriate peak calculated. The decay of peak intensity with increasing gradient strength was fitted to the Stejskal-Tanner equation (Equation 5.2) using the MatLab curve fitting tool.

2.37   Confocal Microscopy of Fluorescent Fibrils

2.37.1   Sample Preparation

A mica surface was attached to a microscope slide using double-sided tape and then cleaved using electrical tape. 20μl of neat Sup35NM LCys N_3 BCN FITC was incubated for 5 minutes before washing with 1ml of syringe filtered mQH_2O and drying under a stream of $N_{2(g)}$.

2.37.2   Acquisition conditions

Samples were visualized, by Dr Dan Mulvihill, using an Olympus IX71 microscope with PlanApo 100x OTIRFM-SP 1.45 NA lens mounted on a PIFOC *z*-axis focus drive (Physik Instrumente, Karlsruhe, Germany), and illuminated using LED light sources (Cairn Research Ltd, Faversham, UK) with appropriate filters (Chroma, Bellows Falls, VT). Samples were visualized using a QuantEM (Photometrics) EMCCD camera, and the system was controlled with Metamorph software (Molecular Devices). Each 3D-maximum projection of volume data was calculated from 31 *z*-plane images, each 0.2 μm apart, using Metamorph or Autoquant X software.

3

Assembly of Amyloid Fibrils and the Quantitative Analysis of their Morphology

3.1   Introduction

The rise of molecular biology since the early 1930s, the subsequent identification of the structure of DNA by Franklin, Watson and Crick in the 1950s and the unlocking of the genetic code by Crick *et al* in 1961 (Crick et al., 1961) are key landmarks in the development of recombinant DNA technologies (rDNA) as a tool for both the understanding of biological systems and the curing of disease (Kellenberger, 2004). rDNA technology has evolved from the seminal work of Jackson *et al* (Jackson et al., 1972) which describes the strategy and techniques required to incorporate exogenous DNA into cell systems (Struhl, 2008).
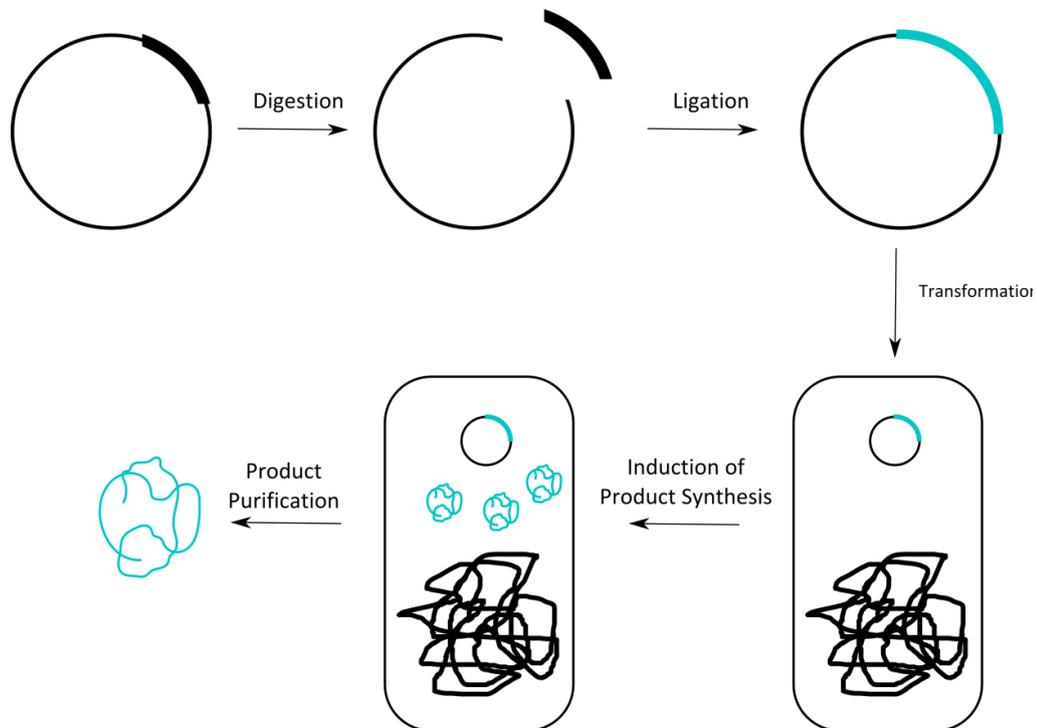
Figure 3.1   The Process of rDNA Material Production
Bacterial plasmid DNA is digested by the action of Restriction Enzymes (Digestion) which is followed by recircularisation of the plasmid DNA by ligation of a DNA fragment encoding for the required product into the plasmid (Ligation). Insertion of the plasmid DNA into a competent bacterial strain (Transformation) is followed by induction of product synthesis (Induction of Product Purification) and then purification of the generated product (Product Purification).

The work of Edward Jenner, 1796, on the use of bovine pox to vaccinate against small pox showed the power of biological materials for the treatment of disease (Riedel, 2005) but it

was not until 1976 with the funding of Genentech and their subsequent use of rDNA to produce somatostatin, 1977, in *E. coli* that the modern biotechnology era started (Itakura et al., 1977). rDNA has become integral to the study of biochemistry, interactions of proteins (Rosano and Ceccarelli, 2014) and the production of biotherapeutics such as the highly specific monoclonal antibodies and has led to a dramatic rise in the market share of biologics compared to their small molecule counterparts (Goodman, 2009).

For all rDNA technologies, the purification of the produced material from the cellular milieu is paramount to their use, whether as therapeutics or for biochemical and biophysical studies (Fig. 3.1). Purification of molecules, small or large, often relies on using their physical/chemical properties to differentially partition between two mediums i.e. size exclusion chromatography and ion exchange chromatography. One of the major benefits of the rDNA technology is the ability to genetically encode modifications to proteins for multiple purposes including the addition of tags which allow for affinity capture of the product onto a resin i.e. His, GST and FLAG tags (Young et al., 2012).

As previously discussed the aggregation of proteins into amyloid is important biologically. Industrially, the aggregation of synthetic biomolecules during their synthesis and storage causes a significant loss in yield associated with the drug (Wang et al., 2009b), and more worryingly, the development of localised amyloidosis can be associated with their administration, for example Insulin (Yumlu et al., 2009). Considering the hypothesis that all proteins under the appropriate conditions can convert into the amyloid state (Chiti et al., 1999; Dobson, 1999), these processes are not surprising, but they highlight another reason why understanding and prevention of amyloid is important.

The conversion of protein into amyloid fibrils follows a nucleated growth mechanism which requires a change from native monomer conformation to one which contains a larger proportion of beta sheet (Serpell, 2000). The optimal conditions required for this

conformational change and subsequent polymerisation can vary considerably, for example different pH, and temperature change. The generation of high concentrations of pure protein, as expected with rDNA, also increases the probability that two misfolded proteins will come into contact and therefore the potential for nucleation also increases.

rDNA technology is utilised to generate some of the amyloid forming proteins investigated in this thesis, including α-Synuclein, Sup35, Lysozyme, β-Lactoglobulin and Insulin B chain.

3.1.1 α-Synuclein

α-Synuclein is a 14.6kDa intrinsically disordered protein (IDC) (Fig. 3.2) whose aggregation into amyloid fibrils has been associated with the onset of the neurological disorder Parkinson's disease. The N-terminal domain of α-synuclein has been shown to interact with other proteins like chaperones (Rekas et al., 2004; Souza et al., 2000) and the C-terminal domain provides solubility but also has a role in affecting the rate of amyloid formation (Uversky et al., 2001). The early stages of α-synuclein fibril formation is promoted by the presence of dopamine (Cappai et al., 2005b) and fibrils formed have been shown to be a major component of Lewy bodies, found in dopaminergic neurons at disease onset (Li et al., 2001). The disease is commonly found in older patients but the faster aggregating A30P and A53T point mutants of the protein have been found to cause the early onset forms of the disease (Conway et al., 1998; Li et al., 2001)(Fig. 1.3 and Fig. 3.2).

```
MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTK
EQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDP
DNEAYEMPSEEGYQDYEPEA
```

Figure 3.2   Amino Acid sequence of α-Synuclein
The primary amino acid sequence of α-Synuclein with the position of the A30P (red) and the A53T (green) mutants marked.

3.1.2 Sup35

The aggregation of the yeast translation termination factor eRF3, alternatively known as Sup35, is central to the [*PSI*⁺] phenotype discovered by Brian Cox in 1965 (Cox, 1965), and is

an example of a prion (Tuite and Cox, 2006). Polymerisation of Sup35 into amyloid aggregates and their subsequent chaperone mediated fragmentation into propagons is key to their transmission from one cell to another (1.10.1) (Byrne et al., 2009; Cox et al., 2003). The Sup35 amino acid sequence can be split into 3 domains, each with different functions, and separated by methionine residues (Fig. 3.3). The N terminal domain is the amyloid forming region, the M domain contains polar residues thought to be important for solubility and the C terminal domain is responsible for the ATP dependent translation termination activity of the protein (Glover et al., 1997). As the [*PSI*⁺] phenotype is a facet of the N and M domains, and not affected by the C domain, a lot of the research into the initiation and propagation of the phenotype have been carried out on the truncated form of the protein containing only the N and the M domains, Sup35NM. Recent work has raised the importance of the C domain showing that mutations within its sequence can affect the phenotype observed in yeast showing the importance of considering the full length protein, Sup35NMC (Krzewska and Melki, 2006; Krzewska et al., 2007). The threonine residue at position 341 in the C domain of the protein is remote from the fibril forming region and therefore should not affect the [*PSI*⁺] phenotype observed but mutation to aspartate or alanine causes lethality and [*PSI*⁺] phenotype weakening respectively (Kabani et al., 2011).

MSDSNQGNNQQNYQQYSQNGNQQQGNNRYQGYQAYNAQAQPAGGYYQNYQGYSGYQQGGY
QQYNPDAGYQQQYNPQGGYQQYNPQGGYQQQFNPQGGRGNYKNFNYNNNLQGYQAGFQPQ
SQGMSLNDFQKQQKQAAPKPKKTLKLVSSSGIKLANATKKVGTKPAESDKKEEEKSAETK
EPTKEPTKVEEPVKKEEKPVQTEEKTEEKSELPKVEDLKISESTHNTNNANVTSADALIK
EQEEEVDDEVVNDMFGGKDHVSLIFMGHVDAGKSTMGGNLLYLTGSVDKRTIEKYEREAK
DAGRQGWYLSWVMDTNKEERNDGKTIEVGKAYFETEKRRYTILDAPGHKMYVSEMIGGAS
QADVGVLVISARKGEYETGFERGGQTREHALLAKTQGVNKMVVVVNKMDDPTVNWSKERY
DQCVSNVSNFLRAIGYNIKTDVVFMPVSGYSGANLKDHVDPKECPWYTGPTLLEYLDTMN
HVDRHINAPFMLPIAAKMKDLGTIVEGKIESGHIKKGQSTLLMPNKTAVEIQNIYNETEN
EVDMYAMCGEQVKLRIKGVEEEDISPGFVLTSPKNPIKSVTKFVAQIAIVELKSIIAAGFS
CVMHVHTAIEEVHIVKLLHKLEKGTNRKSKKPPAFAKKGMKVIAVLETEA

Figure 3.3   Amino Acid Sequence of Sup35NMC/NM
The primary amino acid sequence of Sup35NMC with the sequence N domain coloured purple, M domain coloured red and the C domain in black. The methionine residues contained in the protein are coloured green and the cysteine residues blue.

3.1.3 Insulin

Insulin is a hormone common to many organisms that regulates blood sugar levels (Fig. 3.4, Left). The human protein consists of two chains, A and B, covalently joined through three disulfide bonds and has a molecule weight of 5808Da. Insulin has been observed to form amyloid aggregates as the heterodimer but also each of the constituent peptide chains has been shown to convert into amyloid when isolated (Devlin et al., 2006). The failure of a body to regulate blood sugar levels by producing enough or any insulin are the hallmarks of type 2 or type 1 diabetes respectively. Diabetes is highly prevalent in today's society and one of the methods of treatment is to subcutaneously inject insulin into the body (http://www.diabetes.co.uk/diabetes-causes.html). Before rDNA technology therapeutic insulin was derived from animal sources, which do not have the same amino acid sequence as human insulin (Fig. 3.4, Right), this has led to allergic responses as well as requiring a large amount of animal material. In 1978 a collaboration between City of Hope Medical Centre, California, and Genentech led to a bacterial strain capable of producing the two chains of human insulin which could then be converted into the therapeutic agent (Genentech, 1978).


```
Human Insulin A      GIVEQCCTSICSLYQLENYCN 21
Porcine Insulin A    GIVEQCCTSICSLYQLENYCN 21
Bovine Insulin A     GIVEQCCASVCSLYQLENYCN 21
                     *******:*:***********

Human Insulin B      FVNQHLCGSHLVEALYLVCGERGFFYTPKT 30
Porcine Insulin B    FVNQHLCGSHLVEALYLVCGERGFFYTPKA 30
Bovine Insulin B     FVNQHLCGSHLVEALYLVCGERGFFYTPKA 30
                     *****************************:
```
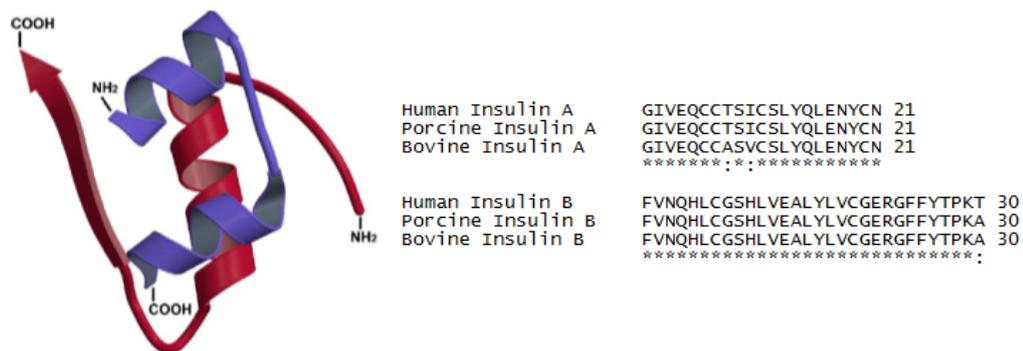
Figure 3.4   3 Dimensional Representation of the Insulin Hormone and Sequence Homolgy Between Species
Left - The structure of insulin with the B-Chain in blue and the A-Chain in red. (Devlin et al., 2006). Right – Comparison of the amino acid sequences of the two chains of Insulin derived from three species; Human, Porcine and Bovine.

The individual chains of the insulin hormone can be isolated and prevented from forming homodimers by oxidation of the sulphur atoms in the cysteine residues to sulfonic acids as observed in the commercial form of bovine insulin B-chain (Fig. 3.5).

```
FVNQHLC(SO₃H)GSHLVEALYLVC(S0₃H)GERGFFYTPKA
```

Figure 3.5   Amino Acid Sequence of Sigma Bovine Insulin B-Chain with Cysteine residues oxidised to sulphonic acids
The primary amino acid sequence of Insulin B-Chain separated from the A-Chain and disulfide bond formation limited by oxidation of the cysteine sulfurs to sulfonic acids (highlighted in green), as supplied by Sigma Aldrich (I6383)

3.1.4   Lysozyme

Lysozyme (Fig. 3.6, Left) was first observed by Alexander Fleming in 1922 as a material capable of lysing bacteria, an ability stemming from its primary function as an enzyme which catalyses the hydrolysis of β-1,4 glycosidic linkages between N-acetylmuramic acid and N-acetylglucosamine, which are found in bacterial cell walls. The determination of lysozyme's 3D structure by X-ray diffraction was the second protein structure solved by this method and the first for an enzyme (Blake et al., 1965; Johnson, 1998).

KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKF
ESNFNTQATNRNTDGSTDYGILQINSRWWCNDGR
TPGSRNLCNIPCSALLSSDITASVNCAKKIVSDGNGM
NAWVAWRNRCKGTDVQAWIRGCRL
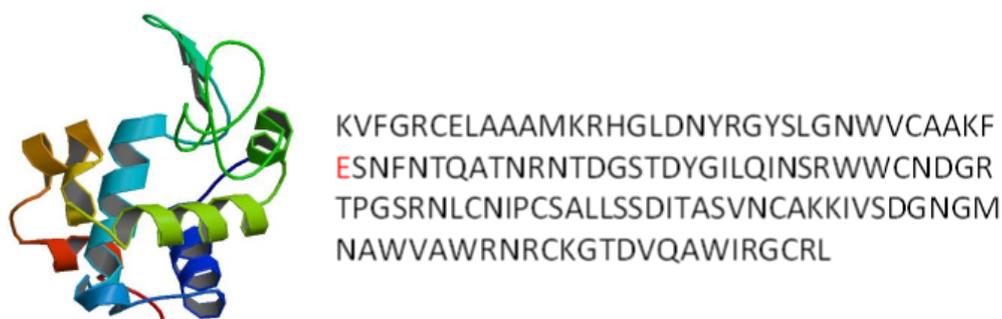
Figure 3.6   3D Structure and Amino Acid Sequence of Hen Egg White Lysozyme
Left – 3D structure of Hen Egg White Lysozyme, Right – Amino Acid sequence of Hen Egg White Lysozyme with the catalytic Glutamic acid residue labelled in red. Structure and sequence from RCSB PDB record 4RLM.

In its native state, the 14.3kDa protein has two domains containing alpha helices and beta-sheet motifs (Merlini and Bellotti, 2005) of which the glutamic acid in position 35 is the residue responsible for its catalytic activity (Strynadka and James, 1991) (Fig. 3.6, Right).

The unfolding of Lysozyme can cause the generation of amyloid aggregates and this is associated with the systemic amyloidosis (Mossuto et al., 2010).

3.1.5 β-Lactoglobulin

β-Lactoglobulin (Fig. 3.7) is a dimeric protein and a major component of cow's milk whey but has no known function apart from a propensity to bind retinol (Hamada and Dobson, 2002) and is the main allergen associated with dairy intolerance (Sélo et al., 1999). Under acidic conditions, the dimeric complex collapses into monomeric protein, and when monomers are denatured they convert into amyloid fibrils (Hamada and Dobson, 2002).

LIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQ
SAPLRVYVEELKPTPEGDLEILLQKWENGECAQKKII
AEKTKIPAVFKIDALNENKVLVLDTDYKKYLLFCMEN
SAEPEQSLACQCLVRTPEVDDEALEKFDKALKALPM
HIRLSFNPTQLEEQCHI

Figure 3.7   3D Structure and amino acid sequence of β-Lactoglobulin
Left- 3D structure of b-Lactoglobulin Right – Amino acid sequence of β-Lactoglobulin.
Structure and sequence both from RCSB PDB record 4IB9.

3.1.6 Chapter Aims

This part aims to develop expression and purification protocols for the generation of these model proteins thus allowing for the investigation of the conditions required for conversion of the protein monomers into fibrillar aggregates suitable for atomic force microscopy analysis. The generation of high quality samples reproducibly will allow investigation of some of the physical characteristics of amyloid fibrils such as height, contour length and persistence length. These fibril stocks will then be used to examine the resistance of different fibril types to mechanical stress, see Chapter 4.

3.2 Results

The *in vitro* assembly of proteins into fibrillar structures requires the generation of protein monomers prior to assembly. Due to the different conditions required to convert the monomeric protein into fibrils, the availability and purification of these materials also varies considerably. Insulin B chain, lysozyme and β-lactoglobulin all require unfolding conditions to convert them to their amyloid form and the protein monomers are supplied in their monomeric form commercially. The formation of α-synuclein and the Sup35 proteins all convert to their fibrillar form at room temperature and neutral pH and require synthesis and purification prior to assembly. Due to α-synuclein and Sup35's propensity to aggregate under ambient conditions their purification strategies need to be tailored to their individual properties to enable isolation of pure monomers.

3.2.1 Protein Purification

3.2.1.1 α-Synuclein

Due to the significance of α-synuclein to the pathology of Parkinson's disease the study of this protein and its fibrillar form has been undertaken on numerous occasions. Our studies of the breakage of α-synuclein amyloid fibrils require the generation of reproducible and homogeneous samples of fibrils from monomers prepared following published methods (Cappai et al., 2005b; Smith et al., 2008; Wu et al., 2009). Briefly, transformation of the chemically competent *E.coli* expression cell line BL21[DE3] pLysS with a pET23a vector containing the gene for wildtype α-synuclein with T7 promoter under the control of the *lac* operon, obtained from (Smith et al., 2008), gave an expression system yielding large amounts of the full length wild type protein. A modified purification protocol where acid precipitation of the cellular residue was followed by a stepped ammonium sulphate precipitation, from 30% to 50% gave very pure material (Fig. 3.8). Salt removal by buffer exchange allowed final purification by strong anion exchange chromatography to remove higher molecular weight materials evident on the interface of the stacking and resolving

components of the SDS PAGE (Fig. 3.8), any large aggregates could also be removed by filtering through a 0.2µm membrane. The purity and identity of the protein was confirmed by SDS PAGE and LC-MS analysis which after deconvolution gave the desired molecular weight of 14459.52Da.

Despite α-synuclein's ability to convert to the amyloid state at close to ambient temperature the expression and purification of the protein is possible under native conditions giving very high yields of pure protein, 28mg/L.
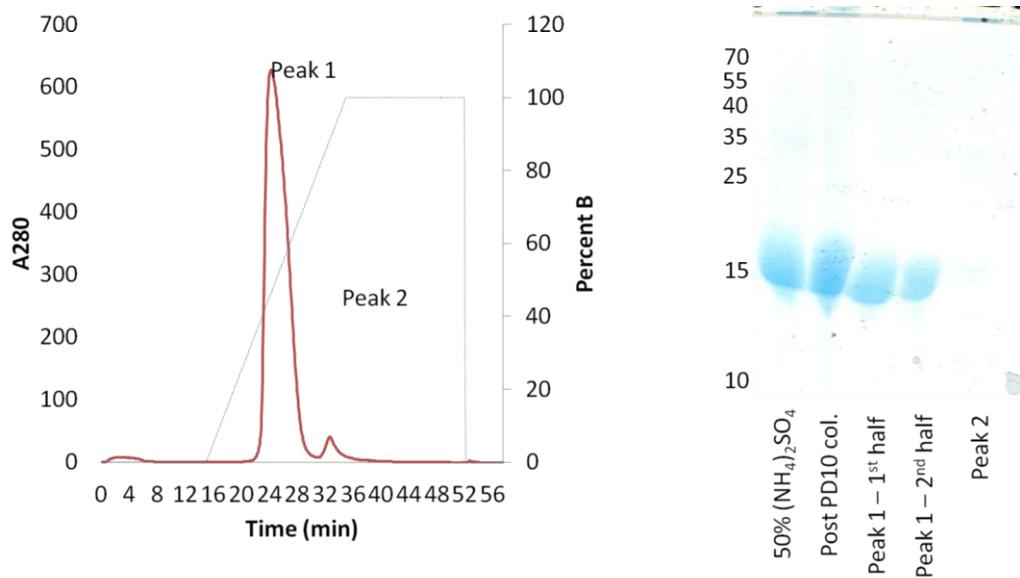
Figure 3.8   α-Synuclein Purification
Left – Chromatogram showing the elution profile of α-synuclein eluting from the HiTrap Q FF anion exchange column under a sodium chloride gradient (%B). Right – SDS PAGE, 15% Tris-Glycine, analysis of the purity of the α-synuclein at different stages of the purification. Ladder not shown but position of markers indicated by number indicative of mass in kDa of marker.

The high yield and ease of purification associated with α-synuclein production is not a facet shared with the Sup35 derived proteins.

3.2.1.2   Sup35NM

Study of the protein Sup35NMC, whose conversion to an amyloid state is at the heart of the [*PSI*[+]] phenotype observed in yeast, is commonly performed on constructs containing only the N terminal and middle domains of the protein, Sup35NM. The removal of the C

terminal, translation termination domain simplifies protein production but also leads to *in vitro* assembled fibrils which can cause the [*PSI*+] phenotype when introduced to yeast cells. It is therefore a good surrogate for the full length protein (King and Diaz-Avalos, 2004; Tanaka et al., 2004).

In comparison to α-synuclein, the majority of Sup35NM expressed in the cell is found in insoluble inclusion bodies, which are postulated to exist in its amyloid form (Ventura 2012). It is therefore necessary to purify the protein under denaturing conditions (Glover et al., 1997). The purification of amyloidogenic proteins, particularly those whch polymerise at pH7, in denaturing conditions offers the advantage of inhibiting the assembly of monomers into fibrils.
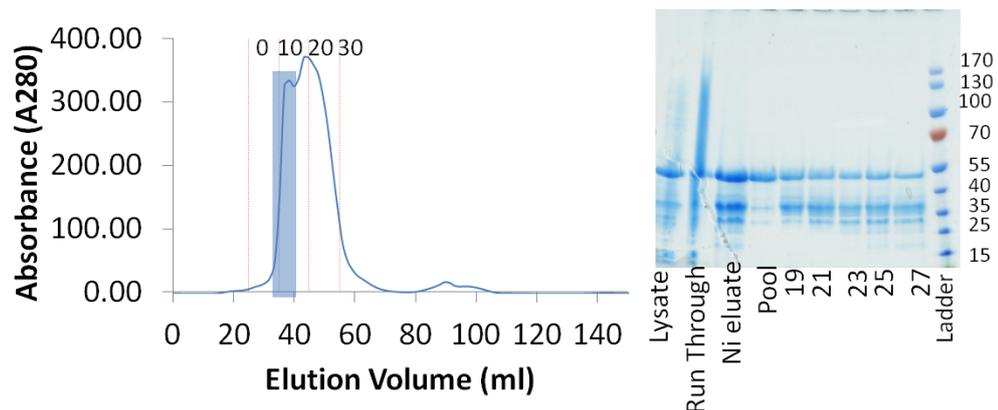
Figure 3.9   **Purification profile of Sup35NM**

Left – Äkta purification profile of Sup35NM lysed under denaturing conditions and then subjected to affinity purification before further purification using size exclusion chromatography at a 0.3ml/min. The start of fraction collection and every 10th fraction was marked with a red dotted line. A broad peak was eluted with the left hand side of the peak containing purest material (the fractions highlighted blue were pooled). Right – Invitrogen 4-12% Bis-Tris SDS PAGE run in MOPS buffer at 180V was used to analysis the fractions from the purification, which were diluted 1 in 3, to reduce interference of the guanidine on running, prior to loading onto the gel. Fractions 19-27 were the right hand side of the peak post pooling. Äkta trace and SDS PAGE supplied by Tracey Purton.

An N-terminally his-tagged DNA construct of Sup35NM was cloned, using a plasmid containing the full length Sup35 DNA sequence, obtained from (Kabani et al., 2011; Krzewska and Melki, 2006; Krzewska et al., 2007), and ligated into a pET15b vector.

Transformation of this plasmid into an *E.coli* expression strain, BL21 [DE3] pLysS produced a clone capable of Sup35NM expression. Protein purification was achieved using a modification to published protocols (Tanaka, 2010). Briefly, cell disruption by sonication in Lysis buffer containing 6M guanidine followed by centrifugation removed cell debris and the resulting supernatant was affinity purified (Ni-NTA with imidazole elution) followed by size exclusion chromatography in buffer containing 6M guanidine. The material obtained from the size exclusion chromatography gave a very broad peak with no baseline resolution between the Sup35NM material and the lower molecular weight material (Fig. 3.9). The material to the left of the peak contained the purest protein. LC-MS analysis of the product shows formation of the desired product, 30549.62Da, with removal of the N-terminal methionine residue by methionine aminopeptidase (MAP) (Frottin et al., 2006).

3.2.1.3   Sup35NMC – Full length protein

It has recently become apparent that the C terminal domain of Sup35 can influence its polymerisation and cause phenotypic variation (Kabani et al., 2011; Krzewska and Melki, 2006; Krzewska et al., 2007).

To investigate fibril assembly of Sup35NMC, a pET15b vector containing the gene for Sup35NMC, obtained from (Kabani et al., 2011; Krzewska and Melki, 2006; Krzewska et al., 2007), was transformed into competent BL21 [DE3] pLysS *E. coli* to give an expression system capable of producing full length Sup35.

Previously published work on the generation of Sup35NMC fibrils has utilised a low yielding native purification strategy (Baxa et al., 2011; Krzewska and Melki, 2006; Krzewska et al., 2007) to produce the protein. Following the success of the purification of Sup35NM using denaturing conditions, the same denatured purification approach was attempted with Sup35NMC. However utilising a completely denaturing purification strategy from lysis onwards not only solubilises more of the Sup35NMC but other cellular material. These

proved difficult to remove in subsequent affinity purification (Ni-NTA) and size exclusion chromatography steps.

While Sup35NM is almost entirely contained in insoluble inclusion bodies during expression in *E. coli* (Glover et al., 1997) some Sup35NMC remains soluble. The purity of Sup35NMC from affinity purification of lysate generated under native, folded, conditions is cleaner than protein isolated from a denatured lysis. Therefore in order to maintain the advantages of the denatured material, a hybrid purification was attempted. Consequently cells were disrupted under native conditions in the presence of protease inhibitor, PMSF, followed by Ni-NTA affinity purification. The protein was then unfolded on the resin by washing and eluting with buffer containing 6M guanidine. This resulted in a Sup35NMC solution with much higher purity prior to the final size exclusion chromatography step and allowed successful purification (Fig. 3.10).
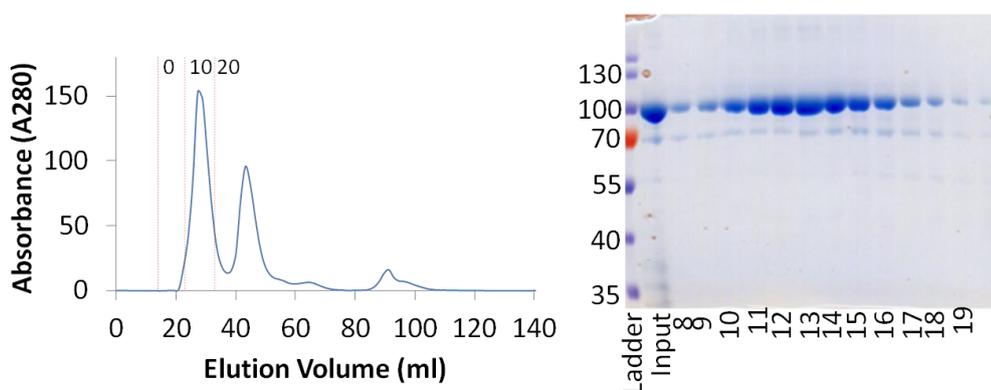
Figure 3.10   **Purification profile of Native -Denaturing Purification**
Left – Äkta purification profile of Sup35NMC lysed under native conditions and then denatured during affinity purification and then further purified using size exclusion purification, 0ml elution = injection, giving rise to a single major peak eluted between 20 and 30mls with the start of fraction collection, the 10th fraction and 20th fraction marked with a dotted line. Right – 10% Tris-Glycine SDS PAGE analysis of the fractions from the purification diluted 1 in 3 prior to loading onto the gel. The lane numbers are numbered according to the fraction number seen in the Äkta trace.

The purification was also attempted using the native conditions previously published but the disadvantage of this process is the continual polymerisation of material during the purification means the whole procedure needs to be completed at low temperature and as

quickly as possible. Disruption of the *E.coli* in the presence of phenylmethanesulfonyl fluoride, PMSF, and 2-mercaptoethanol allowed initial purification by affinity purification, Ni-NTA, allowing analysis by LC-MS and showed the protein was modified by the addition of 2-mercaptoethanol, presumably by disulfide formation.
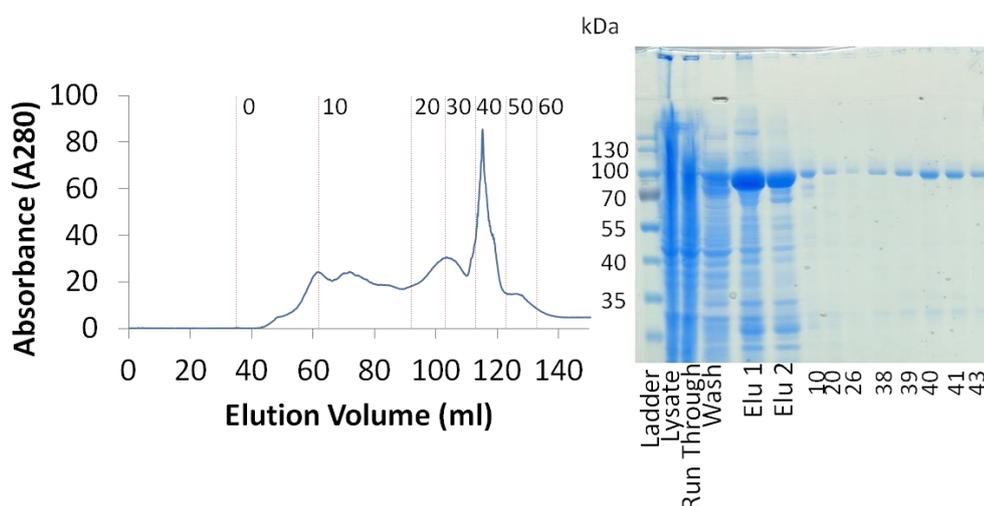
Figure 3.11   Purification of Sup35NMC with DTT in Lysis buffer
Left – Äkta purification profile of Sup35NMC purified natively at 1ml/min after lysis in the presence of Dithiothreitol giving rise to a broad elution profile with a major peak at about 115 minutes, in contrast to the purification with no thiol added. The position of every 10 fractions collected are marked on the graph with a dotted red line. Right – 10% Tris-Glycine SDS PAGE analysis of the fractions from the purification with the run through, wash and elu1 and elu 2 from the Ni-NTA purification. The fractions are then labelled according to the appropriate fraction number from the Äkta spectra.

To avoid this modification the 2-mercaptoethanol used in the lysis buffer was replaced with dithiothreitol (DTT), a stronger reducing agent whose oxidation results in the formation of a cyclic product through intra-molecular disulfide bond formation and therefore has a lower propensity to modify proteins. Using a concentration of DTT low enough not to reduce the Ni(II) based affinity matrix the modification of Sup35NMC was limited, as determined by LC-MS, but the final purification of the protein by size exclusion chromatography was complicated by apparent association of the protein with the gel matrix. The 2ml of injected Sup35NMC was eluted across 80ml with the majority eluted after almost one column volume, 120ml (Fig. 3.11).

Due to the problems associated with the addition of reducing agents purification in the absence of a thiol seemed, by FPLC trace, to give a better elution profile, more akin to the native denaturing protocol, where the protein does not interact with the matrix (Fig. 3.12, Left). Analysis of the protein material by SDS PAGE showed that in fact the material contained a lower molecular weight impurity, which is co-purified due to its association with the Sup35NMC (Fig. 3.12, Right). LC-MS comparison of the material from the purification with the same material treated with 5mM DTT prior to injection into the spectrometer suggests that the material is covalently modified, via a disulphide bond as the only trace of the desired material is in the presence of DTT (Fig. 3.13). This potentially explains the relatively large amount of impurity observed in the material from the size exclusion purification.
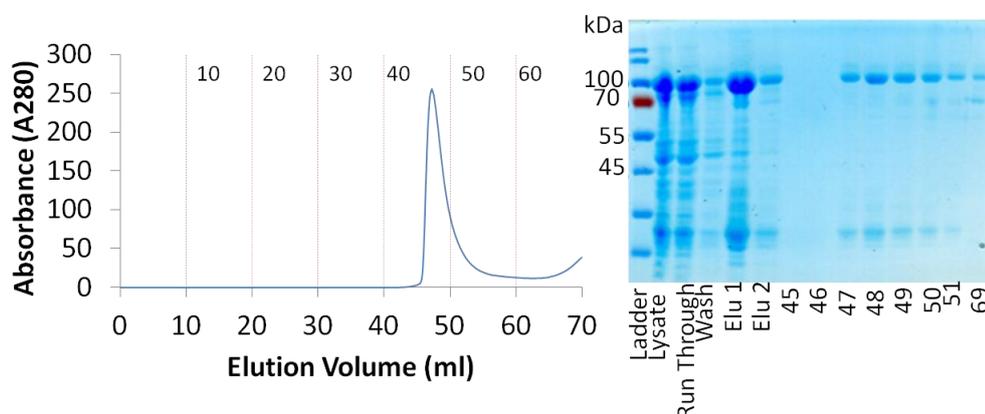
Figure 3.12   Purification of Sup35NMC without thiol treatment
Left – Äkta purification profile of Sup35NMC purified by affinity purification and then size exclusion purification, in native buffer at 1ml/min, after lysis giving rise to a single major peak at about 50 minutes. 1ml fractions were collected from injection with every 10th fraction marked with a dashed line. Right – 10% Tris-Glycine SDS PAGE analysis showed the material from the Ni-NTA purification: Run Through, Wash, Elution 1 and Elution 2 followed by fractions from the size exclusion purification labelled according to their fraction number.

The presence of a reducing agent in the purification obviously has some effect on the structure and ability of the protein to 'stick' to the size exclusion matrix. To determine if this is a property of reduced material on the sepharose column, despite its potential to modify the protein, 2-mercaptoethanol was used in the purification and progressed to size

exclusion chromatography. In the presence of 2-mercaptoethanol the normal elution profile was returned and the lower molecular weight impurities were diminished although not totally removed (Fig. 3.14).
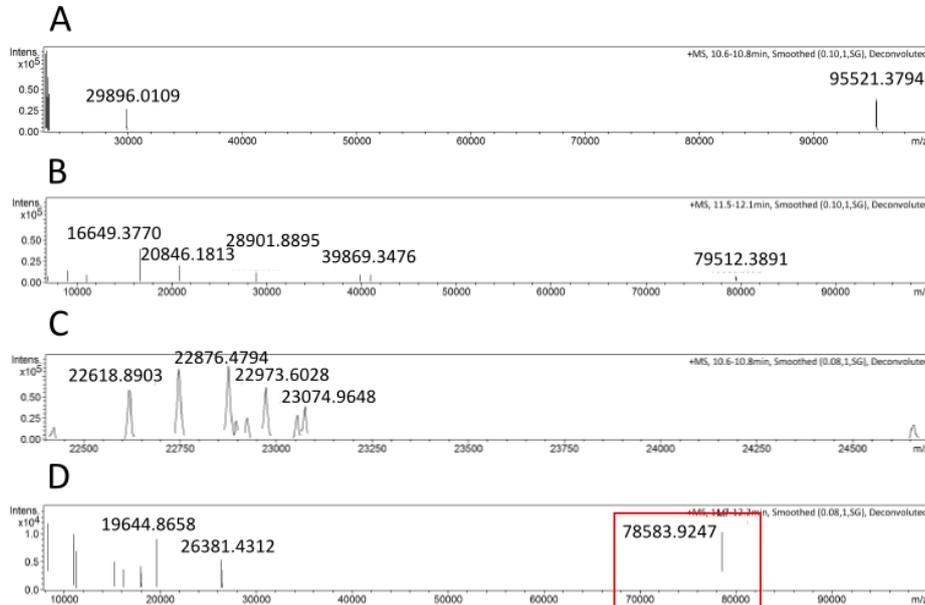
Figure 3.13   MS analysis of Sup35NMC with and without Thiol

Sup35NMC purified in the absence of thiol was analysed by LC-MS with (C and D) and without 5mM DTT (A and B). The deconvoluted MS spectra displaying m/z (mass/charge) on the x axis, where z=1, and therefore m/z is equal to the mass in Da. The two main components earlier (A and C) and later (B and D) observed in the LC were deconvoluted to give parent ions. The data shows that only in the later eluting (higher molecular weight) material which had been treated with DTT is the correct ion, 78583Da, for the product observed see the red box in D.
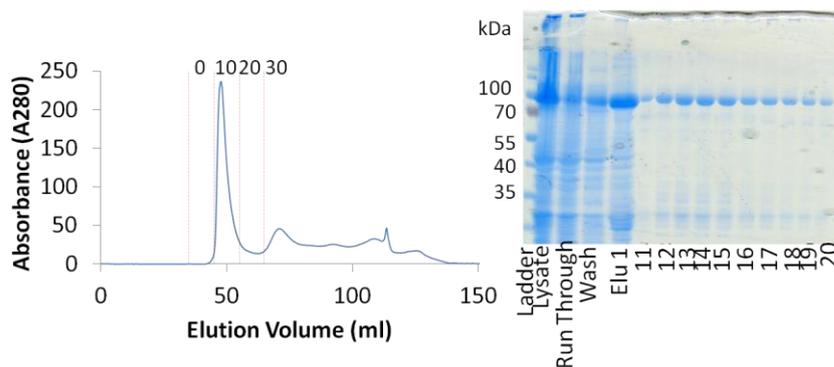
Figure 3.14   Purification of Sup35NMC with Lysis in the presence of 2-mercaptoethanol

Left – Äkta purification profile of Sup35NMC purified after lysis in the presence of 2-Mercaptoethanol giving rise to a major peak at about 50 minutes, as observed in the purification with no thiol added. The start of fraction collection and every 10th fraction were marked with a red dashed line. Right – 10% Tris-Glycine SDS PAGE analysis of the elution from the Ni-NTA purification – Lysate, Run Through, Wash and Elu1 followed by the fractions eluted from the size exclusion column labelled according to the fraction from the Äkta purification.

In contrast to the MS analysis pre-size exclusion purification, the material obtained afterwards showed no evidence of a 2-mercaptothanol adduct but also N-terminal methionine cleavage, 78583.92Da, by MAP as observed in the purification of Sup35NM.

3.2.2   *In Vitro* Assembly

The conditions required to convert monomers to fibrils varies considerably between proteins and can potentially affect the morphology of the fibrils formed. The amyloid fibrils physical properties and fragmentation requires formation of polymers and deposition conditions for AFM imaging which allow for automated tracing.

3.2.2.1   α-Synuclein

Due to the intrinsically disordered nature of α-synuclein, the conversion to amyloid fibrils is straight forward. The dynamic nature of α-synuclein means that it may spend some time in a conformation susceptible to amyloid formation therefore harsh conditions are not required to unfold monomeric α-synuclein into amyloid fibrils. In fact, incubation of >150µM α-synuclein at physiological temperature, 37°C, in a neutral buffer for 1 week is enough to convert α-synuclein into its fibrillar form.
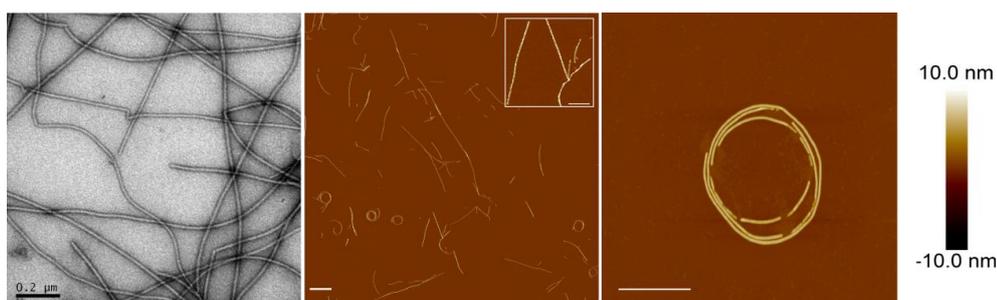
Figure 3.15   α-Synuclein Polymerisation into Fibrils
Left – A negatively stained, uranyl acetate (UOAc$_2$), TEM micrograph of α-synuclein fibrils with 0.2µm scale bar, Middle – An AFM image, 20µm x 20µm at 2048 pixels x 2048 pixels with a 1µm scale bar, Right - An AFM image, 2µm x 2µm at 1024 x 1024 pixel resolution with a 500nm scale bar, focussing on an example of a highly coiled α-synuclein fibre. The AFM height scale is set between -10 and 10nm.

The purified α-synuclein monomers were converted in to fibrils by first removing the high salt solution from the anion exchange medium by buffer exchange using PD10 column into 25mM sodium phosphate, pH7.5. The sample was incubated at 37°C, and agitated at

160rpm for one week. Confirmation that amyloid fibrils were formed in solution was achieved by negative stain TEM imaging as well as AFM imaging (Fig. 3.15).

The assembly progression of monomeric α-synuclein to fibrillar material was shown to have gone almost to completion, 2.4% remaining, by SDS PAGE based residual monomer assay against a serial dilution of monomeric protein.

AFM analysis of the fibrillar material identified the presence of fibrils with highly coiled fibril morphology (Fig. 3.15, Right) the absence of these structures in TEM analysis suggests that they are an artefact of deposition on to mica for AFM analysis and not how these samples exist in solution. The automation of particle tracing using the MatLab application requires well dispersed particles where fibrils cross over is limited, this highly coiled morphology would make this process difficult and therefore alternative deposition conditions this were investigated.

Optimisation of the conditions of α-synuclein deposition onto mica for tracing purposes (Fig. 3.16) highlighted that the drying process after sample incubation and washing was causing fibril breakage.

This breakage would give rise to error in subsequent particle length analysis so reduction of this artefact was also important. To minimise breakage, the drying process was modified such that most of the excess water on the mica surface was removed by filter paper placed at the very edge of the mica surface and then carefully dried with a gentle stream of nitrogen gas. The final conditions for α-synuclein deposition, which give the appropriate particle density for particle length measurement, whilst minimising coiling and particle breakage on the surface involved the dilution of the fibril sample to 0.48μM with fibril forming buffer, incubation for 5 minutes prior to washing with 1ml of syringe filtered mQH$_2$O and drying using the modified procedure.
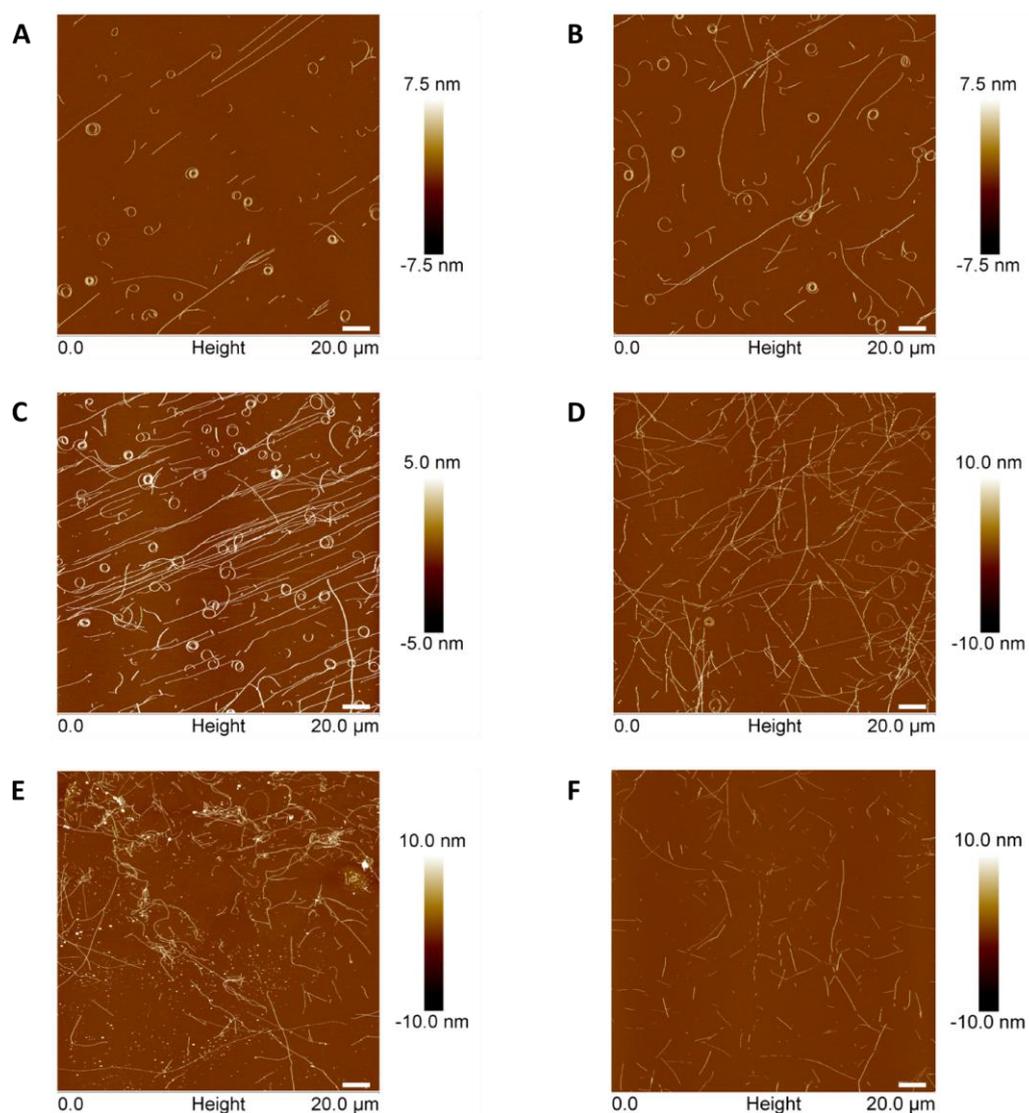
Figure 3.16   Mica Depostion Conditions

AFM images showing the development of AFM deposition conditions. 2µm scale bar in each image. A – 1 in 50 dilution into fibril forming buffer, 5 minutes incubation, 1 ml wash and dried with $N_{2(g)}$. B – 1 in 50 dilution into fibril forming buffer, with 0.001% Sarkosyl, 5 minutes incubation, 1ml wash and dried with $N_{2(g)}$. C – 1 in 50 dilution, with 400µM $NiCl_2$, 5 minutes incubation, 1ml wash and dried with $N_{2(g)}$. D – 1 in 1000 dilution, 10µl dried onto mica surface, washed with 1ml and dried under a stream of $N_{2(g)}$. E – Similar conditions to D but highlighting the problems with drying sample onto the mica surface. F – Fibrillar material diluted with pH2 HCl and incubated on mica before washing with water. The AFM height scales set so as to show fibril height and ranges from -5 to 5nm and -10 to 10nm.

3.2.2.2   *Sup35NM and Sup35NMC*

The N and M domains of Sup35NM and Sup35NMC are relatively disordered meaning they

can exist in numerous conformational states, similar to α-synuclein, and do not require

harsh conditions for fibril assembly. Conversion of denatured Sup35NM and Sup35NMC to

their fibrillar form was initated by buffer exchange using PD10 column into fibril forming

buffer containing 20mM sodium phosphate, 50mM sodium chloride, pH7.5 solution. The

resulting solution was incubated at 30°C, a temperature that mimics the conditions that the

Sup35 protein is exposed to in the optimum growth environment of *Saccharomyces*

cerevisiae.
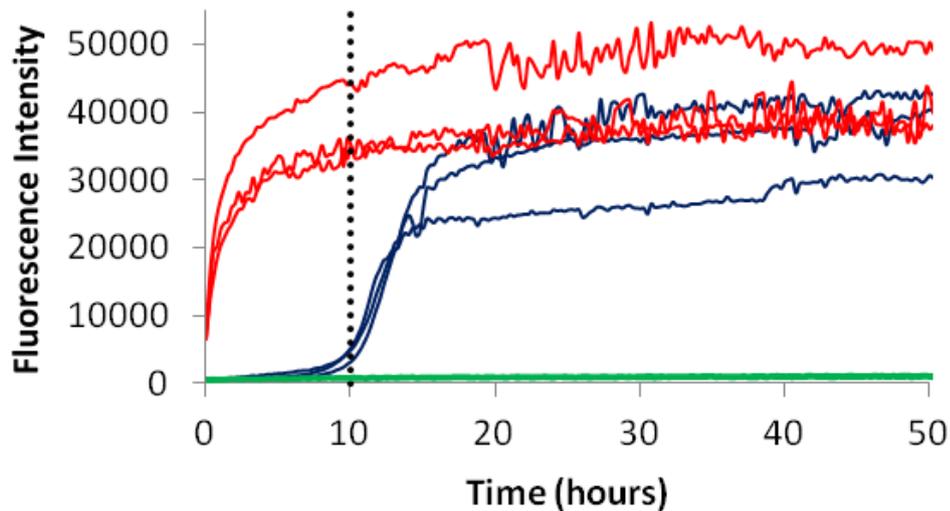
Figure 3.17   Kinetics of Denatured Sup35NM and Sup35NMC Conversion to Amyloid

Thioflavin T curves showing the polymerisation of Sup35NM and Sup35NMC from denatured purification. 3 repeats of a blank corrected *de novo* polymerisation of Sup35NM (Blue) (10uM Sup35NM, Thioflavin T and buffer), seeded Sup35NM polymerisation (Red) (10uM Sup35NM, 1% sonicated Sup35NM seed, Thioflavin T and buffer) and *de novo* Sup35NMC from denatured purification polymerisation (Green) (10uM Sup35NMC, Thioflavin T and buffer). The emission of Thioflavin T was monitored at 493nm. Black dotted line signifying 10 hours of polymerisation.

The kinetics of fibril formation can be followed by measuring the fluorescence associated

with the binding of the profluorescent dye thioflavin T to the cross-beta sheet core of the

amyloid backbone. The *de novo* conversion of monomeric Sup35NM into amyloid fibrils,

follows the classic sigmoidal shaped curve associated with a nucleated growth mechanism.

The assembly has a lag phase of about 10 hours which is followed by an exponential phase

where the proportion of cross-beta sheet in solution increases rapidly (Fig. 3.17). The lag

phase of fibril growth is dominated by the formation of critical nuclei, as observed in the *de*

novo assembly of Sup35NM but can be completely removed when a seed, a small amount

or preformed fibrillar material, is added to the reaction mixture (Fig. 3.17). The polymerisation of Sup35NMC from the denatured purification showed no increase in ThT fluorescence over the course of the reaction even when seeded with fragmented Sup35NM fibrils (Fig. 3.17, *de novo* only shown). As the polymerisation of Sup35NMC has been observed previously using monomeric protein derived from a native purification the lack of polymerisation is potentially due to the failed refolding of the C domain of Sup35NMC and the non specific association of these unfolded materials into amorphous aggregates.

Comparison of the natively purified Sup35NMC in the absence of thiol, in the presence of 2-mercaptoethanol and also material isolated from the final peak of material eluted from the DTT purification were compared by thioflavin T assay. The change in fluorescence over time for these experiments shows that the material purified in the presence of DTT does not polymerise, like the denatured Sup35NMC, where-as the material formed in the absence of thiol and the presence of 2-mercaptoethanol gives rise to an increase in beta-sheet composition over time (Fig. 3.18).

The potential modification of the material generated in the absence of thiol does not seem to alter its propensity to form amyloid fibrils and a lag phase of about 10 hours was observed (Fig. 3.18, A) similar to Sup35NMC purified with 2-mercaptoethanol (Fig. 3.18, B), Sup35NM (Fig. 3.17) and the data shown by Krzewska et al (Krzewska et al., 2007).

Not only does the incorporation of 2-mercaptoethanol in the lysis buffer return the Äkta elution profile back to that of the material purified without thiol but its ability to form amyloid fibrils was also returned. Although thioflavin T binding showed that the *de novo* formation had a lag phase of just less than 10 hours in comparison to the thioflavin T curves for Sup35NM and Sup35NMC without thiol the lag phase exhibited a decrease in fluorescence prior to the exponential phase which means the curves are flatter in look but do exhibit the sigmoidal shape (Fig. 3.18, B).
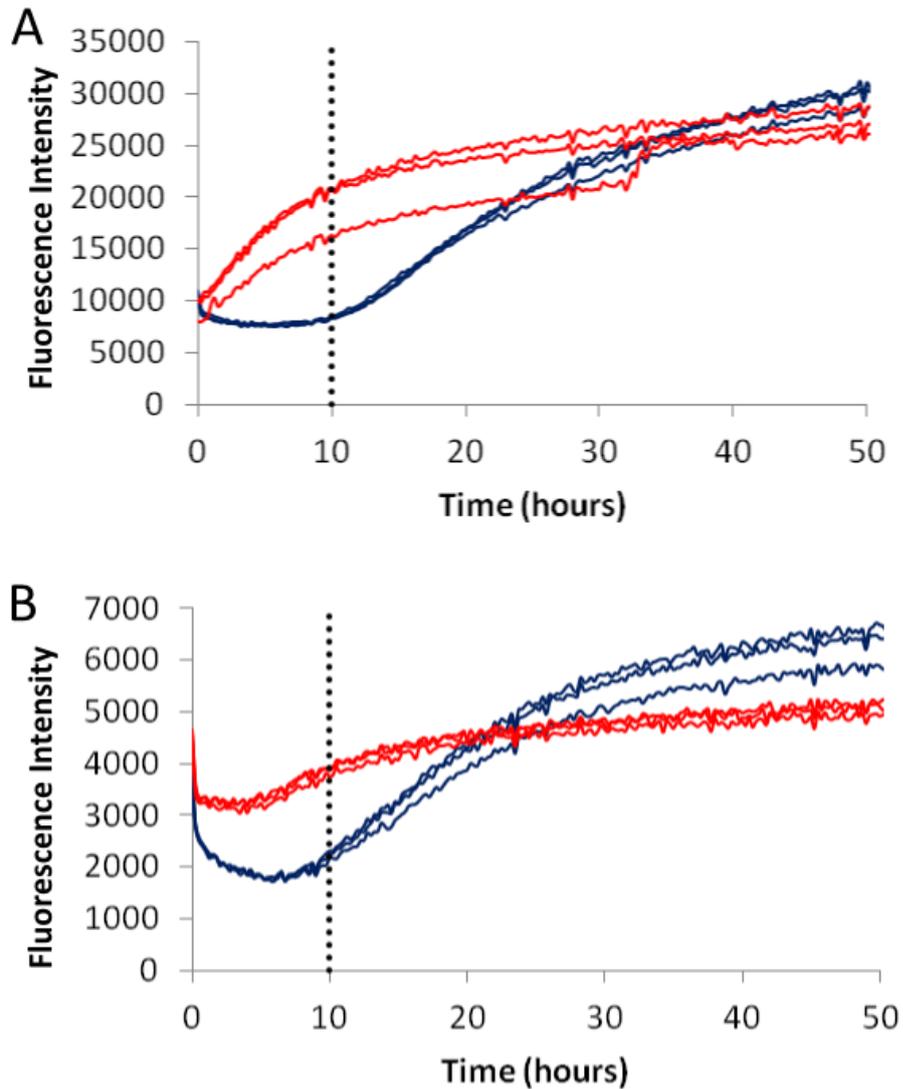
Figure 3.18   Kinetics of Native Purified Sup35NMC Conversion to Amyloid

Thioflavin T curves showing the polymerisation of Sup35NMC derived from native purification protocols. The emission of Thioflavin T was monitored at 493nm. The black dotted line signifies 10 hours of polymerisation.

A – No thiol in purification - 3 repeats of blank corrected *de novo* polymerisation (Blue) (10uM Sup35NMC, Thioflavin T and buffer) and seeded polymerisation (Red) (10uM Sup35NMC, 1% sonicated Sup35NM seed, Thioflavin T and buffer).

B – 2-mercaptoethanol purification - 3 repeats of blank corrected *de novo* polymerisation (Blue) (10uM Sup35NMC, Thioflavin T and buffer) and seeded polymerisations (red) (10uM Sup35NMC, 1% sonicated Sup35NM seed, Thioflavin T and buffer).

This abnormal profile is potentially due to the defrosting of the monomer solution prior to fibril formation and the sample being partially aggregated before it depolymerises as it warms up in the plate reader, although it is unclear exactly what is happening. The

addition of Sup35NM seed molecules to both the experiments removed the lag phase completely as would be expected with a nucleated growth mechanism.
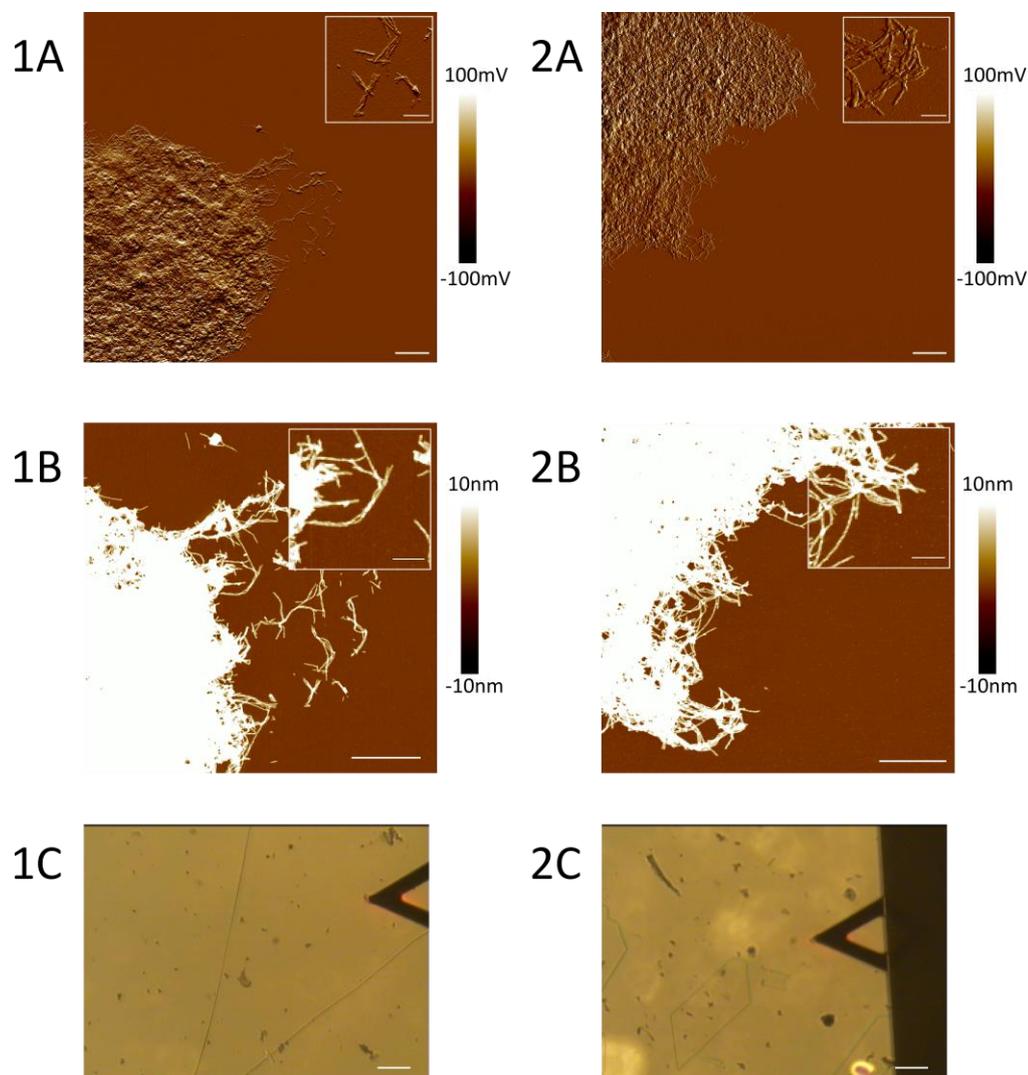
Figure 3.19   AFM Analysis of Fibrils Derived from Sup35NM

AFM imaging of Sup35NM polymerisation 1 = Sup35NM 10μM de novo polymerisation, 2 = Sup35NM 10μM seeded with Sup35NM fibrils (1% seed, sonicated at 20% amplitude for 4 minutes, 5 second pulses and on ice). A = Peak force error image showing 10μm x 10μm area with 1μm scale bar, 1μm insert which is magnified 3x with a 250nm scale bar and a peak force scale on the left hand side; B = Height based image showing a 5μm x 5μm area with a 1μm scale bar, 1μm insert which is magnified 2x with a 250nm scale bar and a height scale on the left hand side; C – Microscope image from the 10x optical microscope attached to the AFM showing a 1mm scale bar showing the presence of large aggregates on the mica surface; D – 3 separate 2μm sections were compared to show consistency in height between different purification methods. The scale bar for the peak force error spectra is set between 100 and -100mV and the height phase is set between 10 and -10nmm.
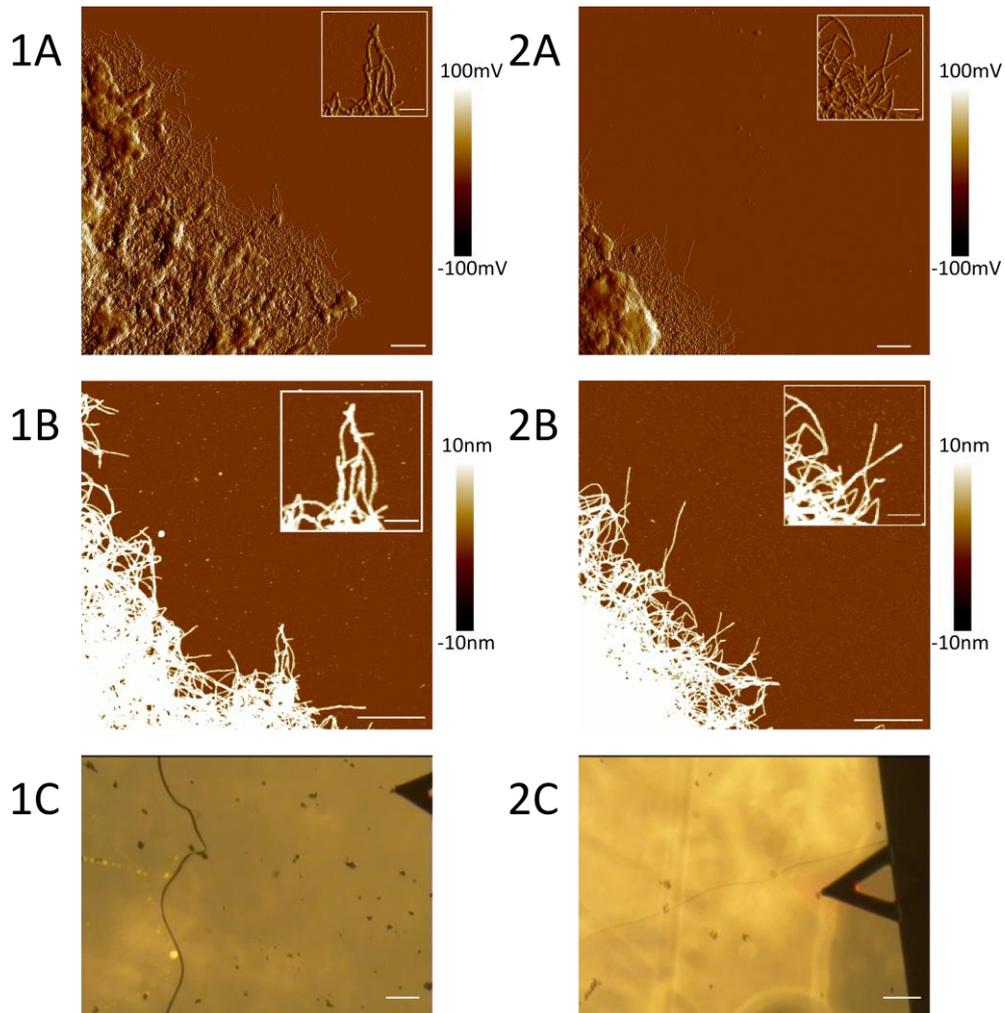
Figure 3.20   **AFM Analysis of Sup35NMC polymerisation**

AFM imaging of de novo (unless stated) Sup35NMC polymerisation 1 = Sup35NMC purified without thiol added, 2 = Sup35NMC purified in the presence of 2-mercaptoethanol. A = Peak force error image showing 10μm x 10μm area with 1μm scale bar, 1μm insert which is magnified 3x with a 250nm scale bar and a peak force scale on the left hand side; B = Height based image showing a 5μm x 5μm area with a 1μm scale bar, 1μm insert which is magnified 2x with a 250nm scale bar and a height scale on the left hand side; C – Microscope image from the 10x optical microscope attached to the AFM showing a 1mm scale bar showing the presence of large aggregates on the mica surface; D – 3 separate 2μm sections were compared to show consistency in height between different purification methods (3D contains a dashed line at 10nm). The scale bar for the peak force error spectra is set between 100 and -100mV and the height phase is set between 10 and -10nmm.

AFM analysis of the Sup35NM and Sup35NMC materials after the assembly has reached the stationary phase of the assay shows the presence of fibrillar structures but in contrast to the dispersed nature of α-synuclein fibrils the fibrils of Sup35NM and Sup35NMC, purified without thiol and with 2-mercaptoethanol, exist in a highly associated state in clumps of fibrils which are also visible using the optical microscope used for AFM laser alignment,

almost 400μm in height (Fig. 3.19 and 3.20, 1-2C). AFM imaging of these aggregates by peak force and height showed that the internal structure is that of a large network of fibrils (Fig. 3.19 and 3.20, 1-2A and 1-2B).

By comparison material, taken at the same point from the denatured Sup35NMC polymerisation showed that large aggregates were formed but unlike those formed by Sup35NM, which were formed by the association of amyloid fibrils, these were amorphous in agreement with ThT data. AFM imaging of the assembly reactions derived from the DTT purified material did not give rise to any aggregates, fibrillar or otherwise but the seeded sample gave rise to large amorphous aggregates still visible by the optical microscope but not a complex association of fibrils.

3.2.2.3   **Insulin B-Chain**

Literature precedence shows that the conversion of insulin B-chain into amyloid fibrils can be readily and reproducibly achieved by heating pH2 solutions of insulin B-chain to 60°C for periods of 1 day (Devlin et al., 2006; Hong and Fink, 2005).
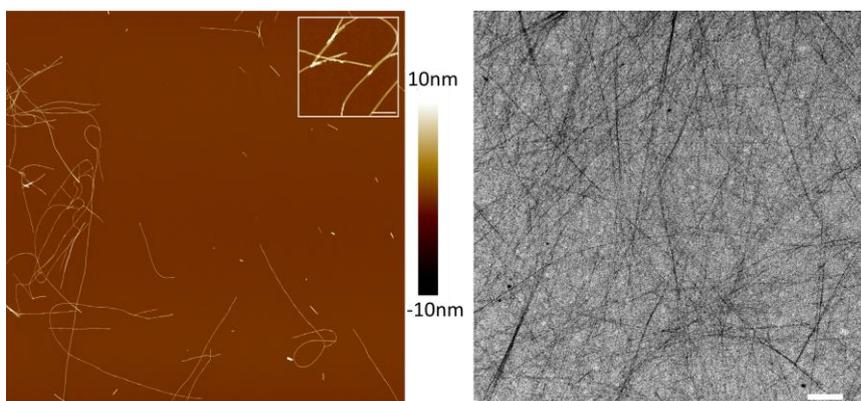
Figure 3.21   **AFM and EM Analysis of Fibrils derived from Insulin B-Chain**
Insulin B-chain fibrils Left - deposited onto a mica surface, washed, dried and imaged at 20μm x 20μm at a resolution of 2048 pixels x 2048 pixels, with a 1μm scale bar and 1μm insert magnified x5 containing a 250nm scale bar. The AFM height scale is set from -10 to 10nm. Right – Fibrils deposited onto a carbon/formavar coated copper surface and then imaged using TEM, with a 1μm scale bar.

A 1.8mM solution of insulin B-chain was syringe filtered and heated to 60°C for 48 hours at which time conversion to amyloid was confirmed by AFM and TEM imaging (Fig. 3.21).

Recent work by Piejko *et al* (Piejko et al., 2015) has shown that the conversion of Insulin under acidic conditions is promoted by limited proteolysis with Pepsin and the generation of smaller, more amyloidogenic, peptides is the root cause of the conversion of insulin into amyloid fibrils. The amide bonds which make up the peptide back bone are usually robust but under heated acidic conditions are susceptible to breakage and therefore a hydrolytic mechanism can also explain the conversion of insulin and its component chains, A and B, into fibrils in the absence of Pepsin.

3.2.2.4   Lysozyme

The conversion of Lysozyme into amyloid fibrils also follows an acid-promoted mechanism. Following similar conditions to insulin B-chain, the heating of syringe filtered monomer solution to $60^{o}C$ for 48 hours at pH2 yielded fibrils but far less efficiently than observed with Insulin B-Chain.

Assuming that amide hydrolysis is the key step in converting monomers to amyloid the efficiency of fibril formation could be improved by increasing the amount of hydrolysis in the system. An increase in the hydrolysed proportion of the protein can be achieved by lengthening the heating time or by increasing the temperature of the reaction (Fig. 3.22).

Increasing the incubation time and temperature will both increase the amount of hydrolysis experienced by the sample but this does not necessarily correlate with an increase in the amount of fibrils formed. As a rule of thumb, derived from the Arrhenius equation, the rate of a thermodynamically controlled reaction doubles for every $10^{o}C$ increase in reaction temperature. By increasing the temperature for $60^{o}C$ to $90^{o}C$ an eight fold increase in the speed of fibril formation would be expected but the result observed is of no fibril formation (Fig. 3.22, C).
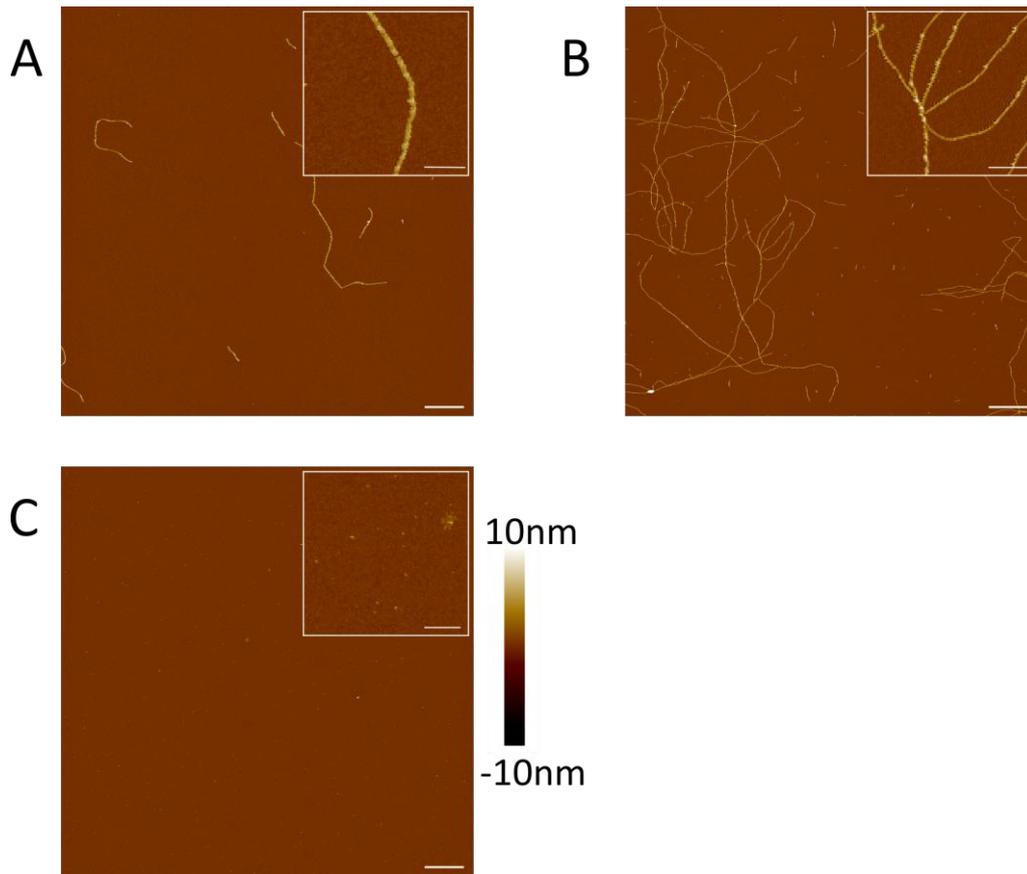
Figure 3.22   Temperature and Incubation Time Affect Lysozyme Fibril Formation
10μm x 10μm AFM scans at 1024 pixels x 1024 pixels resolution with a height scale of between -10 and 10nm showing the change in composition of fibril forming reactions after heating at pH2 for **A** = 60°C for two days, **B** = 60°C for six days and **C** = 90°C for 18 hours. 699μM (10mg/ml) Lysozyme was heated then diluted 1 in 1500 in pH2 HCl and deposited onto mica according to 2.34.1.

There are several possible explanations for the absence of fibrils:

1. The increase in temperature gives a higher rate of hydrolysis but also rate of critical nucleus formation such that it is the overwhelming process and therefore retards fibril elongation.

2. The increase in reaction temperature amplifies the rate of hydrolysis but to a degree which hydrolyses the key sequences required to interact for fibril formation.

3.   The increase in reaction temperature could also cause any amyloid formed to disaggregate.

Rather than trying to augment the hydrolysed proportion by increasing the rate of hydrolysis it could be altered by heating the samples for longer. By extending the incubation time at 60°C to six days the degree of polymerisation observed in the reaction was amplified, as seen by a greater number of particles observed on the mica surface (Fig. 3.22, A and B), giving the final conditions for fibril assembly.

To qualitatively examine the role of hydrolysis in fibril formation for lysozyme, monomeric protein and fibrillar samples were examined by SDS PAGE (Fig. 3.23). Two different denaturation temperatures, 70°C and 95°C, were examined in the presence and absence of 2-mercaptoethanol, to show that gel sample preparation isn't the cause of the degradation observed (Fig. 3.23). The fibril samples show that the process of fibril formation generates a large amount of lower molecular weight materials, and confirms the hydrolytic mechanism.
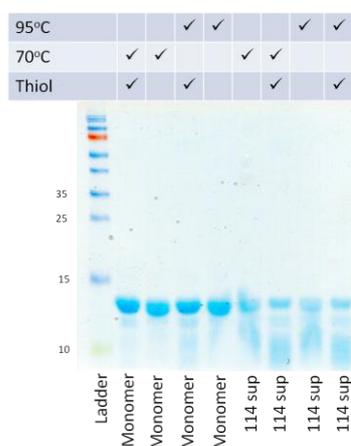
Figure 3.23   SDS PAGE analysis of the degradation of Lysoyme during fibril formation

15% Tris-Glycine gel run at 180V showing the stability of Lysozyme monomers to SDS PAGE sample preparation (left 4 lanes) and the hydrolysis of the monomers during fibril formation (right 4 lanes).

Due to the hydrolytic mechanism of fibril formation at pH2 and elevated temperature the efficiency of fibril formation can be linked to the consumption of lysozyme, SDS PAGE comparison of the supernatant of an ultracentrifuged, 75000rpm for 15 minutes, fibril

sample to a serial dilution of the parent protein can give a percentage conversion of monomer to fibrils (Fig. 3.24). Densitometry of the serial dilution bands and comparison to the fibril samples shows that the conversion is 55-60% complete.

It has subsequently be shown that the mechanism for conversion of lysozyme from monomer to fibril does indeed occur via a hydrolytic mechanism whereby heating in acid breaks the amide linkages in the protein (Frare et al., 2006) and if heated long enough the major component of the fibrils is in fact a 6 residue peptide, ILQINS, (Lara et al., 2011a, 2013, 2014) (Fig. 3.25).
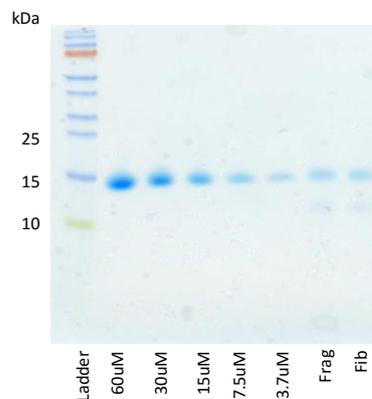
Figure 3.24   SDS PAGE Analysis of the Conversion of Lysozyme Monomer to Fibril

Residual monomer assay showing the diminishing of Lysozyme monomer under the conditions used to form amyloid fibrils. Ladder on the left followed by a serial dilution of Lysozyme and then (far right) the top 10µl of a 120µM centrifuged fibril sample was diluted 1 in 10 and then analysed on a Tris-Tricine gel SDS PAGE. The assays shows 54µM residual monomer therefore the reaction has progressed to 54% completion. The Frag sample is of the same sample subjected to fragmentation.

```
KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDGSTDYGILQINS
RWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIVSDGNGMNAWVAWRNRCKGTDV
QAWIRGCRL
```

Figure 3.25   Amino Acid sequence of Hen Egg White Lysozyme with Core Amyloid Region Highlighted

The primary amino acid sequence of Hen Egg White Lysozyme with the amino acid sequence found to form fibrils highlighted in yellow.

3.2.2.5   β-Lactoglobulin

The conversion of β-lactoglobulin into amyloid fibrils can also be achieved by heating in acidic medium similar to insulin B chain and lysozyme. β-Lactoglobulin forms a dimeric structure at neutral pH but at pH2 the dimeric structure collapses into monomer (Hamada and Dobson, 2002). Initial experiments using the same conditions as insulin B chain, pH2 at 60°C for 48 hours, did show evidence of some fibril formation but this, like lysozyme, was very inefficient.

Previous experiments (Bolisetty et al., 2011; Gosal et al., 2004; Krebs et al., 2009) have shown that the conversion of β-lactoglobulin to fibrils required a higher temperature, 80°C, for 24 hours. Modification of these conditions to heating at 90°C in pH2 for 24 hours showed no evidence of monomeric proteins in the sample (Fig. 3.26) and a much higher efficiency of fibril formation as observed by the number of fibrils observed by AFM and TEM (Fig. 3.27).
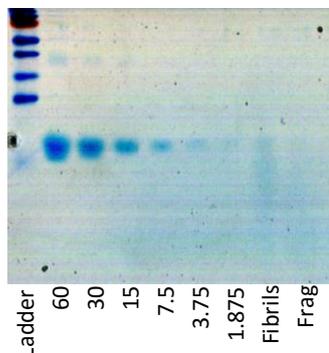
Figure 3.26   SDS PAGE Analysis of the Conversion of β-Lactoglobulin Monomer to Fibrils

Residual monomer assay showing the diminishing of β-Lactoglobulin monomer under the conditions used to form amyloid fibrils. Ladder on the left followed by a serial dilution of β-Lactoglobulin and then (second from right) the top 10µl of a 120µM centrifuged fibril sample was then analysed on a Tris-Tricine gel SDS PAGE. The assays shows 2µM residual monomer therefore the reaction has progressed to 98% completion. The Frag sample is the same as for the Fibrils sample except that it has been subjected to fragmentation.

Similar to the observation that the hydrolysis of Lysozyme and Insulin are important for fibrillogenesis, this also seems to be an important aspect of β-lactoglobulin and the recent work of Akkermans *et al* (Akkermans et al., 2008) has confirmed this hydrolytic mechanism.
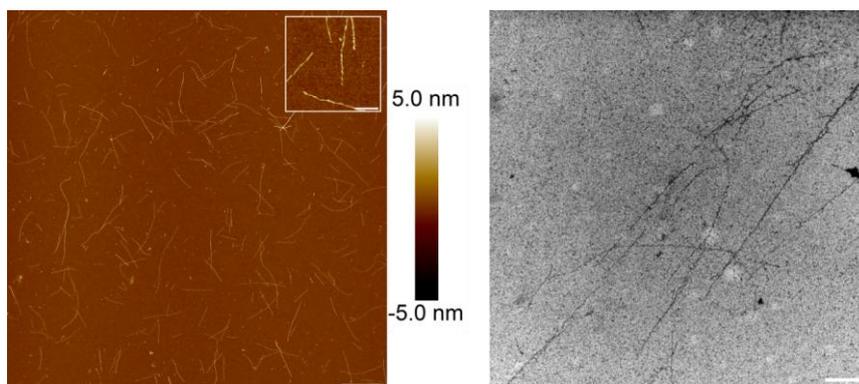
Figure 3.27   TEM and AFM Analysis of β-Lactoglobulin Fibrils

β-Lactoglobulin fibrils Left - deposited onto a mica surface, washed, dried and imaged at 20μm x 20μm at a resolution of 2048 pixels x 2048 pixels, with a 1μm scale bar and a 1μm insert magnified x5 containing a 250nm scale bar. The AFM height scale is set between -5 and 5nm. Right – Fibrils deposited onto a carbon/formavar coated copper surface and then imaged using TEM, with a 1μm scale bar. The AFM height scale bar is set between -5 and 5nm.

3.2.3   AFM Imaging and Comparison of Fibril Structures

3.2.3.1   Qualitative Comparison

Comparison of the fibril species formed through *in vitro* polymerisation shows that the fibrils of Sup35NM and Sup35NMC aggregate into extremely large clusters whereas all other fibrils species examined form discrete particles (Fig. 3.28).

Visual inspection of the surfaces shows that β-lactoglobulin forms straight fibrils of a maximum length of roughly 4μm in comparison to α-synuclein, lysozyme and insulin fibrils which form longer, more flexible particles with rough maximum length of 10μm, 10μm and 15μm respectively (Fig. 3.28).

To compare the heights of fibril samples cross sections were taken (Fig. 3.29). The large aggregated state of Sup35NM and Sup35NMC fibrils means that the edges of the aggregates is the only place where individual fibrils are visible and therefore 2μm sections around this region were taken. Sup35NM and Sup35NMC fibrils both appear to have an approximate height of about 7-8.5nm (Fig. 3.29). The discrete particles observed for insulin
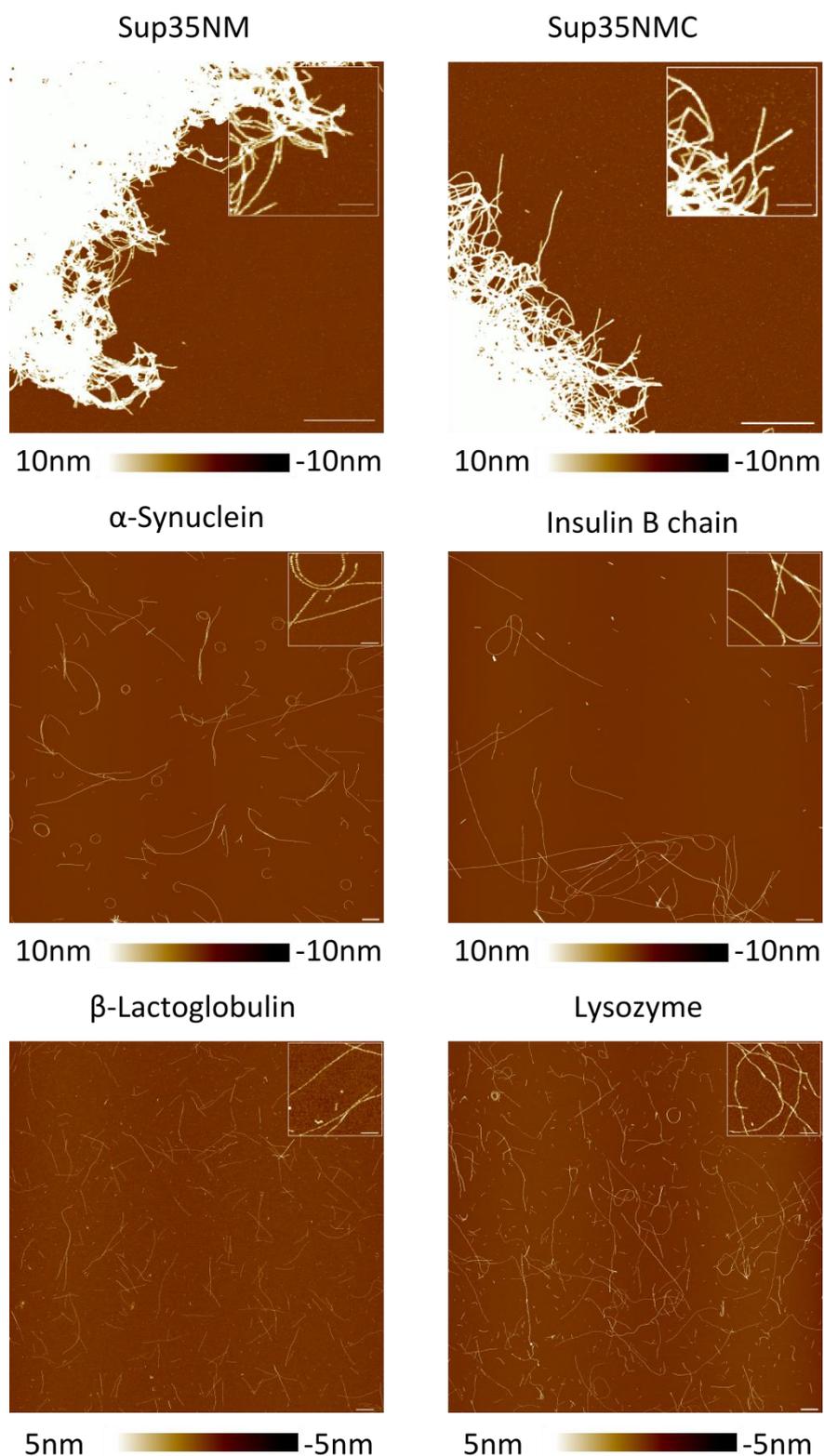
Figure 3.28   AFM Images of Fibrils Produced

AFM images of fibrils produced. Sup35NM and Sup35NMC - 10μm x 10μm images at 1024 x 1024 pixels with 2μM scale bar and a 1μm insert magnified x 3. The other proteins imaged at 20μm x 20μm and 2048 x 2048 pixel resolution and 1μm scale bar and 1μm insert magnified x 5 and a 250nm scale bar. Sup35NMC, Sup35NM, α-synuclein and insulin B-chain have height scales of between -10 and 10nm where as the scale for β-Lactoglobulin and lysozyme are between -5 and 5nm.

Sup35NM

Sup35NMC

α-Synuclein

Insulin B chain

β-Lactoglobulin

Lysozyme
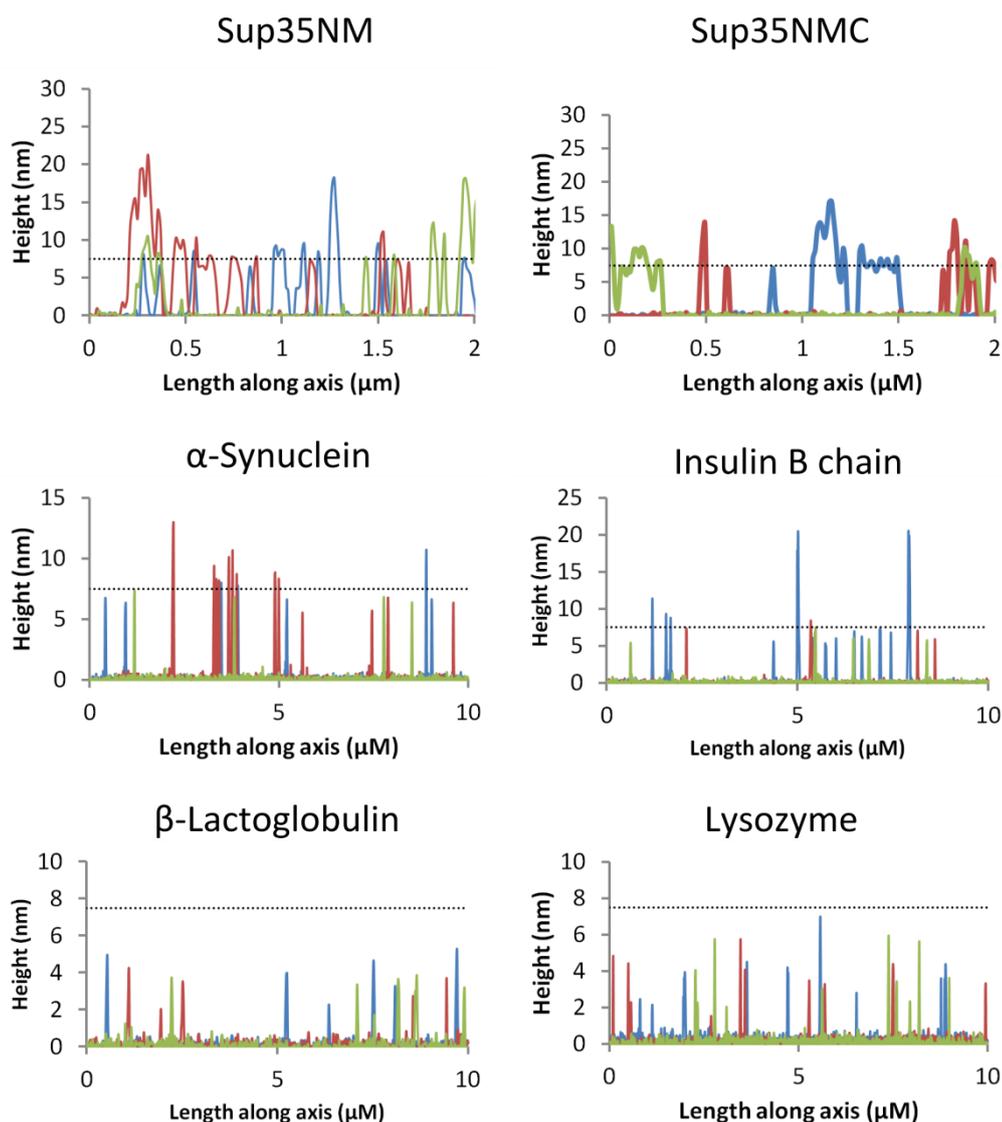
Figure 3.29   Cross Section Analysis of Fibril Heights

For Sup35NM and Sup35NMC 3 x 2µm cross sections were selected to analyse the edges of the fibrillar aggregates where the most unclumped examples exist. For the other proteins 3 x 10µm cross sections were taken. To each graph a dotted line at 7.5nm was added to show the differences in fibril samples.

B chain, α-synuclein, β-lactoglobulin and lysozyme allowed 10µm cross section areas to be

examined. The data shows that α-synuclein and insulin B chain are slightly thinner than

Sup35 fibrils with heights of 7nm and 5-6nm respectively and β-lactoglobulin and lysozyme

fibrils are the thinnest with heights from 2-4nm (Fig. 3.29).

3.2.3.2   Quantitative Analysis

Quantitative analysis of fibril properties such as contour length, height and persistence

length can be calculated by transfer of AFM images to a MatLab application, Tracey (Xue,

unpublished).

The MatLab application locates and traces the fibrils by first finding the ends of fibrils and then connecting them following predefined parameters such as approximate fibril width in pixels and an angle (θ) which defines how 'wavy' the fibrils are such that the application can differentiate between overlaps/breakages of fibrils (Fig. 3.30). These values are changed between fibrils of different proteins but are fixed for samples derived from the same protein. The application traces the particles and allows particle information such as height along the particle and contour length to be extracted.
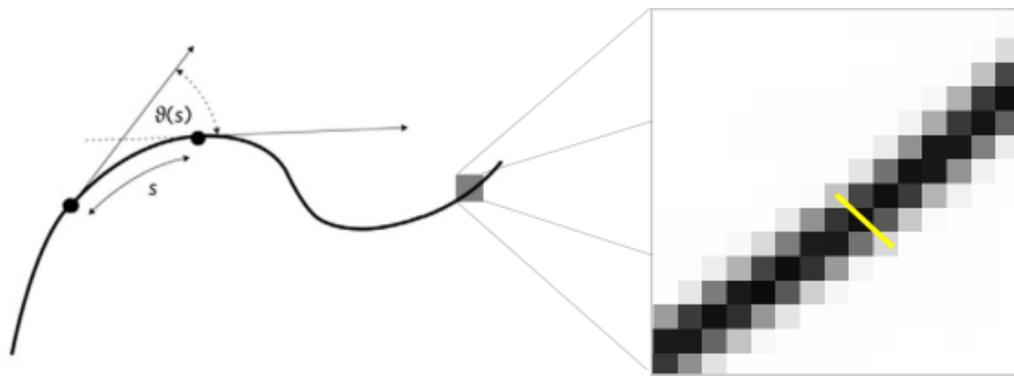
Figure 3.30   MatLab Input Parameters
The required parameters θ and fibril width in pixels can be adjudged by inspection of AFM images and trial and error analysis of surfaces using the application to see which conditions give the best results. θ is the angle of bend allowed in the fibril, see left , whilst the fibril width in pixels can be judged by zooming in on a fibril at the resolution scanned, see zoom in section and yellow line. Adapted from Adamcik and Mezzenga (2012) (Adamcik and Mezzenga, 2012)

The height of the amyloid fibrils is a better representation of the diameter of amyloid fibrils than their width due to the tip radius of the probe making the observed width larger than the actual diameter. The contour length of a polymer is the length from one terminus to the other following the path of the polymer. A polymer of contour length L can be divided into smaller rigid units, the length of which is the persistence length, L_p. A smaller value for L_p indicates a less stiff species due to it consisting of more, small rigid species rather than less, larger ones. The L_p is therefore a quantifiable property of a polymer used to describe its stiffness. The persistence length L_p can be calculated in several ways (Equation 1). The top equation is a probability based method for defining L_p where as the middle equation

determines the L_p by looking for the point on a fibre when correlation between two tangents starts to diminish and the lower equation utilises the end to end distance of individual particles to generate the L_p. The L_p is calculated here using a MatLab script (Xue, unpublished) which utilises the first equation (Equation 3.1).

Equation 3.1 Determination of Fibril Persistence Length

$$P\left(\theta_{2D}(L)\right) = \sqrt{\frac{L_p}{2\pi L}} e^{-\frac{L_p\theta_{2D}^2}{2L}}$$

$$\langle\cos(\theta_{2D})\rangle = e^{-\frac{L}{2L_p}}$$

$$\langle R_{2D}^2\rangle = 4L_pL\left[1 - \frac{2L_p}{L}\left(1 - e^{-\frac{L}{2L_p}}\right)\right]$$

Multiple images, processed to remove image curvature and streaking, per protein were traced using MatLab and the particles from each concatenated before assessment of variables such as longest fibril, weight average contour length (which is an alternative to the arithmetic mean which places a larger emphasis on longer particles, $\frac{\sum(CL^2)}{(\sum CL)^2}$), weight average contour length standard error of mean (SEM), mean height and mean height standard deviation (Table 1). The height and contour length data is in agreement with the qualitative assessment of their dimensions. The persistence length quantifies the stiffness of the particles and therefore the 'waviness' observed on the mica surface and shows that β-lactoglobulin forms the most rigid polymers followed by insulin and then α-synuclein and lysozyme, whose similar persistence lengths make them the most flexible polymers. This stiffness information is in agreement with visual inspection of the morphology of particles observed on the mica surface.

Table 3.1   **Quantitative Comparison of Fibril Dimensions/Properties***

Protein	Longest Fibril (nm)	Weight Average Contour Length (nm)	Weight Average Contour Length SEM (nm)	Mean Height (nm)	Mean Height SD (nm)	L_p (µm)	N_p
α-Synuclein	7444.10	1862.10	56.02	6.94	2.42	1.86	1105
β-Lactoglobulin	3806.40	1117.80	23.97	2.49	1.18	5.12	2175
Lysozyme	10970.00	2757.40	93.38	3.38	1.91	1.30	872
Insulin B chain	15983.00	5937.30	488.04	10.74	21.05	2.48	148

*Table showing fibril dimensions and persistence length from a number of traced fibrils, N_p.

3.3   Discussion and Conclusion

Despite the assumption that fibril formation is potentially a facet of all proteins the development of *in vitro* amyloid systems is complicated from synthesis and purification to conversion into fibrils. Where the conversion from monomer to fibril occurs under mild conditions the generation and purification of these species <u>can</u> be complicated by their aggregation. The expression and purification of α-synuclein, which although polymerises at the same temperature as the protein expression, gives rise to a large amount of pure, monomeric protein. In contrast to α-synuclein the purification of Sup35NM and Sup35NMC is complicated by a large amount of polymerisation during expression. The purification of Sup35NM requires denaturing conditions to extract and purify the protein from the cellular residues but the subsequent fibril formation must also tolerate the protein coming from a denatured purification, which for Sup35NM is the case. The C domain of Sup35NMC and its inability to refold post denaturation, appears, to seriously retard its ability to form fibrils. Native purification enables, albeit in a lower yield, isolation of material susceptible to polymerisation and although a less subtle observation than the mutants observed to affect fibril formation by Kabani *et al* (Kabani et al., 2011) the folding of the C-terminal domain does indeed have a significant effect on the propensity of the protein to form fibrils.

Insulin

GIVEQCCTSICSLYQLENYCN

FVNQHLCGSHLVEALYLVCGERGFFYTPKA

Lysozyme

KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDGSTDYGILQINS
RWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIVSDGNGMNAWVAWRNCKGTDV
QAWIRGCRL

β-Lactoglobulin

LIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQSAPLRVYVEELKPTPEGDLEILLQK
WENGECAQKKIIAEKTKIPAVFKIDALNENKVLVLDTDYKKYLLFCMENSAEPEQSLACQ
CLVRTPEVDDEALEKFDKALKALPMHIRLSFNPTQLEEQCHI

Figure 3.31   Peptides at Core of Amyloid Fibrils Formed by Heating in Acid

Top – Bovine Insulin A (top) and B (bottom) with regions protected from hydrolysis in green. Middle – Lysozyme sequence with peptide found at the core of the amyloid fibrils marked in yellow and the area around which observed in the fibrils marked in purple. Bottom – β-Lactoglobulin sequence with one of the regions at the core of the fibril marked in blue.

Whilst the conversion of monomer to fibril under mild conditions offers difficulties with purification pre-assembly, proteins which seem to require extra activation to polymerise offer the potential of easier purification but the problem of working out the conditions required for fibrils to form. The hydrolytic mechanism of fibril formation has recently been shown with insulin (Piejko et al., 2015), β-lactoglobulin (Akkermans et al., 2008) and lysozyme (Lara et al., 2011a, 2013, 2014) (Fig. 3.31).

As proteins are a chain of amino acids covalently bonded together by amide groups the acidic conditions utilised promotes non-specific hydrolysis (Fig. 3.32) and therefore potentially leads to a highly heterogeneous monomer set.

Although the hydrolytic mechanism of fibril formation is tempting it is also worth contemplating the protective nature of the cross-beta sheet core (Fig. 1.12) and is the final monomer unit of a fibril the same as the monomer required for fibril formation? The hydrolysis of one or two amide bonds may well be enough to allow monomers to enter conformations not normally allowed and therefore form fibrils. The cross-beta sheet region formed is protected from hydrolysis but leaves the exposed fibril 'side-chains'

susceptible to further acidic hydrolysis. The proposed hydrolytic mechanism when proteins are subjected to heat in an acidic environment means that the fibrils formed in this process are derived from the protein and not fibrils of the protein.
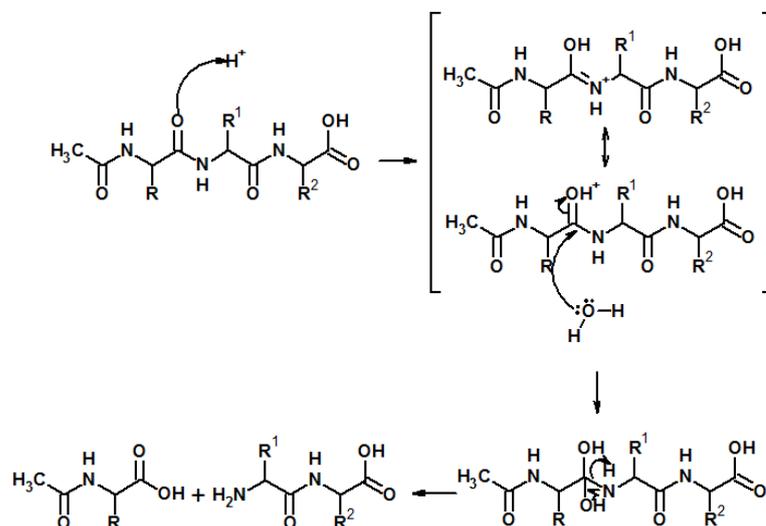
Figure 3.32   Reaction Mechanism Showing the Acid Catalysed Hydrolysis of a Peptide bond

The acid catalysed addition of water to an amide bond with subsequent elimination of an amine molecule is shown. The brackets denoting the interconversion of two tautomeric forms of the protonated amide bond before addition of water. The amine is then eliminated leaving a carboxylic acid.

The conversion of β-lactoglobulin, lysozyme and insulin B chain to fibrils can also be achieved *in vitro* using denaturing conditions and therefore potentially without hydrolysis: β-lactoglobulin (Hamada and Dobson, 2002), lysozyme (Krebs et al., 2000) and insulin (Nielsen et al., 2001). The tertiary structure of a protein protects it from hydrolysis as in its folded form some residues are not exposed to water and therefore protected. Whilst denaturing conditions are not enough to break standard amide bonds the removal of protein tertiary structure exposes residues more susceptible to hydrolysis to water and therefore potentially increases hydrolysis. Work by Vlad *et al* (Vlad et al., 2011) has shown that α-synuclein has autoproteolytic activity which can liberate peptides, 71-140, from the amyloid forming region (Uéda et al., 1993), 61-95, whose reaction kinetics are enhanced when compared to the full length protein. Although no evidence for this truncation was

observed here it offers evidence that hydrolytic enhancement of amyloid assembly is a more common phenomenon than just for fibrils formed in an acidic environment.

Whilst the evidence for the role of hydrolysis, whether chemical or enzymatic, in the preparation of *in vitro* amyloid fibrils derived from β-lactoglobulin (Akkermans et al., 2008), insulin (Piejko et al., 2015), lysozyme (Frare et al., 2004, 2006; Lara et al., 2014), α-synuclein (Vlad et al., 2011) is apparent its importance *in vivo* is unclear. The identification of fibrillar materials *in vivo* tends to be via histological staining of tissue but the proteomic analysis of these samples can enlighten what the monomeric units of the fibrils are. *Ex vivo* analysis of lysozyme fibrils, by N-terminal sequencing, derived from patients suffering from systemic lysozyme amyloidosis has shown that the monomers are mutated full length lysozyme (Booth et al., 1997; Pepys et al., 1993) (Fig. 3.33).

```
HEWL        1    MRSLLILVLCFLPLAALGKVFGRCELAAAMKRHGLDNYRGYSLGNWVCAA    50
                               |||.|||||..:||.|:|.|||.||.||:|.|
Human Lys 1    ----------------KVFERCELARTLKRLGMDGYRGISLANWMCLA      32

HEWL       51    KFESNFNTQATNRNT-DGSTDYGILQINSRWWCNDGRTPGSRNLCNIPCS    99
                  |:||.:||:|||.|. |.|||||..|||||:|||||:|||:.|.|::.||
Human Lys 33   KWESGYNTRATNYNAGDRSTDYGTFQINSRYWCNDGKTPGAVNACHLSCS    82

HEWL      100    ALLSSDITASVNCAKKIVSDGNGMNAWVAWRNRCKGTDVQAWIRGCRL    147
                  |||..:|..:|.|||::|.|..|:.|||||||||:..||:.:::||.:
Human Lys 83   ALLQDNIADAVACAKRVVRDPQGIRAWVAWRNRCQNRDVRQYVQGCGV    130
```

Figure 3.33   Comparison of the human and hen egg white lysozyme sequences.
A sequence alignment of hen egg white lysozyme (HEWL, Top) and human lysozyme (Lower). The HEWL sequence labelled with the ILQINS sequence determined by Lara et al (Lara) in red and the positions susceptible to acidic hydrolysis as shown by Frare et al (Frare) in green. The residues of human lysozyme whose mutations are central to conversion to amyloid fibrils without hydrolysi in vivo by Pepys et al (Pepys) are highlighted in yellow.

Frare et al (Frare et al., 2004, 2006) have shown that wild type human and hen egg lysozyme do not enter the conformations required for amyloid formation without hydrolysis where as the mutated forms found in systemic amyloidosis are capable of entering this conformation and therefore converting to the amyloid state. Insulin derived amyloidosis is observed at the injection sites of recombinant insulin. Storage of the insulin formulations over time have shown that the protein structure is the subject of amide

hydrolysis, both side chain and peptide backbone (Brange et al., 1992; Nilsson and Dobson, 2003).

Beyond the proteins studied here there is evidence that hydrolysis can play a role in the aggregation of other proteins such as the S52P mutant of transthyretin (TTR) (Bergström et al., 2005; Mangione et al., 2013) and β2-microlglobulin (Heegaard et al., 2005; Linke et al., 1989; Stoppini and Bellotti, 2015).

Aside from the differences observed in fibril formation conditions the dimensions and properties of the fibrils varies despite the same core cross-beta sheet. The Sup35 proteins are the largest in terms of monomer size at 35kDa and 78kDa for Sup35NM and Sup35NMC respectively but by qualitative analysis form fibrils of a similar diameter 7-8.5nm. In comparison insulin B chain is the smallest protein investigated at 5.8kDa and yet gives rise to fibrils of comparable height with a mean traced height of 10.74nm and a particle height mean of 8.18nm, both values appear high when compared to the qualitative assessment, 5-6nm, and this is due to the presence of some very high material skewing calculations. α-synuclein, β-lactoglobulin and lysozyme are all similar molecular weights as protein monomer, ~15kDa, yet give rise to fibrils of different height at 6.94nm, 2.49 and 3.38nm respectively. This suggests that there is no association between monomer size and fibril diameter but this assessment is based on the starting monomers size and does not take into account hydrolysis observed during conversion or fibrils propensity to laterally associate.

The quantitative assessment of the longer fibril particles is in agreement with the qualitative assessment of their length and the length trend concurs with the order of fibrils weight average contour length.

Finally, characterisation of the polymers in terms of there nanomechanical properties was carried out. The elasticity of a material, its non permament deformation, is quantified by

the Young's modulus (Pa) of the material, the smaller the value the more elastic the polymer. The persistence length (L_p) of a polymer is the length where the correlation between two tangent angles starts to diminish and is a measure of the rigidity of a polymer. The relationship between Young's modulus and persistence length is defined by Equation 3.2.

Equation 3.2 The correlation of Young's modulus and L_p

$$ L_p = \frac{EI}{\kappa_B T} $$

E= Young's modulus, I= Moment of Inertia of the particle cross section (for rigid rods of diameter R, I = $1/4\pi$R), κ_B=Boltzmann constant and T = temperature (K).

Calculation of the fibril persistence length of each type of fibril in this chapter shows that L_p is not correlated with the diameter of the fibrils, as shown by the thinnest fibril, β-lactoglobulin, being the most rigid and the widest, α-synuclein, having an intermediate persistence length. The size of the regions shown to be at the core of the amyloid structures shows that β-lactoglobulin has the largest region, ~30 residues (Fig. 3.31) (Akkermans et al., 2008), whilst lysozyme (Lara et al., 2014) and insulin B chain (Piejko et al., 2015) have the smallest, ~15 residues roughly following the order of decreasing L_p. Is it possible that the number of inter-molecular interactions between monomers is key to the rigidity of the resulting polymer? The observed L_p for a-synuclein and this trend suggests the size of the cross-beta sheet core would be similar/slightly smaller than insulin B chain. The core 35 residue sequence for α-synuclein is known, NAC (Uéda et al., 1993), and the work of Vlad *et al* (Vlad et al., 2011) shows that the actual core maybe slightly smaller, ~25 residues, which suggests its L_p should be in the region of β-lactoglobulin rather than insulin. This comparison though assumes that there is not further redundancy in the 25 residue cross-beta sheet region observed for α-synuclein. For Sup35 the N domain is the amyloid forming region and is 254 residues in size but the importance of 16 residues, 8-24, on

phenotypes which promote and prevent fibril formation (ASU/PNM) suggests that this region is key to the cross-beta sheet (DePace et al., 1998).

Table 3.2   Published persistence length data for the fibres generated in chapter 3

Protein	Subtype	Height	L_p	No. of particles	Reference
β-Lactoglobulin	i	2nm	0.968µm	50	(Adamcik et al., 2010)
β-Lactoglobulin	ii	3.5-4nm	1.62µm	108	(Adamcik et al., 2010)
β-Lactoglobulin	iii	5.5-6.4nm	3.244µm	60	(Adamcik et al., 2010)
β-Lactoglobulin		2.3nm	3.818µm	110	(Schleeger et al., 2013)
α-Synuclein	Curly	<17nm[†]	0.17µm	50	(Bhak et al., 2010)
α-Synuclein	Straight	<17nm[†]	140µm	50	(Bhak et al., 2010)
Lysozyme		2.4nm	2.5µm	66	(Lara et al., 2011b)
Lysozyme peptide*		13nm	5.57µm	-	(Lara et al., 2014)
Bovine Insulin		5-6nm[ᵶ]	42±30µm	76	(Smith et al., 2006)

[†] Values extracted from TEM images in paper. [ᵶ] Values extracted from the height scale of AFM images.

*Fibrils derived from the synthetic ILQINS peptide.

The persistence length data previously published for fibrils of β-lactoglobulin shows that the entwining of fibrils together into higher structures makes a significant difference to the rigidity of polymers (Table 3.2). By associating two fibres together a doubling of the persistence length is observed and when three fibres interact the persistence lengths trebles (Adamcik et al., 2010). Whilst the height of the single fibres observed in (Adamcik et al., 2010) are similar to the heights of β-lactoglobulin fibrils generated in chapter 3 the calculated persistence lengths differs greatly, as do the values observed for α-synuclein (Bhak et al., 2010) and insulin (Smith et al., 2006). Whilst those fibres differ from the previous literature, the data generated in chapter 3 is in good agreement with the previously published data for lysozyme (Lara et al., 2011b) and that of a separate publication regarding β-lactoglobulin fibrils (Schleeger et al., 2013). In all cases the methodology used in chapter 3 has enabled the persistence length data to be generated on much larger populations of fibrils than has previously been possible which is of the utmost importance for generating accurate values of the persistence length.

The (Adamcik et al., 2010) data although apparently differing in rigidity to the data shown here does show that the structural organisation of fibres has a large influence on the persistence length of materials and suggests that height does play some role in determining persistence length.
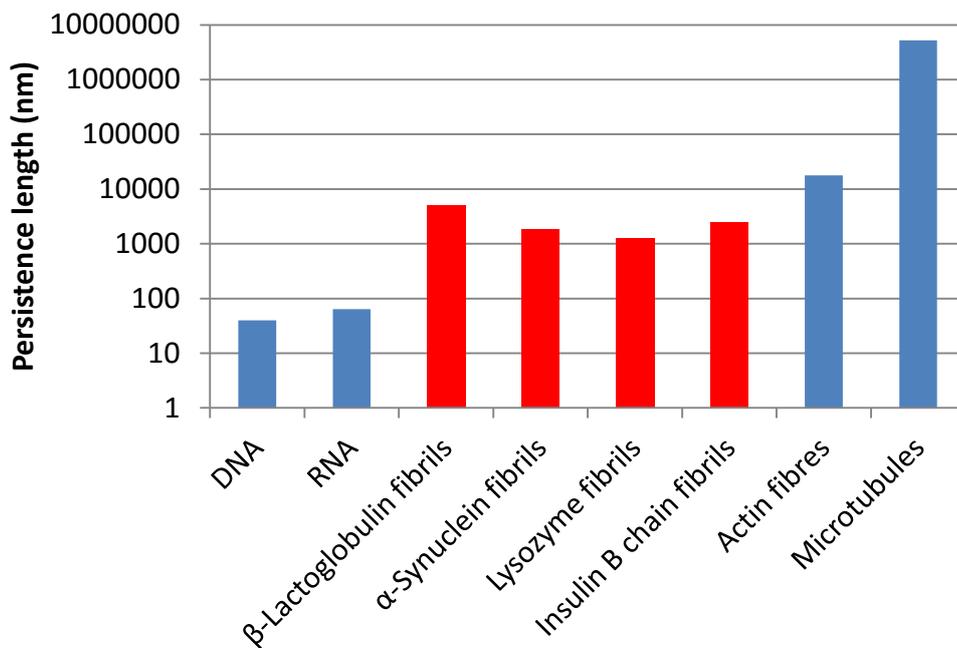
Figure 3.34 **Bar chart showing a comparison of the fibrils formed in this chapter with other biological polymers**
A bar chart with a log plot showing published persistence length values (blue) for DNA (Hansma et al., 1996), RNA (Abels et al., 2005), actin fibres (Gittes et al., 1993) and microtubules (Gittes et al., 1993) compared to the persistence length values extracted from the fibrils of lysozyme, β-lactoglobulin, α-synuclein and insulin B chain.

Published analysis of apoliprotein-C II fibrils shows particles of a height similar to β-lactoglobulin and lysozyme but have a significantly lower persistence length, 36nm (Hatters et al., 2003). X-ray diffraction analysis of these particles showed that the core structure is that of a single protofilament with no further structural organisation (Hatters et al., 2003). Whilst (Adamcik et al., 2010) show that β-lactoglobulin is susceptible to inter-fibril aggregation leading to differing persistence lengths, the particles generated in chapter 3 appear to be, by comparison of the particle height with published data, protofilaments with

a much higher rigidity. β-Lactoglobulin has been shown to form two distinct morphologies of fibrils: worm like and straight with low (92.7nm) and high (3818nm) persistence length respectively, under different fibril formation conditions (VandenAkker et al., 2011). The difference between worm like and straight fibrils derived from β-lactoglobulin was shown by vibrational sum frequency spectroscopy to be correlated with the the proportion of beta-sheet observed in the samples (Schleeger et al., 2013; VandenAkker et al., 2011). The particles formed in this research have similar height (2.3nm) and persistence length values (3818nm) to the values generated here and are that of a single protofilament (VandenAkker et al., 2011). These data suggest that the nanomechanical properties of amyloid fibrils are profoundly influenced by the molecular structure of the fibres.

Although the values for L_p in Table 3.1 show that the L_p of the amyloid fibrils generated in this chapter are of different lengths, comparison to other biologically relevant polymers such as DNA, RNA, actin and microtubules the amyloid structures of lysozyme, β-lactoglobulin, α-synuclein and insulin B chain exhibit similar L_p values to each other (Fig. 3.34, red bars). The fibrils formed here are much more rigid than the polynucleotides DNA and RNA (Fig. 3.34, left hand blue bars) (Abels et al., 2005; Hansma et al., 1996) but more flexible than actin polymers and microtubules (Fig. 3.34, right hand blue bars) (Gittes et al., 1993).

To conclude expression and purification systems have been developed for the generation of α-synuclein, Sup35NM and Sup35NMC monomers. Utilising these monomers as well as commercial sources of β-lactoglobulin, hen egg white lysozyme and bovine insulin B-chain an array of different fibrils have been generated, characterised in terms of their physical dimensions/properties. Comparison of this data leads to interesting questions relating the structure of the cross-beta sheet and polymer properties such as L_p. These fibril stocks will allow the investigation of their resistance to mechanical stress (Chapter 4).

3.4 Future Work

To develop *in vitro* polymerisation conditions which do not result in the aggregation of

Sup35 amyloid fibrils into large clumps.

4

Comparative Analysis of the Fragmentation of Amyloid Fibrils

4.1   Introduction

The formation of a nucleus from protein monomers and subsequent growth of amyloid fibrils by monomer addition are pivotal events in the nucleated growth mechanism of amyloid fibrils. However, the kinetics of fibril formation are profoundly influenced by secondary processes such as surface nucleation and fibril fragmentation. Without these secondary processes, the exponential phase of growth observed would be considerably slower.

The influence of amyloid fibril fragmentation extends beyond the kinetics of fibril formation into their functional or deleterious roles in biological systems. As discussed in 1.8.2 the chaperone mediated fragmentation of amyloid deposits is an integral component of the transmission mechanism associated with the prion phenotype (Byrne et al., 2009; Cox et al., 2003). In yeast, this can provide an adaptive mechanism, for example the utilisation of different nitrogen sources, but in mammalian organisms so far only gives rise to disease association in for example CJD and Kuru. A key observation associated with yeast prions, [$prion^+$], is the presence of different phenotypes, weak and strong, caused by the formation of different fibril morphologies with different propagation propensities (Tanaka et al., 2006). The central role that fragmentation has on prion propagation and the potential differences of the morphologies with respect to their resistance to fragmentation poses an important question: Does a "weaker" morphology of fibrils, more susceptible to breakage, give rise to a stronger prion phenotype? (Tanaka et al., 2006)

The fragmentation of fibrils also offers a potential explanation to the lack of correlation between amyloid load and the symptoms observed with sufferers of diseases such as Alzheimer's. MTT dye release assays have been used to demonstrate that the smaller fibril particles generated by fragmentation interact with membranes more efficiently and increase cytotoxicity compared to the unfragmented fibrils (Fig. 4.1) (Xue et al., 2009b).

Recent work using the same β_2-Microglobulin (B2m) fibril samples has shown that fragmented fibrils are internalised into the cell via endocytosis more efficiently than unfragmented longer particles, and the cytotoxicity observed is due to interaction of the particles with lysozomes and impairment of their function (Jakhria et al., 2014).
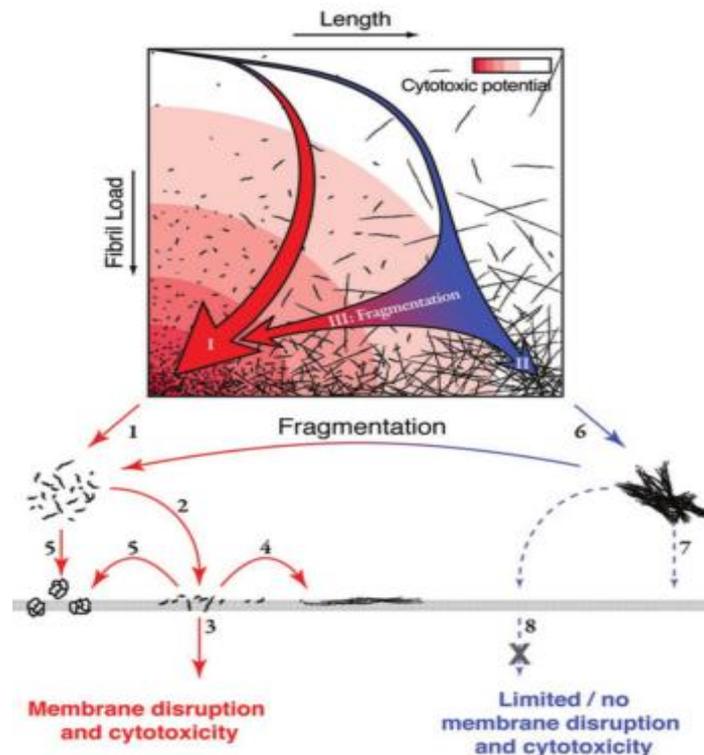
Figure 4.1   Fibril Fragmentation Increases Cytotoxicity
A schematic showing that the fragmentation of amyloid fibrils increases the cytotoxicity of the particle by allowing greater interaction with, and disruption, of cell membranes. The red arrow shows a fibril formation where fragmentation is an active process leading to shorter particles. The blue arrow is an example of a fibril formation where fibril fragmentation is minimal leading to longer particles. Fragmentation of these longer particles gives rise to an increase in cytotoxicity. Taken from (Xue et al., 2009b)

As previously discussed, the properties of amyloid fibrils have marked them out as potential nanomaterials (Knowles and Buehler, 2011). By correlating structural characteristics and physical properties with the composition of the material it is possible to build up a structure activity relationship, and generate rules that will enable the *in silico* design of new amyloid materials with enhanced properties (Li et al., 2014; Scheibel et al., 2003). The formation of fibrils of different morphologes from the same monomers can

generate different properties, as phenotypically observed with prions and shown *in vitro* with β-lactoglobulin (VandenAkker et al., 2011), suggesting that fibril formation conditions are also important in generating different amyloid types, which could be differentiated between by examination of their resistance to fragmentation.   Thus, the ease of fragmentation of a fibril can potentially be correlated to their yield strength and/or rigidity.
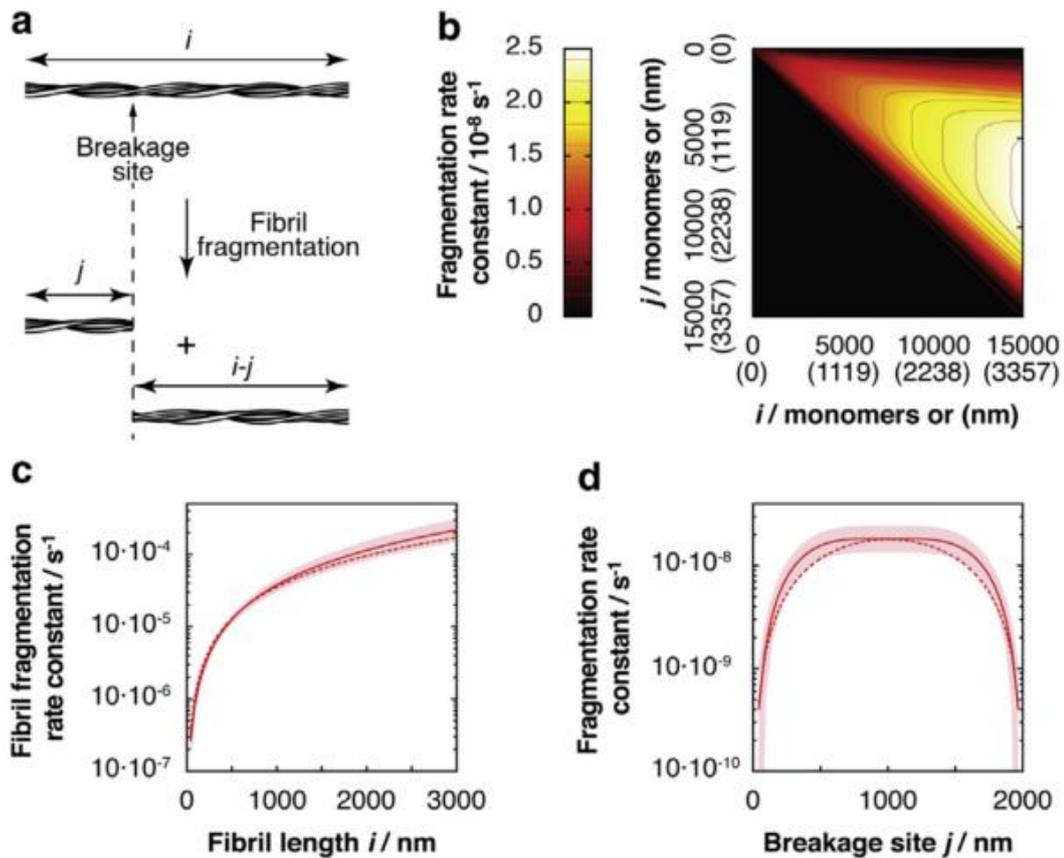
Figure 4.2   Best fit Fragmentation Model for β2-microglobulin.
The breakage of amyloid fibrils into smaller components could potentially occur in between monomers at any point on the fibril, length i, to give a fragment, length j (a). The first order fragmentation rates of fibril breakage derived from best fit analysis of β2-microglobulin fragmentation which shows that breakage is length dependent with longer fibrils more likely to break than short (b). (c) A graph showing the change in fragmentation rate of β2-microglobulin fibrils as a function of fibril length. (d) A graph showing the rate of fragmentation as a function of position on the fibril with the middle exhibiting the faster rate. Figure taken from (Xue and Radford, 2013).

The fragmentation of β_2-microglobulin fibrils under mechanical perturbation has previously been investigated by analysis of the particle dimensions for the fibril population over the course of the fragmentation. This was achieved by using the data from large numbers of

individually imaged fibrils using Tapping Mode Atomic Force Microscopy (TM-AFM) and an automated MatLab application which is capable of tracing fibrils in the images and extracting information about them (Xue, unpublished). Applying different periods of fragmentation, by stirring, showed how these materials fracture over time (Xue and Radford, 2013; Xue et al., 2009b). The change in contour length distribution over time enabled the generation of a model which best fits the contour length distribution profile observed experimentally when β2-microglobulin is fragmented, in which the rate of breakage for longer fibrils is faster than short and that it is more likely for the breakage occur in the centre of a fibril rather than the end of amyloid fibrils (Fig. 4.2)(Xue and Radford, 2013).

4.1.1   Chapter Aims

It is evident that fibril fragmentation and particle size are important considerations for understanding the biological function of amyloid fibrils and their future development as nanomaterials. Therefore, the development of methods to quantify amyloid fibrils resistance to fragmentation is important. In this chapter, the resistance of amyloid fibrils to mechanical stress for fibrils derived from different proteins, developed in chapter 3, will be assessed using the physical fragmentation assay described previously (Xue and Radford, 2013; Xue et al., 2009b) and further developed here. The data captured here will enable the comparison of amyloid fibrils' stability towards fragmentation, and show fibrils derived from different monomers have altered physical strengths despite the same core cross-beta sheet architecture. Results here will enable the fibril fragmentation and fibrils mechanical characteristics to be mathematically modelled. Quantitative understanding of these properties will allow the prediction of fibril fragmentation for other assemblies and allow for *in silico* design of fibrils resistant to mechanical stress.

4.2 Results

To enable comparison between experimental studies the conditions used to fragment the fibrils generated in chapter 3 were identical to the work of Xue *et al* (Xue and Radford, 2013; Xue et al., 2009b). Briefly, 500µl of 120µM fibril solutions, in fibril forming buffer, were stirred at 1000rpm by a 3 x 8mm magnetic stirrer bar in a 1.5ml glass chromatography vial using an Ika Squid stirrer plate with a digital display (Fig. 4.3). The fragmentation experiments were always carried out in fibril forming buffer as recent evidence has shown that altering the composition of buffer and temperature has the potential to change or disaggregate the fibrils formed (Ikenoue et al., 2014; Tipping et al., 2015).
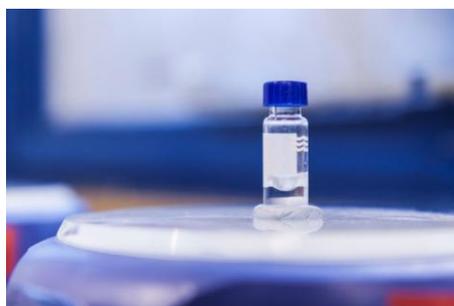
Figure 4.3   Experimental Setup of Fibril Fragmentation
A 1.5ml clear LC-MS vial containing a 3x 8mm PTFE stirrer bar and 0.5ml of a 120µM fibril solution were stirred at 1000rpm by an IKA Squid stirrer, the experiment was covered to avoid light damage of the sample. Image courtesy of Wei-Feng Xue.

The analysis of the particle size distributions were achieved by AFM image analysis.

At different time points aliquots of the fragmentation experiment were removed and diluted to concentrations previously shown to give good particle coverage, and incubated on a freshly cleaved mica surface before washing and drying under a stream of nitrogen gas. 20µm x 20µm scans at 2048 pixels x 2048 pixels resolution were captured and the images were flattened using the nanoscope analysis software, to remove any curvature of the image plane. The processed image files generated could then be exported as a text file for recapitulation in MatLab or transferred directly to the MatLab application (See 3.2.3.2).

An important consideration for the successful analysis of fibril length distribution is to find fibril deposition conditions which allow for the capture of appropriate AFM data that gives a high density of discrete particles where overlap is minimised. The aim of each timepoint is to generate a large number of particles, at least 500, across multiple images to minimise the potential that the data is not a true representation of the surface, and therefore the sample. Calculation and subsequent plotting of mean fibril lengths against fragmentation time visualises the fragmentation and allows for calculation of rates of fragmentation under these conditions.

4.2.1   Qualitative Assessment of Fibril Fragmentation

Quantitative analysis of fibril size distributions is dependent on image quality and requires a large number of discrete particles where overlap is minimised to enable successful tracing of the fibrils. The suitability of the fibril samples produced in chapter 3 to quantitative particle size analysis during fragmentation is investigated qualitatively in this section.

4.2.1.1   Sup35

As observed in 3.2.2.2 the polymerisation of Sup35NM and Sup35NMC gave rise to large networks of amyloid fibrils (Fig. 4.4, 1A and 1B), which were also visible by the optical microscope used for aligning the probe with the laser and the mica surface (Fig. 4.4, 1C). The concentration required for Sup35NM fibril assembly is in the region of 10µM and as such the targeted fragmentation experiment concentration of 120µM was not possible to reach without monomer or fibril concentration processes which further exacerbate aggregation problems. Also, fragmentation of these aggregates was not possible by stirring as it was not powerful enough to break the inter fibril interactions holding the network together but mild sonication, 20% amplitude for 4 minutes on ice with 5 second bursts, did allow for the network to be dispersed. After sonication, the large aggregates visible pre sonication were no longer visible using the optical microscope and had been broken into smaller clusters shown by AFM (Fig. 4.4, 2A, 2B and 2C).
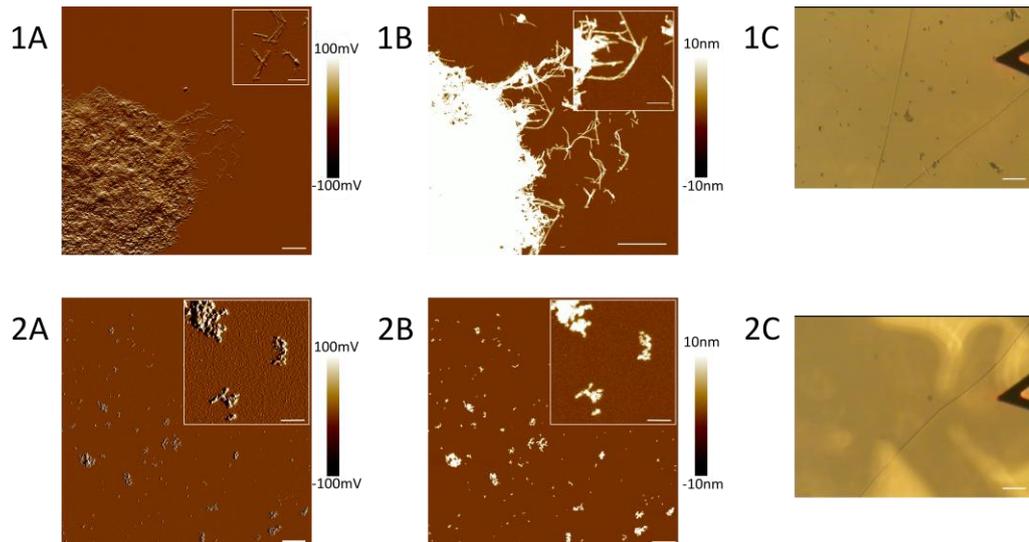
Figure 4.4   Analysis of Fragmentation of Sup35NM

1, 10uM Sup35NM Fibrils, unsonicated. 2, Sonication of 10uM Sup35NM fibrils on ice for 4 minutes with 5 second bursts at 20% amplitude A – Peak force error scan of 10μm x 10μm at 1024 x 1024 pixels with 1μm scale bar and 1μm insert magnified x 5 with 200nm scale bar. B – Height scan of the same image with 1μm scale bar and 1μm insert magnified x 5 with 200nm scale bar. C – Optical microscope image of the mica surface. The scale for the peak force error images are set between -100 and 100mV and the height scales are set between -10 and 10nm.

.

The large networks of fibrils at the beginning of the experiment and the material after sonication were not conducive to particle length analysis or fragmentation studies so were not progressed to quantitative size analysis.

4.2.1.2   Insulin B-chain

Exposure of fibrils derived from insulin B-chain, formed in 3.2.2.3, to fragmentation under the conditions used previously, and described in 4.2, resulted in a gradual loss of affinity of the particles to the mica surface (Fig. 4.5) which was not returned even with altered deposition conditions. Whilst the disappearance of the fibrillar material from the mica surface was obvious another change in mica surface coverage is less evident and involves the deposition of a large amount of small material generated during the fragmentation onto the mica surface (Fig. 5, inserts). This reproducible loss of affinity consistently occurred after about 9 hours of stirring and although the AFM background had changed TEM analysis showed the presence of fibrillar material in the fragmentation experiment still persisted (Fig. 4.6).
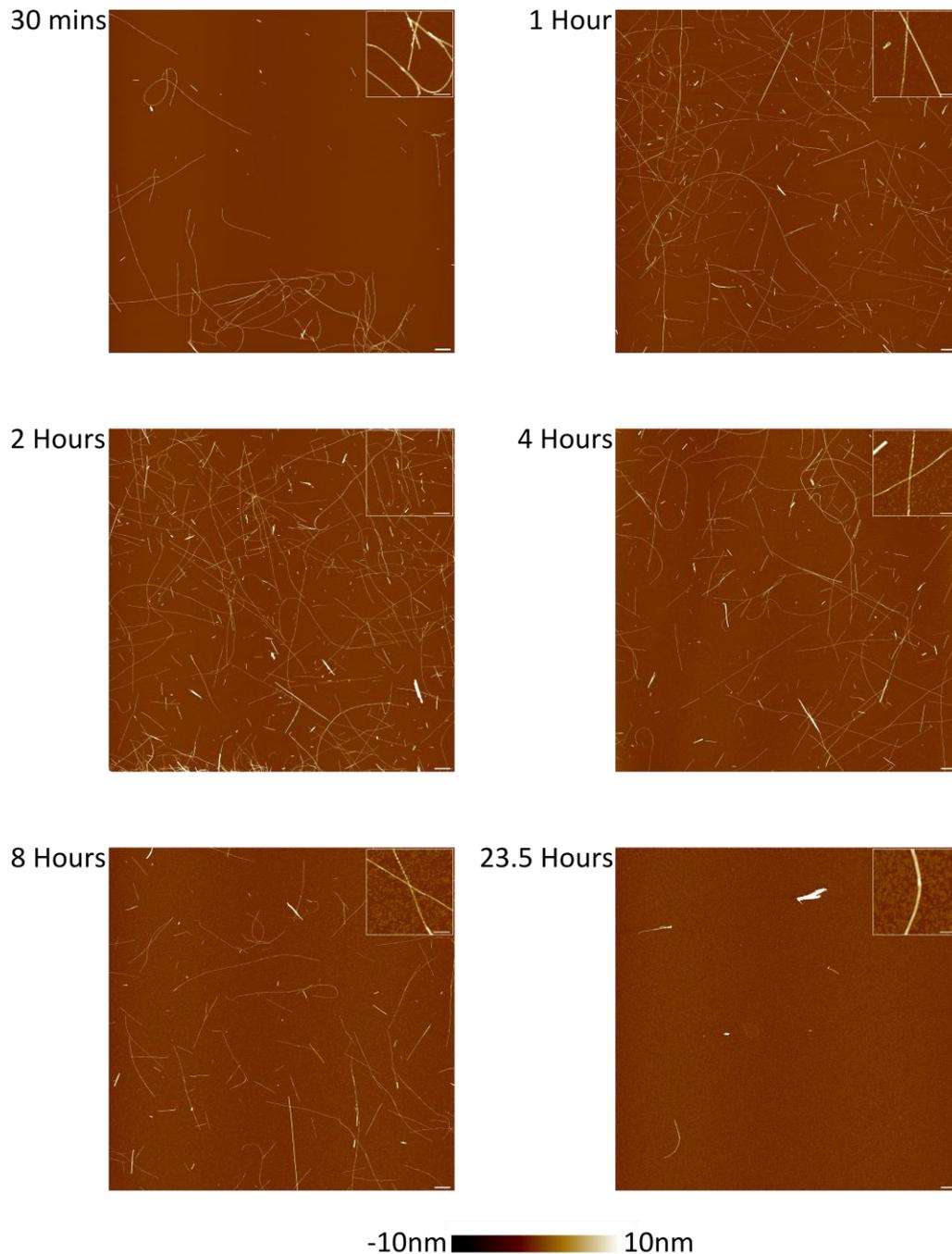
Figure 4.5 **AFM analysis of the Fragmentation of Insulin Fibrils**
A 120µM Insulin fibril suspension was stirred at 1000rpm and at different timepoints material was deposited onto a mica surface and a 20µm x 20µm area imaged at 2048 x 2048 resolution with a 1µm scale bar. The height scale for each image is set between -10 and 10nm.

There are several possibilities as to the origin of this difference. The fibrils derived from insulin B-chain fragmentation may occur by a different mechanism compared to other proteins, liberating very small 'lumps' of material. As the formation of insulin derived fibrils occurs at pH2 the fibrils are positively charged which aides binding to the negatively charged mica surface. One explanation for fibril disappearance was that the small clusters

formed interact with the fibrils stopping the fibrils from binding to the mica surface but this mechanism is unlikely due to the pH of solution meaning that, like the fibrils, the small clumps are protonated and therefore will electrostatically repulse the fibrils. A more probable explanation is that due to their positive charge the clumps and the fibrils are competing for binding to the negatively charged mica surface. As the image appears to show a covering of the surface with this material it suggests the smaller material is more tightly bound to the surface essentially covering it with positive charges, which would repel the fibrils from the surface due to electrostatic repulsion.

The inability to measure the height and length distributions of the insulin B-chain fibrils as the fragmentation proceeds meant the data generated was not suitable for particle size analysis during fibril fragmentation using this method.
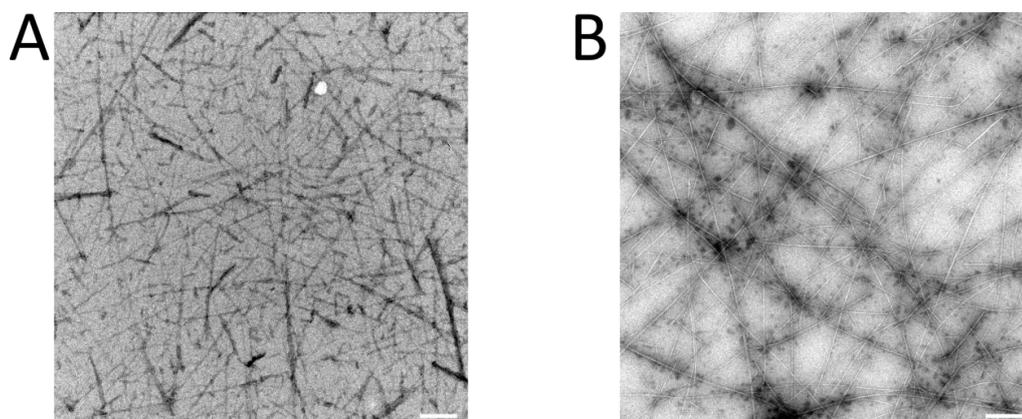
Figure 4.6   TEM Analysis of Fibrils Pre and Post Fragmentation
A – TEM analysis of Insulin fibrils before fragmentation, with 1µM scale bar. B – TEM analysis of Insulin fibrils fragmented until after the disappearance of fibrils from the AFM surface was observed, with 1µm scale bar.

4.2.1.3   Lysozyme, β-Lactoglobulin and α-Synuclein
The fragmentation of fibrils derived from hen egg lysozyme, β-lactoglobulin and α-synuclein were all achievable using the predefined conditions of 500µl of 120µM fibril solutions and using the appropriate deposition conditions gave rise to discrete particles observable by AFM (Fig. 4.7, 4.8, 4.9).
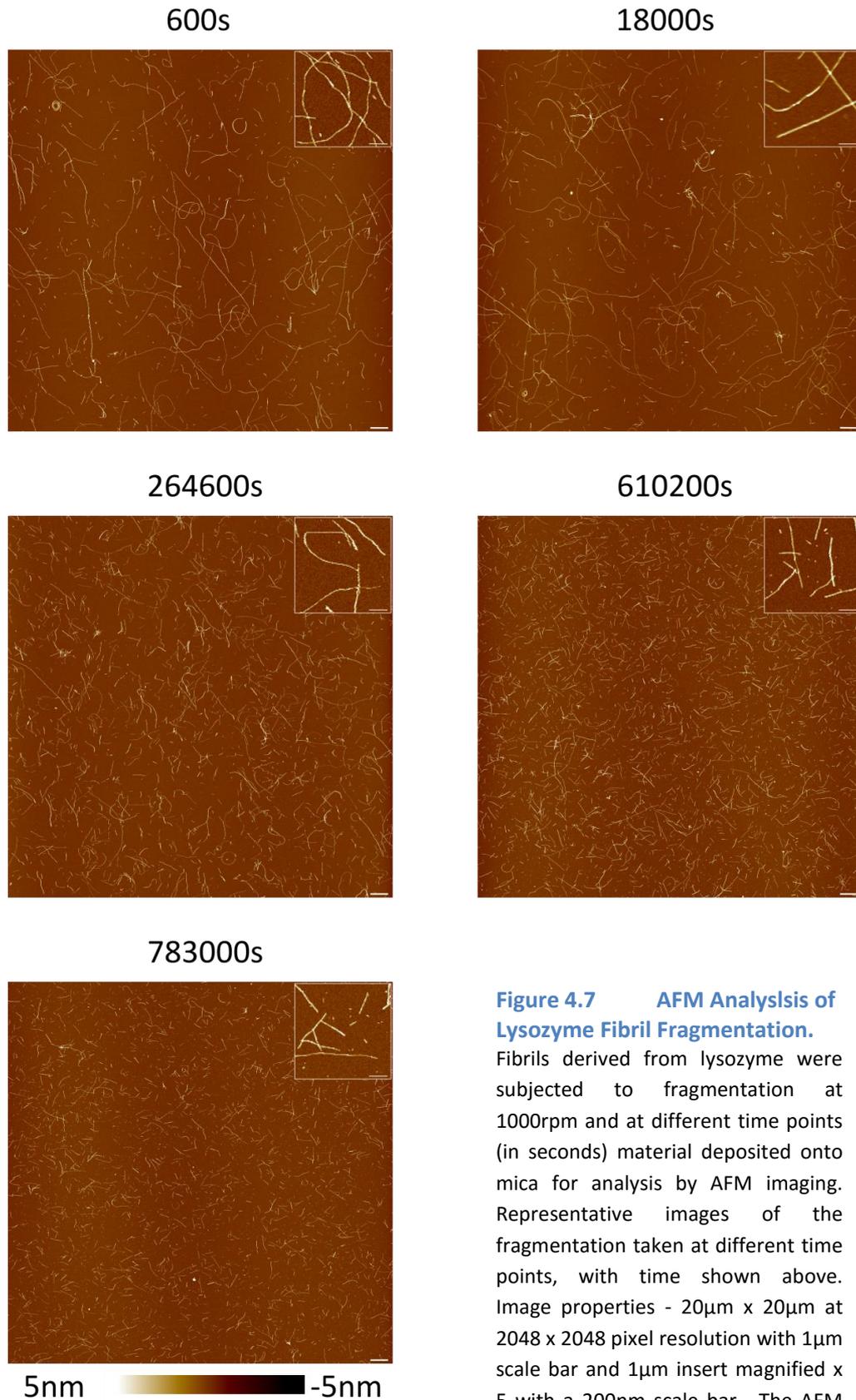
Figure 4.7 AFM Analyslsis of Lysozyme Fibril Fragmentation.

Fibrils derived from lysozyme were subjected to fragmentation at 1000rpm and at different time points (in seconds) material deposited onto mica for analysis by AFM imaging. Representative images of the fragmentation taken at different time points, with time shown above. Image properties - 20μm x 20μm at 2048 x 2048 pixel resolution with 1μm scale bar and 1μm insert magnified x 5 with a 200nm scale bar. The AFM height scale is set between -5 and 5nm.

300s

7200s

86400s

432000s

1128600s

5nm ▤ -5nm
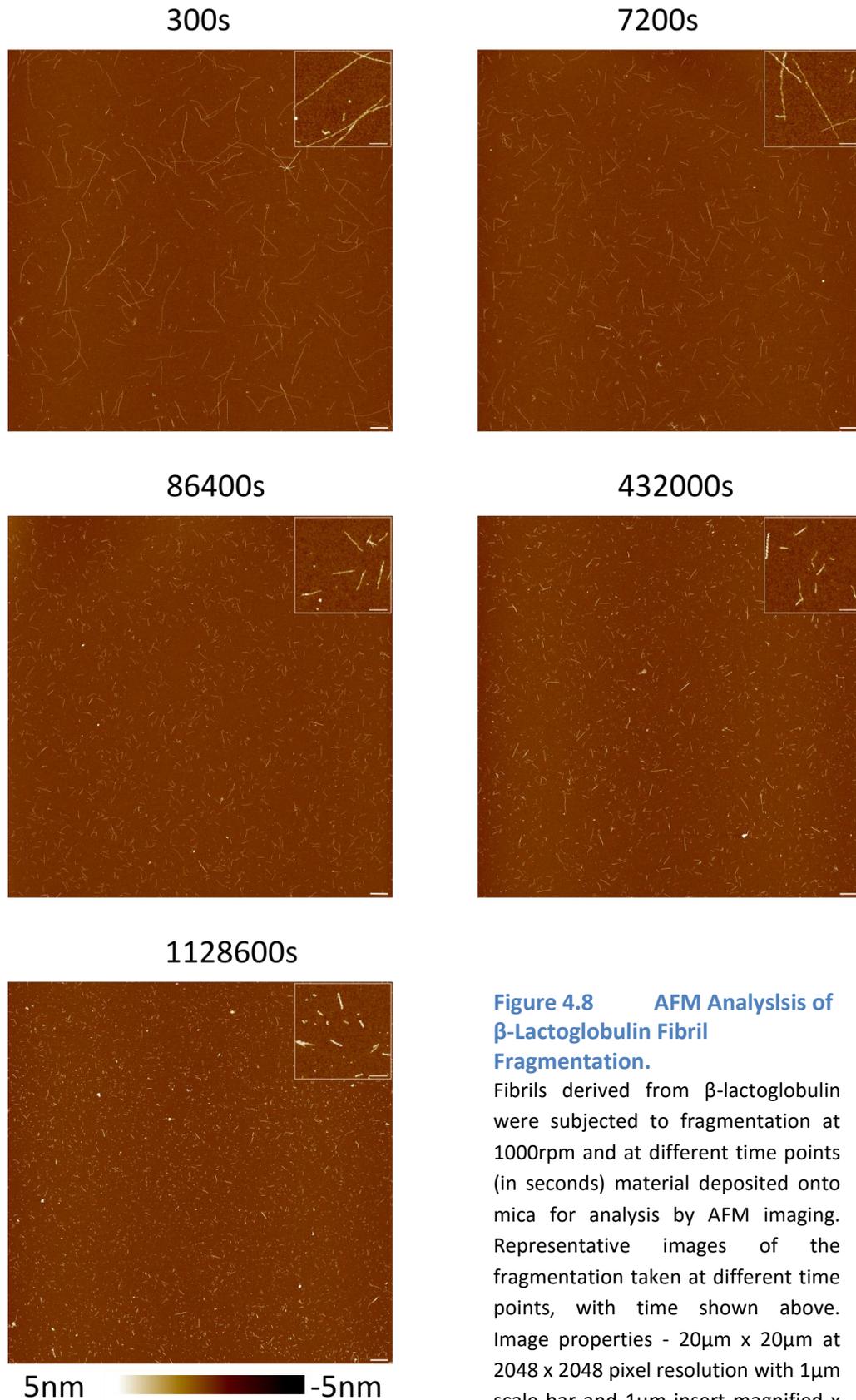
Figure 4.8   AFM Analyslsis of β-Lactoglobulin Fibril Fragmentation.

Fibrils derived from β-lactoglobulin were subjected to fragmentation at 1000rpm and at different time points (in seconds) material deposited onto mica for analysis by AFM imaging. Representative images of the fragmentation taken at different time points, with time shown above. Image properties - 20μm x 20μm at 2048 x 2048 pixel resolution with 1μm scale bar and 1μm insert magnified x 5 with a 200nm scale bar. The AFM height scale is set between -5 and 5nm.

300s

18900s

257100s

606300s

1293300s

10nm ▬▬▬▬ -10nm
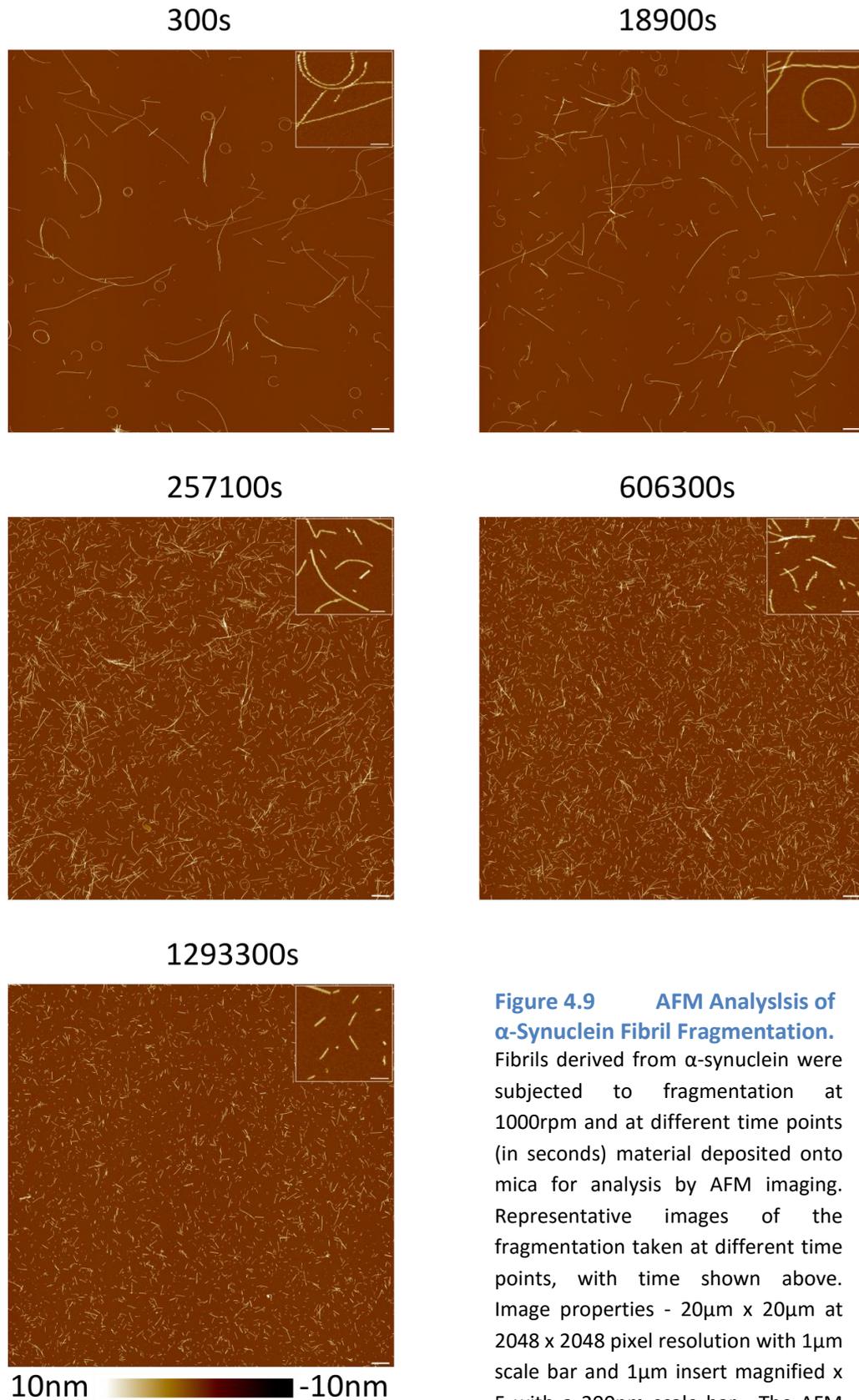
Figure 4.9   AFM Analyslsis of α-Synuclein Fibril Fragmentation. Fibrils derived from α-synuclein were subjected to fragmentation at 1000rpm and at different time points (in seconds) material deposited onto mica for analysis by AFM imaging. Representative images of the fragmentation taken at different time points, with time shown above. Image properties - 20µm x 20µm at 2048 x 2048 pixel resolution with 1µm scale bar and 1µm insert magnified x 5 with a 200nm scale bar. The AFM height scale is set between -10 and 10nm.

Qualitative analysis of the fibrils observed in the first time point of the fragmentation show that hen egg lysozyme and α-synuclein both have a propensity to form longer, more coiled fibrils where as β-lactoglobulin fibrils appear on the mica as shorter, straighter particles. When deposited onto mica unfragmented α-synuclein fibrils appear regularly as coiled structures (Fig. 9, 300s image), something that is observed with lysozyme rarely but not at all with β-lactoglobulin. Inspection of the individual AFM images captured during the fragmentation of these three proteins showed the each change from long fibrils at the start of the experiment to much smaller particles at the end of the experiments (Fig. 4.7, 4.8, 4.9).

Of the five proteins successfully polymerised to amyloid fibrils in chapter 3 only three were suitable for assessment of their particle size distributions during mechanical perturbation: lysozyme, β-lactoglobulin and α-synuclein.

4.2.2 Quantitative Assessment of Lysozyme, β-Lactoglobulin and α-Synuclein Fibril Fragmentation

The fragmentation of lysozyme, β-lactoglobulin and α-synuclein fibrils was carried out twice using different time points over a period of about two weeks to show reproducibility.

Table 4.1 General Statistics for the Fragmentation of Fibril Samples

	Total Fragmentation Time (s)	No. Of Time Points	No. Of Images Captured
Lysozyme 1	1123200	12	26
Lysozyme 2	1048500	14	26
β-Lactoglobulin 1	1126800	12	18
β-Lactoglobulin 2	1296000	13	33
α-Synuclein 1	1293300	14	36
α-Synuclein 2	1380900	15	36

During this time a minimum of twelve aliquots were removed for deposition onto mica and imaging. For each time point multiple images were captured so as to achieve the 500 particles minimum target for analysis of the particle size (Table 4.1).

4.2.2.1   Analysis of Changes in Monomer Composition

To confirm that the fragmentation of lysozyme, β-lactoglobulin and α-synuclein fibrils was by physical breakage and not an artefact of monomer shedding the composition of the supernatant pre- and post- fragmentation was compared by residual monomer assay.
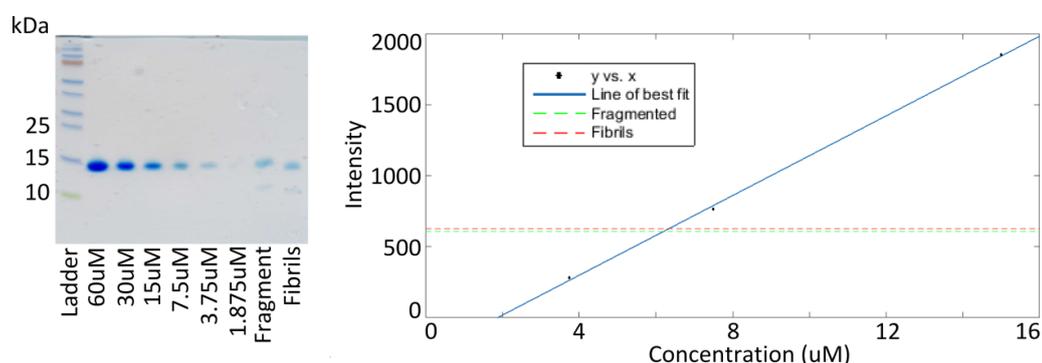
Figure 4.10   Lysozyme Fragmentation Residual Monomer Assay
Left – Coomassie stained SDS PAGE (Tris-Tricine) gel of a serial dilution of lysozyme and 120µM fibrils and fragmented fibrils centrifuged at 75k rpm with 10µl of the very top of the supernatant removed and diluted 1 in 10 for SDS PAGE analysis. The gel was then Coomassie stained. Right – Intensity values of the coomassie stained bands, from Aida densitometry analysis, plotted against µm concentration of dilution. Intensity of the fragmented and parent fibrils plotted onto the graph. Equation of line of best fit, y = 140.34x - 262.79

The values obtained from lysozyme pre- and post- fragmentation were 61.8µM and 63.5µM respectively (Fig. 4.10), β-lactoglobulin less than 5µM respectively (Fig. 4.11) and α-synuclein 4.2µM and 5.7µM respectively (Fig. 4.12). The different level of residual monomer composition before fragmentation for the different proteins is due to different efficiencies of conversion to fibrils. With respect to the starting residual monomer concentration for the fragmentation experiment the fragmented samples exhibit similar monomer compositions with lysozyme differing by only 1.4% (pre- 51.5%, post- 52.9%), α-synuclein by 1.25% (pre- 3.5%, post- 4.75%) and β-lactoglobulin showing no difference at all (pre-/post- 4%).These values show that the major process observed during mechanical agitation is fragmentation and not monomer release.
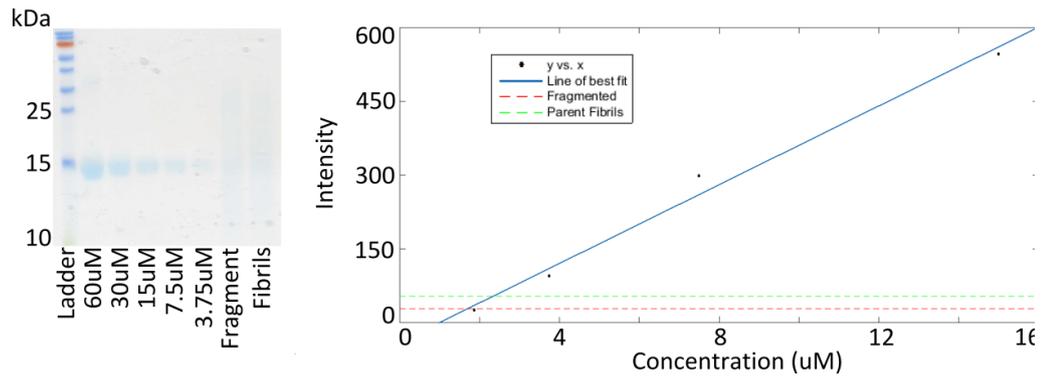
Figure 4.11   β-Lactoglobulin Fragmentation Residual Monomer Assay

Left – Coomassie stained SDS PAGE (Tris–Tricine) gel showing a serial dilution of β-lactoglobulin and 120µM fibrils and fragmented fibrils centrifuged at 75k rpm with 10µl of the very top of the supernatant removed and diluted 1 in 10 for SDS PAGE analysis. The gel was then Coomassie stained. Right – Intensity values of the coomassie stained bands, from Aida densitometry analysis, plotted against µM concentration of dilution.   Intensity of the fragmented and parent fibrils plotted onto the graph. Equation of line of best fit, y = 140.34x - 262.79.
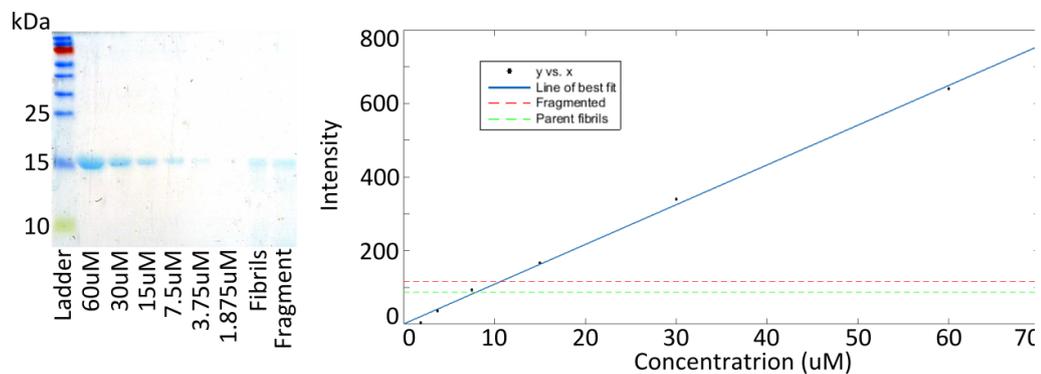
Figure 4.12   α-Synuclein Fragmentation Residual Monomer Assay

Left – SDS PAGE (15% Tris-Glycine) showing a serial dilution of α-synuclein and 120µM fibrils and fragmented fibrils centrifuged at 75k rpm with 10µl of the very top of the supernatant removed and diluted 1 in 10 for SDS PAGE analysis. The gel was then Coomassie stained. Right – Intensity values of the coomassie stained bands, from Aida densitometry analysis, plotted against µM concentration of dilution. Intensity of the fragmented and parent fibrils plotted onto the graph. Equation of line of best fit, y = 10.81x + 0.5979

.

4.2.2.2   Analysis of Particle Size Distributions

MatLab analysis of the fragmented fibril sample images allowed quantification of the

particle size associated with individual particles in solution and allowed a distribution to be

built based on combination of data from multiple fibrils imaged over multiple images.

This data was used to generate mean and deviation values for the contour length and

heights of the particles traced (Tables 4.2, 4.3, 4.4)

Table 4.2   Table Showing the Height and Weight Average Contour Length Details Extracted from the Second Lysozyme Fragmentation

Lysozyme 2						
Seconds	No. Of Particles	No. Of Images	Mean Height (nm)	Height SD	Weight Mean Contour Length (nm)	Weight Mean Contour Length SEM
600	1823	3	3.6848	1.37	2334.036	54.665604
5400	1384	2	3.5426	1.4801	1841.263	49.493465
18000	1497	2	3.0821	1.4582	1775.156	45.880241
91800	855	1	4.0031	1.6653	1295.904	44.318981
171000	2723	2	2.8641	1.4731	1119.322	21.450197
264600	4893	3	3.4697	1.5126	929.2263	13.284153
346500	6504	3	3.1168	2.3732	751.8851	9.3231192
432900	7733	3	2.8033	1.4206	677.0684	7.6994295
518400	5469	2	3.5205	1.5799	612.7526	8.285738
610200	2949	1	3.5279	1.5608	551.0062	10.146565
688500	4255	1	3.666	1.5386	465.355	7.1340237
783000	4151	1	3.3858	1.8248	396.0665	6.1474037
873000	5763	1	2.0055	1.0922	371.1635	4.889235
1048500	4961	1	3.2582	1.9201	356.4145	5.060235

Table 4.3   Table Showing the Height and Weight Average Contour Length Details Extracted from the First β-Lactoglobulin Fragmentation

β-Lactoglobulin 1						
Seconds	No. Of Particles	No. Of Images	Mean Height (nm)	Height σ	Weight Mean Contour Length (nm)	Weight Mean Contour Length SEM
300	2175	2	2.4939	1.1793	1117.791	23.96796
7200	2434	2	2.4133	1.2981	627.5675	12.720383
27000	1972	1	2.683	1.2608	396.9125	8.9380207
86400	2616	1	2.6647	1.5016	281.9121	5.5118179
172800	4050	1	2.42	2.1555	246.7354	3.8770728
331200	8653	2	2.5237	1.9119	200.9448	2.1601973
370800	3626	1	2.8086	1.9598	220.9903	3.6699436
432000	6821	2	3.2308	2.0881	227.5278	2.7549291
520200	5732	2	2.6571	2.0754	213.0976	2.8146567
691200	5770	2	3.9405	2.8816	217.1023	2.8580935
867600	3341	1	3.9146	2.7824	180.9802	3.1310697
1126800	5694	1	3.9262	2.9719	117.3773	1.5555192

Table 4.4   Table Showing the Height and Weight Average Contour Length Details Extracted from the First α-Synuclein Fragmentation

α-Synuclein 1						
Seconds	No. Of Particles	No. Of Images	Mean Height (nm)	Height σ	Weight Mean Contour Length (nm)	Weight Mean Contour Length SEM
300	1105	6	6.9378	2.4188	1862.12	56.017857
7200	560	5	5.7856	2.402	2124.553	89.77874
18900	754	3	7.0573	2.4883	1341.227	48.844589
78600	1525	2	7.3825	2.501	758.1974	19.415445
106800	1648	2	6.1357	1.8992	888.4842	21.886238
175500	3524	2	7.1689	2.2866	544.5122	9.1725397
257100	4615	2	6.8489	2.1034	489.1743	7.2007508
340200	4867	2	6.7875	2.2465	453.9283	6.5066371
441900	3115	2	6.9964	2.1753	412.3436	7.3880563
530700	6067	2	6.4133	2.2122	338.7706	4.3492937
606300	7320	2	6.8247	2.4654	343.6742	4.0169039
780300	8112	2	6.4494	2.2001	295.6831	3.282937
1042200	7827	2	6.1335	2.1109	263.4355	2.97767
1293300	6563	2	7.0129	2.3565	241.8443	2.9852765

The extracted data shows changes in the mean heights and lengths of traced particles over the time of fragmentation such as the decrease in weight average contour length for lysozyme from 2.34µm to 356.41nm, β-lactoglobulin from 1.17µm to 117.38nm and for α-synuclein from 1.86µm to 241.84nm. This decay in contour length can also be observed by histograms (Fig. 4.13, 4.14, 4.15, Left).
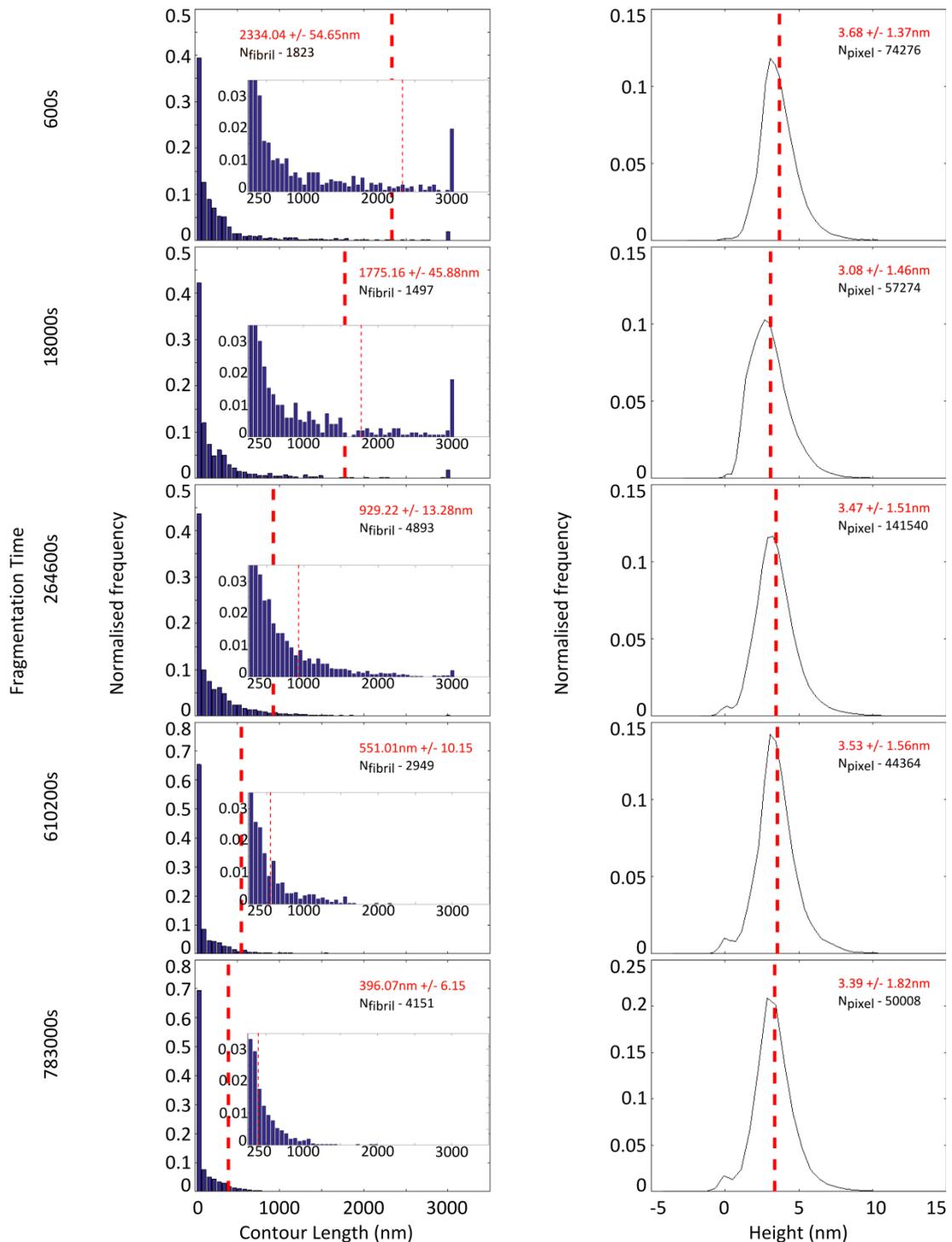
Figure 4.13   Height and Length Distribution of Amyloid Fibrils Analysed from AFM Analysis of a Lysozyme Fragmentation Experiment.
Left – Extraction of contour length data for the fibrils traced by MatLab from multiple images per time-point and plotted as a histogram with weight mean contour length shown with dotted red line and value with standard error mean in same colour. N_{fibril} is the total number of particles traced to give this value. Inserts were of the contour length region from 250nm to 3500nm with a lower normalised frequency to highlight the change in proportion of longer particles. Right – The distribution of heights of pixels traced by MatLab from multiple images per time point with the mean height shown by a dotted red line and the value and standard devation in the same colour. N_{pixel} is the total number of pixels traced to give this data
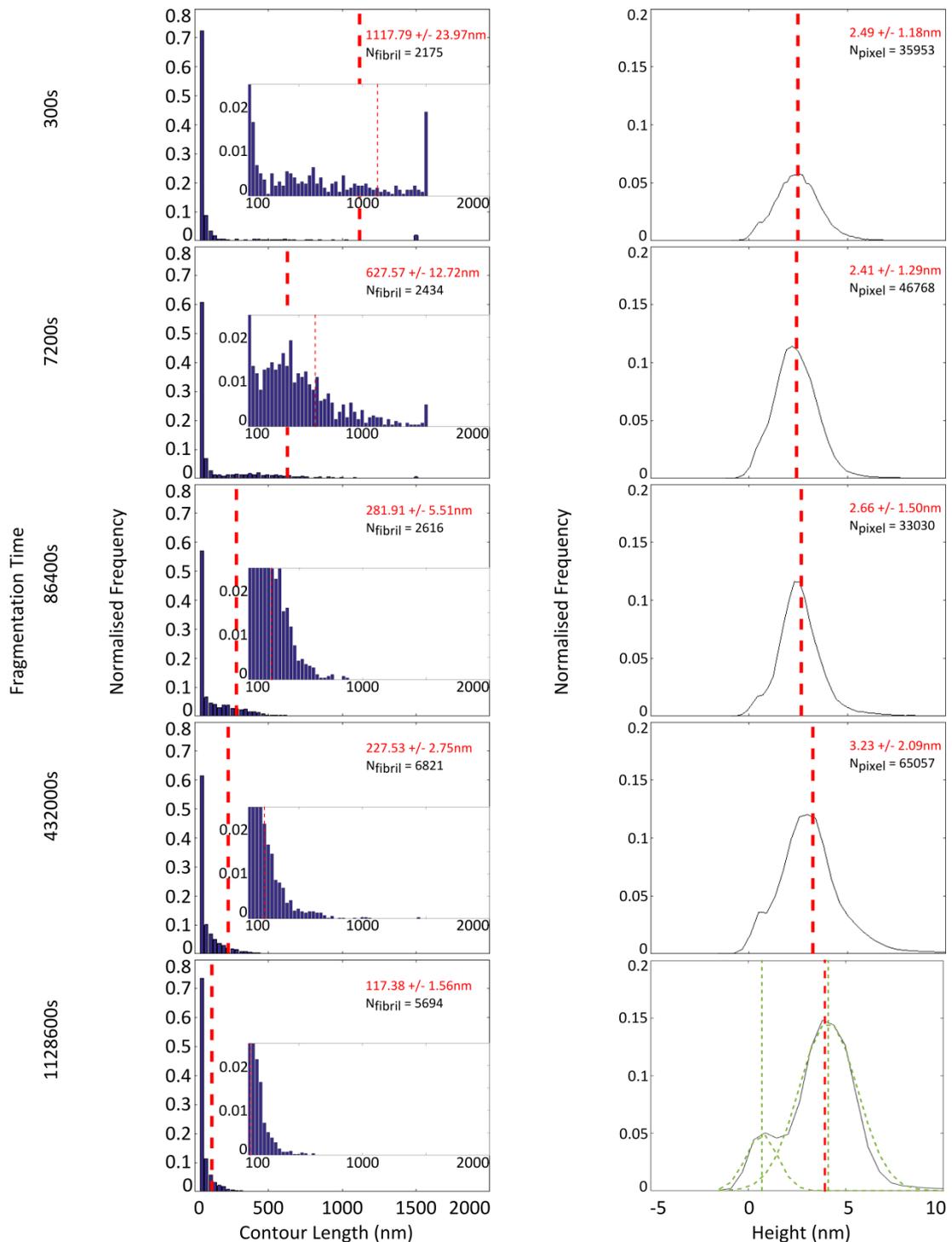
Figure 4.14 Height and Length Distribution of Amyloid Fibrils Analysed from AFM Analysis of a β-Lactoglobulin Fragmentation Experiment.

Left – Extraction of contour lengths for the traced fibrils from multiple images per time-point and plotted as a histogram, weight mean contour length shown with dotted red line and value with standard error mean in same colour. N_{fibril} is the total number of particles traced to give this value. Inserts: contour length region from 100nm to 2000nm with a lower normalised frequency to highlight the change in proportion of longer particles. Right – The distribution traced heights with the mean height shown by a dotted red line and the value and standard deviation in the same colour. N_{pixel} is the total number of pixels traced to give this data. The final height graph shows two normal distributions, coloured green, fitted to the data highlighting the two different populations contained in the sample. The means of the two populations are also plotted in green.
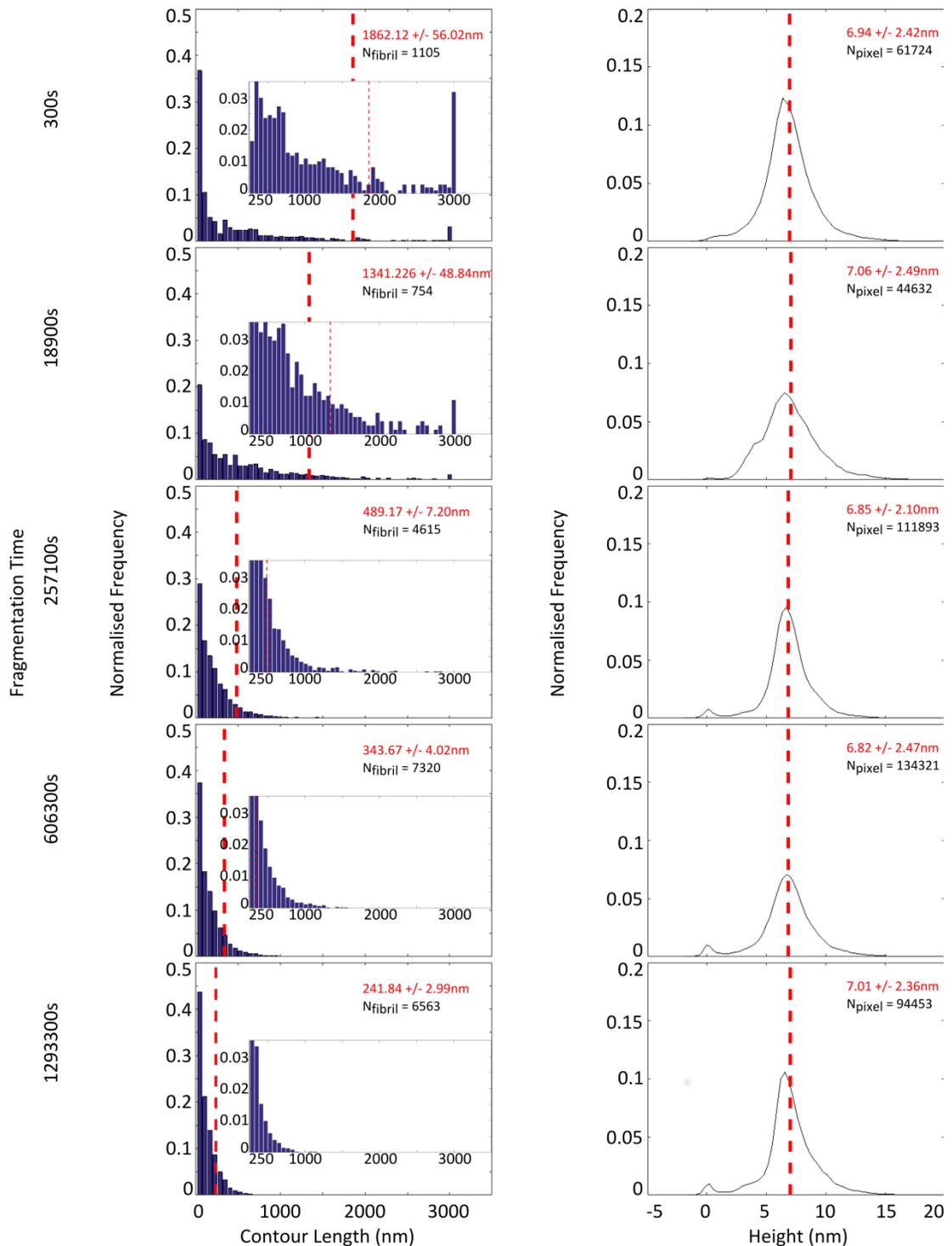
Figure 4.15   **Height and Length Distribution of Amyloid Fibrils Analysed from AFM Analysis of a α-Synuclein Fragmentation Experiment.**

Left – Extraction of contour length data for the fibrils traced by MatLab from multiple images per time-point and plotted as a histogram with weight mean contour length shown with dotted red line and value with standard error mean in same colour. N_{fibril} is the total number of particles traced to give this value. Inserts were of the contour length region from 250nm to 2000nm with a lower normalised frequency to highlight the change in proportion of longer particles. Right – The distribution of heights of pixels traced by MatLab from multiple images per time point with the mean height shown by a dotted red line and the value and standard deviation in the same colour. N_{pixel} is the total number of pixels traced to give this data.

Characterisation of the heights throughout the fragmentation experiment is important for subsequent analysis as any change in lateral association and therefore height alters the fibril concentration on the mica surface and makes bias correction difficult. The mean height of lysozyme and α-synuclein fibrils was consistent between experiments and remained constant throughout the fragmentation with values of 3.15/3.28nm and 6.71/7.04nm respectively. The mean of the mean heights observed through the fragmentation experiments was consistent between β-lactoglobulin fragmentation experiments with values of 3.93nm and 3.99nm but the shape of the distribution of heights showed that this mean is based on a disparity between the beginning and end of the fragmentation where the mean height increases during the fragmentation experiment (Fig. 4.14, Right and 4.16) in both experiments, in fact both fragmentations show a 50% increase from the mean of the early time points to the later. As well as the increase in mean height of the sample observed for the later β-lactoglobulin fragmentation samples there is an increase in the proportion of a thinner subpopulation of particles. The fitting of two normal distributions to the data shows that the thinner materials have a mean of 0.697nm and the larger particles have a mean of 4.099nm which is very similar to the mean of the entire sample. The thinner particles formed are potentially protofilaments derived from the unentwining of the filaments observed at the beginning of the fragmentation, although this does not fit with the assessment that β-lactoglobulin fibrils of this height are protofibrils (Schleeger et al., 2013; VandenAkker et al., 2011). The liberation of these particles and then reattachement to the larger particles could give rise to the larger fibrils observed, in the later time points. The process of fragmentation could promote this process and would explain the change in height profile during fragmentation. Analysis of the distribution of traced heights (Fig. 4.13, 4.14, 4.15, Right) over the course of the fragmentation reactions for lysozyme and α-synuclein (Fig. 4.13, 4.15, Right) suggests that there is no change in lateral association of the fibrils during fragmentation.
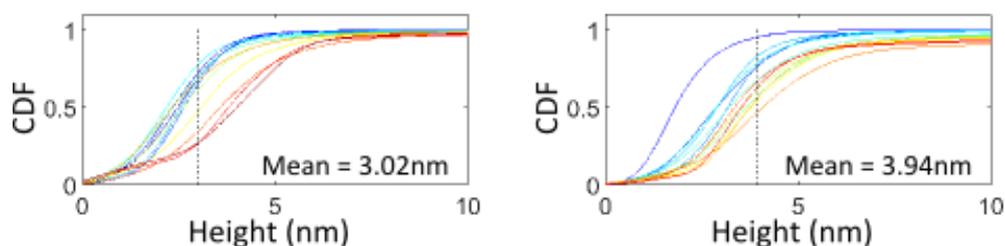
Figure 4.16   Comparison of the Height Distribution of β-Lactoglobulin Fibrils During Different Fragmentation Experiments
Left – A plot of the probability of the different heights found in solution, CDF, showing the distribution of heights observed over the course of the fragmentation of β-lactoglobulin 1, with earliest time points in blue/green and the late in red. Right – As Left but for β-lactoglobulin 2. Early time points in blue progressing to later time points in red.

4.2.2.3   Analysis of Particle Length Decay During Fibril Fragmentation

To visualise the breakage of fibrils over the course of the fragmentation experiments log-log plots of the weight average contour length (nm) were plotted against fragmentation time, in seconds for every image in both fragmentations for each protein (Fig. 4.17, circles).

β-Lactoglobulin exhibits a linear decay (Fig. 4.17, middle) for the first part of the fragmentation until the point when the distortion in fibril height was observed. Lysozyme and α-synuclein differ from β-lactoglobulin as they are linear at the end of the fragmentation (Fig. 4.17, top and bottom, circles) but appear to have a flatter appearance and therefore slower decay at the beginning. This biphasic rate of decay in weight average contour length over the duration of the fragmentation experiment, when visualised on a log/log plot, is similar to what is observed in the fragmentation of β2-microglobulin (Xue and Radford, 2013). To remove any skew in the data which could be derived from extremely high particles or materials which are less than the ability of the AFM to determine if they elongated any particles higher than 99% of the height distribution and less than 19nm in length were discarded from the analysis (Fig. 4.17, Top/Middle/Bottom, crosses). The removal of these extremes did not alter the biphasic shape of the decay or the weight average contour length values of the samples. There are two potential sources of bias associated with the automated tracing of AFM images which could alter the shape

of weight average contour length over fragmentation time decay curve. Firstly the deposition of fragmentation samples onto mica potentially assumes that the affinity of long and short particles to the mica surface would be similar, which cannot be assumed. Secondly the tracing of fibrils in an AFM scan by the MatLab application can also introduce error. Fibrils deposited onto a surface can overlap with each other, form larger untraceable aggregates and cluster together, the MatLab application uses the approximate fibril width, angle θ and a height cut off (z) (See 3.2.3.2) to trace the particles and provide parameters for acceptance or failure of the trace. The process of tracing for long fibrils is more likely to fail due to their greater waviness, potential to overlap and becoming part of larger aggregates than small particles. This means that the tracing potentially underestimates the influence of longer particles on the population and therefore give rise to a low weight average contour length for the image. This underestimation would be more pronounced at early time points than late due to the greater proportion of long particles at this time and might add to the biphasic shape of the decay. To correct for any bias introduced by differing affinities and tracing, the deposition of fibrils onto the mica surface must be identical throughout the fragmentation experiment. When this is the case it can be assumed that the total length of particles deposited onto the mica surface will be the same independent of fragmentation time. A plot of total length traced against fragmentation time will, for an unbiased sample give a flat horizontal line but where bias is incorporated will show a greater length of particles traced at later time points. To assess this bias and determine an appropriate weighting for particles of different sizes a MatLab script (Xue, unpublished) was used which looks at the weight average contour length and total traced length per image to determine an appropriate correction method. Comparison of the sum total contour length of particles on a surface with the weight average contour length of the sample shows that α-synuclein and lysozyme are underestimated at the beginning of the fragmentation where as β-lactoglobulin is not. To assess this bias and determine an

appropriate weighting for particles of different sizes a MatLab script (Xue, unpublished) was used which looks at the weight average contour length and total traced length per image to determine an appropriate correction method. α-Synuclein was corrected using a linear correction method whilst lysozyme was subjected to a power law correction method and due the lack of skew β-lactoglobulin observed did not require any correction.
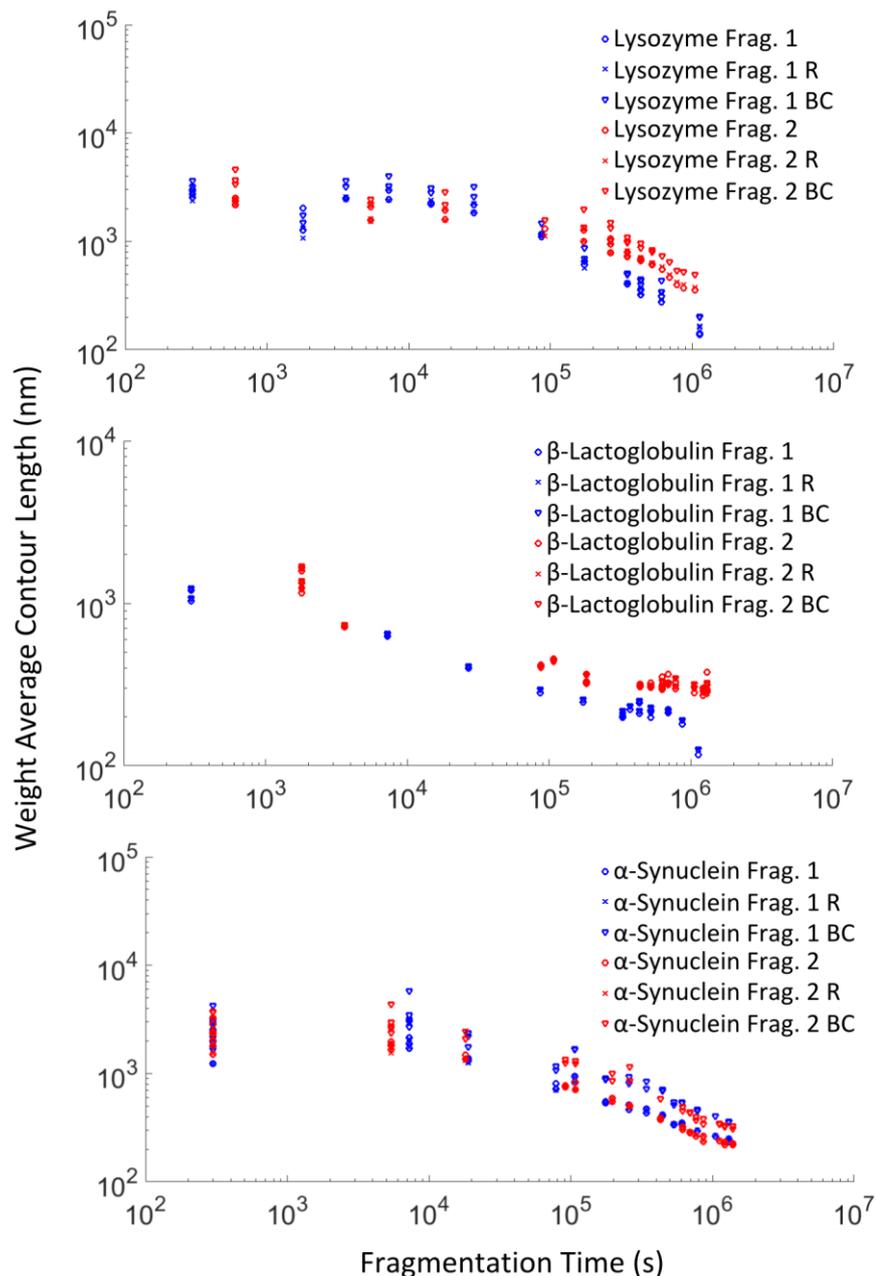
Figure 4.17   Comparison of the fragmentation over time.
Top – Lysozyme, Middle – β-Lactoglobulin and Bottom – α-Synuclein. Circles represent uncorrected data. Crosses represent data with small particles removed. Inverted triangles represent bias corrected data.   Blue shapes represent the first fragmentation and red denote the second fragmentation.

Comparison of this data to the uncorrected data (Fig. 4.17, Top/Middle/Bottom, inverted triangles) shows that whilst the data points for each α-synuclein and lysozyme particle are at a higher weight average contour length the shape of each data remains the same. Whilst this plot shows the difference that bias correction has on the weight average contour length during the fragmentation the change of the distribution of particle contour lengths over the duration of the fragmentation is better described by the cumulative density function (CDF) (Fig. 4.18).
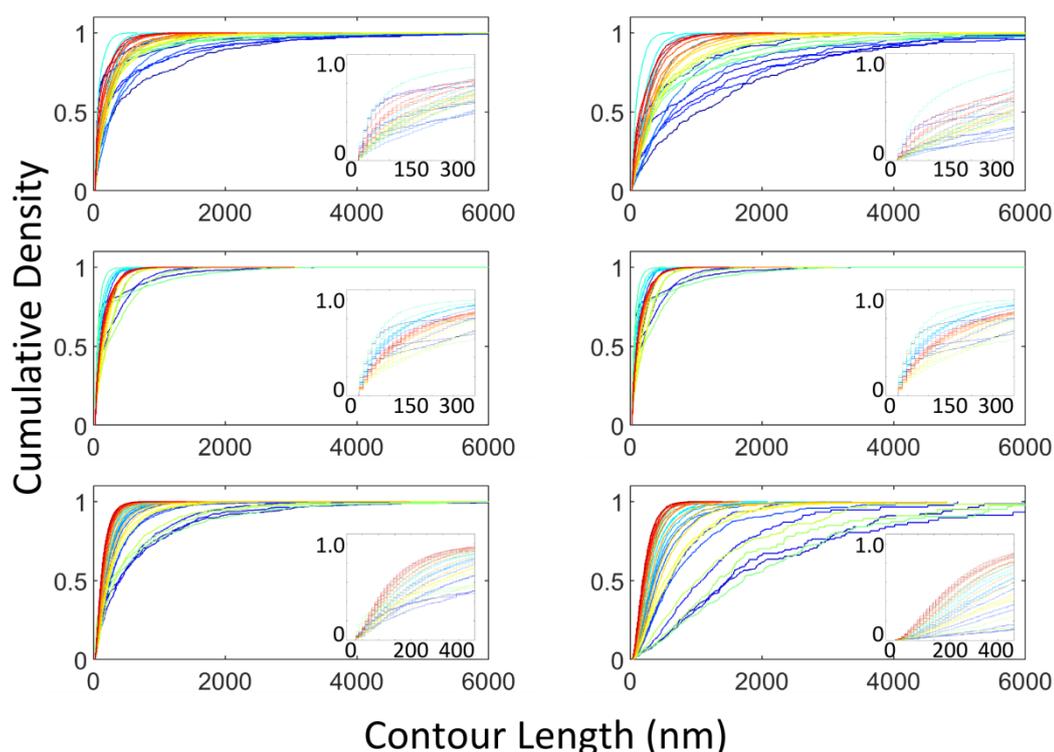
Figure 4.18   Cumulative Density Plots Showing the Change in Contour Length Population during Fragmentation.
Left – Uncorrected data, Right – Bias Corrected data. Top – Lysozyme, Middle – β-Lactoglobulin and Bottom – α-Synuclein. Cumulative density for each time point for both fragmentation experiments with the earliest time in blue progressing to the late time points in red. Inserted into each graph is an expansion of regions 0 to 300 for β-lactoglobulin and lysozyme and 0 to 400 for α-synuclein.

Apart from β-lactoglobulin whose data did not require correction the bias corrected data shows an increase in the proportion of longer particles at the early time points in comparison to the uncorrected data. This data shows a more pronounced change than when just looking at the weight average contour length (Fig. 4.17) and allows for the rate of

change of the population distribution to be assessed. As β-lactoglobulin did not require

correction the bias corrected data is unaltered but does show that the early timepoints are

skewed by the presence of small particles which diminish after a short duration of stirring

(Fig. 4.18, Middle) something which is also visible from the histogram distribution of

particle contour lengths at the beginning of fragmentation (Fig. 4.14).

4.2.2.4   Calculation of the Persistence Length (L_p) of Fragmented Amyloid Fibrils

The MatLab analysed fragmentation images allowed calculation of the L_p for the polymers

observed in the experiments in batches of six using a a Matlab algorithm (Xue unpublished)

(Table 4.6). The samples were concatenated and randomised to avoid the different

fragmentation times skewing the analysis. The L_p values obtained were then averaged to

give the L_p for the entire dataset.

Table 4.6   Table Showing the Calculated Persistence Length for the Fragmented Fibrils*

	α-Synuclein 1	α-Synuclein 2	β- Lactoglobulin 1	β- Lactoglobulin 2	Lysozyme 1	Lysozyme 2
L_p (nm)	1754.72126	1906.64144	2741.92202	3780.535	1259.951	1883.071
σ (nm)	363.184241	353.12246	604.089368	605.05157	133.7309	188.3912
N_P	57602	83823	52884	61604	34530	54960
N_I	36	36	18	33	26	26

*Where L_p = persistence length, σ = standard deviation, N_P = Number Particles analysed and N_I = the number of images scanned.

In comparison to the data collected in chapter 3, Table 3.1, the persistence length values

maintain the same ranking with respect to increasing persistence length and similar

magnitude. The L_p values for fibrils of α-synuclein (1754.72nm and 1906.64nm) and

lysozyme (1259.951nm and 1883.071nm) are similar in their magnitude and therefore

exhibit similar rigidities. The L_p values for β-lactoglobulin (2741.92nm and 3780.535nm) are

more similar to previously calculated values (VandenAkker et al., 2011) than the values in

chapter 3. The data collected from the fragmentation experiments allows for a huge

number of particles to be assessed in the calculation of persistence length, in fact ~1000

times more particles than the values previously published (Table 3.2). The AFM images of

the initial fragmentation time points are in agreement with this observation as lysozyme

and α-synuclein fibrils are observed as long wavy fibrils where as β-lactoglobulin fibrils are

shorter and straighter (Fig. 4.7, 4.8 and 4.9).

4.2.2.5   Determination of the Rate of Fibril Fragmentation

Rather than use an approach requiring *a priori* knowledge of the fragmentation process

Xue et al (Xue and Radford, 2013) used a data driven approach determine which model

best describes the fragmentation of β2-microglobulin fibrils. By determining lines of best

fit for the contour length distributions at each time point the ability of each model to depict

the fragmentation was assessed using Akaike information criterion (AIC) scoring method.

This method together with the Hill model (Hill, 1983) describing the fragmentation of stiff

rods was employed to analyse the change in contour length distribution over time

(Equation 4.1)

Equation 4.1   Hill Model of the fragmentation of stiff rods*

$$\kappa_F(i,j) = \kappa_1[j(i-j)]^{\kappa_2-1}\left[\frac{(i-j)lnj + jln(i-j)}{i^{\kappa_2+1}}\right]$$

*Where κ_F= Fragmentation rate, i=total fibre length, j=the fragmented fibre length (see Fig. 4.2), κ_1 and κ_2 are floating parameters determined quantitatively using maximum likelihood estimations derived from contour length distribution derived from AFM images (Hill, 1983; Xue and Radford, 2013).

The contour length distributions generated for the fragmentation of β-lactoglobulin, α-

synuclein and lysozyme (Xue, unpublished) where used in conjunction with the Hill model

(Hill, 1983) to calculate the constants κ_1 and κ_2 (Equation 4.1) required to generate a rate

of fragmentation for specific lengths of fibril using a Matlab script (Xue, unpublished).

The Hill model describing the fragmentation of stiff rods was fit to the length distribution

data by maximum-likelihood estimation method (W. F. Xue, Homans, & Radford, 2009).

Best fit parameters κ1 and κ2 then describe the change in particle length distribution

during the fragmentation of the amyloid fibrils observed during the experiments. These

best fit parameters described in equation 4.1 were used to determine the rate of fragmentation of particle with length i. A plot of the rate of amyloid fragmentation rate against the length of the particle gives rise to a graph showing longer particles fragment faster than short (Fig. 4.19).

Over the course of the fragmentation experiment the rate of fibril fragmentation changes with longer particles exhibiting a consistent rate until a particle size is formed which is more resistant to breakage dominates the population. At this point a significant drop in fragmentation rate is observed (Fig. 4.19), this data is in agreement with the results observed with β2-microglobulin (Xue and Radford, 2013).
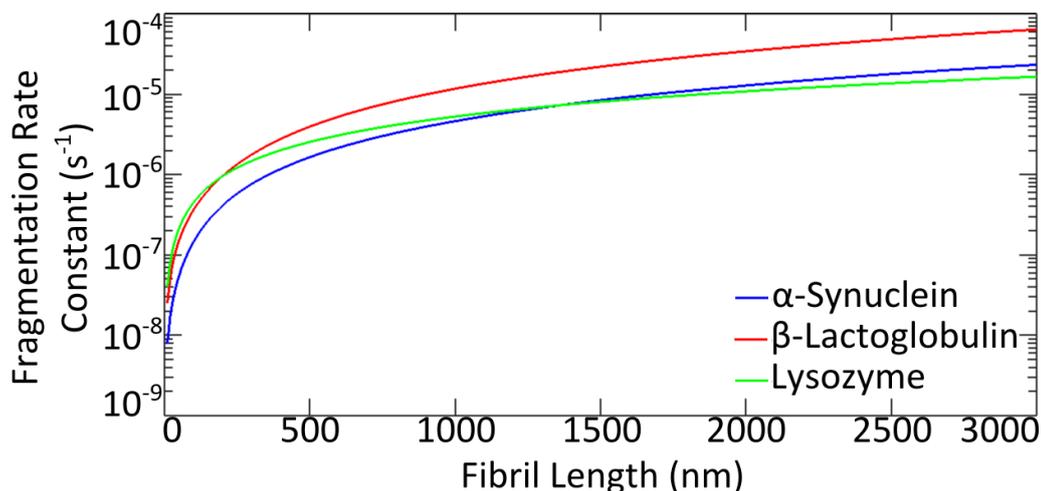
Figure 4.19   Graph Showing the Relationship of the Fragmentation Rate Constant with Fibril Length
A semi-log plot showing the change in fragmentation rate constant of α-synuclein (blue), β-lactoglobulin (red) and lysozyme (green) with fibril length.

Table 4.7   Table showing the Fragmentation rate of the 3 studied proteins at 3μm

	α-Synuclein	β-Lactoglobulin	Lysozyme
Fragmentation Rate at 3μm	0.00002325s^{-1}	0.0000639s^{-1}	0.000016525s^{-1}

Comparison of the fragmentation rate constant for the longer particles shows that β-lactoglobulin is more sensitive to breakage than lysozyme and α-synuclein which exhibit a

similar rate. When the fragmentation rates of the longer particles, 3µm, (Table 4.7) are compared to the persistence length of the polymer a correlation starts to appear where more flexible polymers fragment slower than more rigid polymers (Fig. 4.20).
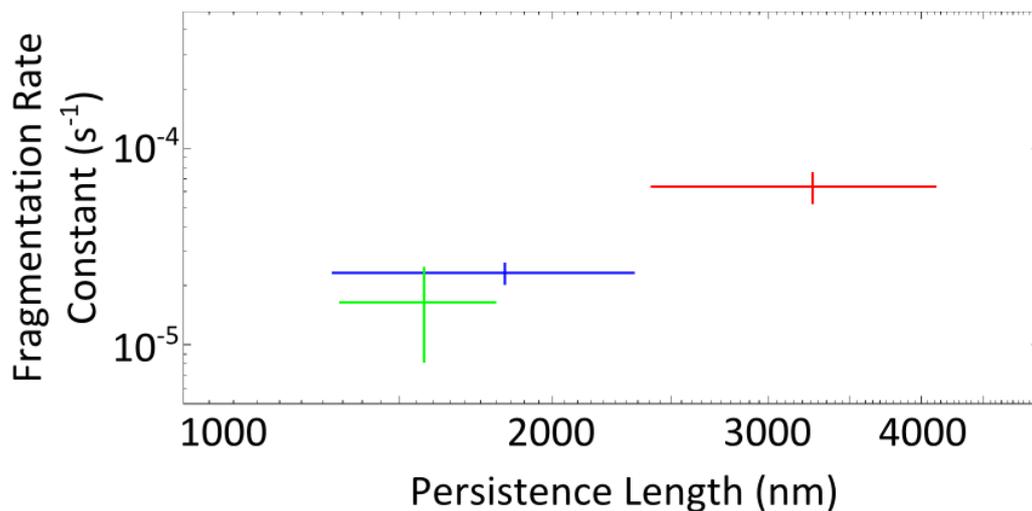
Figure 4.20   Graph Showing the Relationship Between Persistence Length and Fragmentation Rate Constant of 3µm Fibrils.
A log-log plot showing the Fragmentation rate constant of fibrils 3µm in length compared to the persistence length of the fibrils. Lysozyme (green), α-Synuclein (blue) and β-Lactoglobulin (red).

4.3   Discussion

Due to the role amyloid plays in disease, and the physical properties that make them useful nanomaterials, it is important to gain an understanding of the role and influence of fibril fragmentation on the kinetics of fibril formation, the stability of amyloid fibrils and their cytotoxic potential. AFM imaging is the perfect tool for quantifying changes in fibril dimensions for individual particles as well as enabling the deconvolution of fibril properties such as persistence length, L_p.

Although AFM instrumentation is sensitive enough to observe the changes in length and height required to delineate the size distribution of particles contained within a sample, the development of deposition conditions that allow for optimum particle analysis at the beginning and the end of a fragmentation experiment are key to a successful fragmentation

analysis. Examples of where deposition conditions prevent this analysis are insulin B chain and Sup35. The formation of large aggregates from the Sup35NM species formed under current conditions makes proper fragmentation analysis impossible. Insulin B chain appears at the onset of mechanical perturbation to be a good candidate for fragmentation analysis, but the loss of affinity for the surface during fragmentation means that calculating particle distributions for the time points is possible but of limited value. The loss of affinity for the surface and generation of the background observed as insulin B chain fibrils fragment suggests that the fragmentation may occur via a different mechanism to that observed with lysozyme, β-lactoglobulin and α-synuclein fibrils.

Table 4.8   Table Comparing Height Data and Persistence Length*

	Height (nm)	σ	L_p (nm)	σ	N_p	N_i
α-Synuclein 1	6.47	1.41	1754.72	363.18	57602	36
α-Synuclein 2	6.58	1.39	1906.64	353.12	83823	36
β-Lactoglobulin 1	2.68	1.02	2741.92	604.09	52884	18
β-Lactoglobulin 2	3.06	0.94	3780.54	605.05	61604	33
Lysozyme 1	2.88	1.24	1259.95	133.73	34530	36
Lysozyme 2	3.03	1.11	1883.07	188.39	54960	36

*A table showing the height, persistence length, L_p, the number of particles, N_P, and number of images captured, N_I, to calculate this data.

The data extracted from the fragmentation of amyloid fibrils derived from lysozyme, β-lactoglobulin and α-synuclein shows that although these proteins as monomers have similar molecular masses, their aggregation into fibrils with the same cross beta sheet architecture gives rise to fibrils of different types of different diameters, 3.5nm, 3.9nm and ~7nm respectively. It is also evident by comparison with their height with persistence length that fibril flexibility is not related to the diameter of the fibril (Table 4.8) and therefore monomer size.

The height of particles throughout the fragmentation remains static for α-synuclein and lysozyme but β-lactoglobulin shows an increase in fibril height towards the end of the experiment. This increase in height would potentially alter the particles resistance to mechanical stress. The weight average contour length of the images showing deviations in the mean height are in the region of the fragmentation where the rate of fragmentation was slowing so whilst it potentially affects fragmentation here it is does not interfere with the rate observed for longer fibrils.

In terms of their physical appearance, once deposited onto mica, the fibrils of α-synuclein and lysozyme appear far more flexible than β-lactoglobulin due to their propensity to bend and form highly coiled structures (Fig. 4.7, 4.9 and 4.20) and is also in agreement with their calculated persistence lengths (L_p) 1259.951nm and 1883.071nm for lysozyme, 1754.721nm and 1906.641nm for α-synuclein compared to 2741.922nm and 3780.535nm for β-lactoglobulin (Table 7). These values of L_p suggest that β-lactoglobulin forms stiffer polymers than lysozyme and α-synuclein.

The change in the contour length distribution of the fibrillar samples over fragmentation time (Fig. 4.21) shows the progression of contour length distribution from longer particles to shorter ones over time but visually gives no idea if the rates of change are similar.

The conversion of these data into fragmentation rate constants allows for the comparison of the different proteins fibrils fragmentation propensity as well as the dependency on fibril length. The rate of fibril fragmentation, like β2-microglobulin, is dependent on length whereby shorter particles are more resistant to breakage than longer particles. This is in agreement with the model proposed by Xue *et al* (Xue and Radford, 2013) whereby fibril breakage occurs centrally, distal from the ends of the fibres. The smaller the particles the closer the middle of the particles is to the end of the fibrils making them more difficult to fracture.
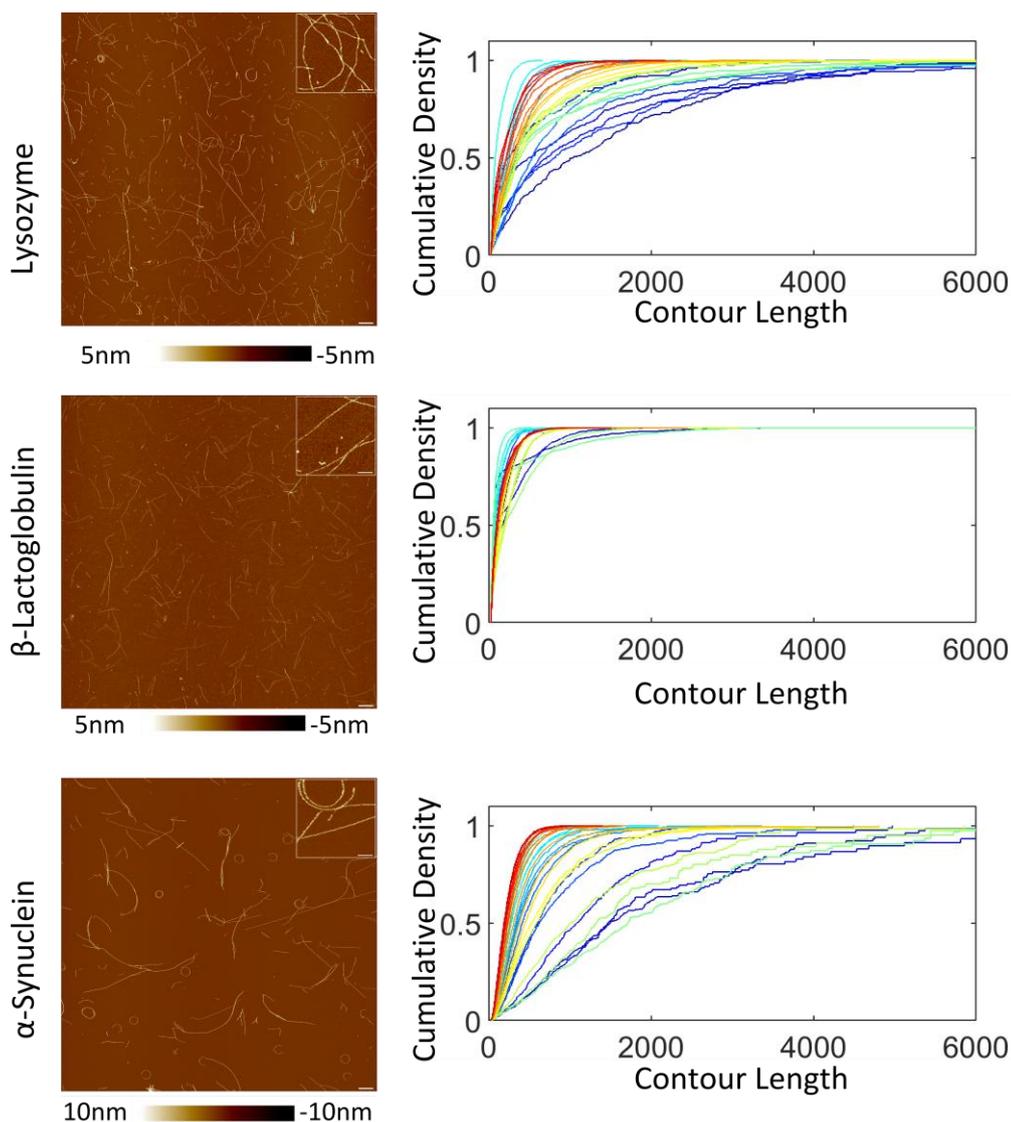
Figure 4.21   Comparison of Initial AFM images and Fragmentation Decay Experiments

Left – AFM images of lysozyme, β-lactoglobulin and α-synuclein at the start of fragmentation. 20µm x 20µm at 2048 x 2048 pixel resolution with 1µm scale bar and 1µm insert magnified x 5 with a 200nm scale bar. The height scales for β-lactoglobulin and lysozyme are set between -5 and 5nm and the α-synuclein is set between -10 and 10nm. Right – Cumulative density plots showing the change in contour length distribution from the beginning of the fragmentation (blue lines) to the end (red lines).

There appears to be no change in the approximate size where this change in fragmentation rate occurs but the initial rates of fibril fragmentation differ between proteins.   The difference in resistance to mechanical stress appears to be linked to the fibrils persistence length with the rate of fragmentation for the more rigid β-lactoglobulin being higher than that of the more flexible α-synuclein and lysozyme (Fig. 4.20).
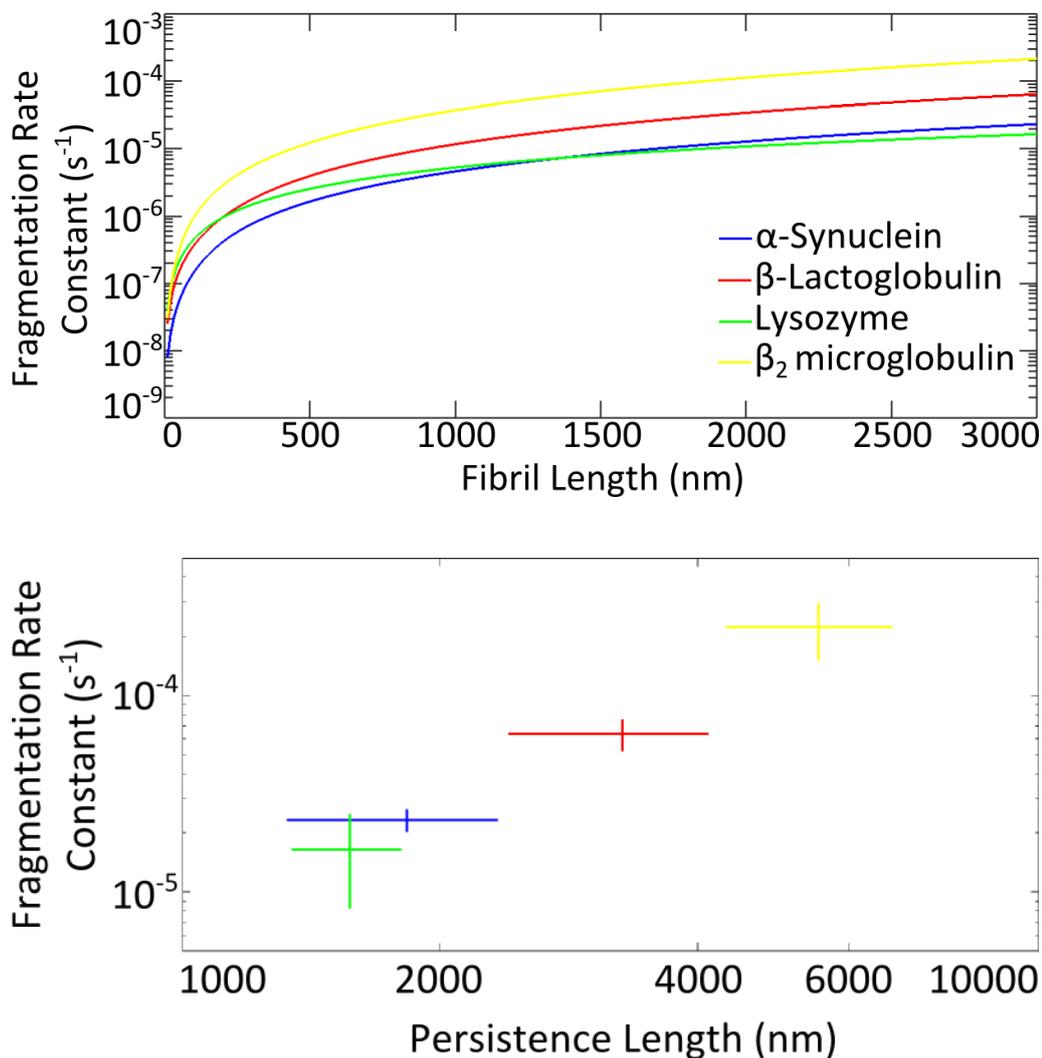
Figure 4.22   Comparison of the Kinetic Data and Persistence Length Data from this Thesis with the B2m data of Xue et al (Xue and Radford, 2013)**.**
Top – A semi log plot showing the dependence on fragmentation rate constant on the length of amyloid fibrils, as 4.19 but including the data for B2m (yellow). Bottom – A log-log showing how the persistence length of the polymer affects the fragmentation rate constant, as 4.20 but including the data for B2m (yellow).

When this data is compared to the previously published data regarding the fragmentation of B2m fibrils the correlation between the persistence length, L_p, of the particle and their resistance to mechanical stress continues (Fig. 4.22). The relationship between L_p and resistance to mechanical stress in the model systems so far analysed is linear with higher values of L_p giving rise to species more sensitive to mechanical stress (Fig. 4.23). The

correlation of L_p with resistence to mechanical stress can also be considered in combination with the work of vandenAkker *et al* (VandenAkker et al., 2011), previously discussed in chapter 3, who show, when investigating β-lactoglobulin only, that there is a relationship between persistence length and the proportion of cross-beta sheet contained within the structures.   An increase in cross-beta sheet proportion observed in the sample, by vibrational sum frequency generation, was shown to give rise to more rigid particles and therefore higher persistence length. In fact, the source of fibril flexibility has been further elucidated to be from the interactions between atoms in the cross-beta and not the sidechains of the amyloid fibrils (Knowles et al., 2007a). Whilst this analysis does work for amyloid derived from large protein molecules it does pose the question as to what this means for small amyloidogenic peptides whose proportion of cross beta sheet in the sample is much higher due to much less sequence redundancy in the amyloid core. These materials should therefore exist as straight fibres with high persistence length and this is indeed the case for TTR(105-110) (Knowles et al., 2007b) and ILQINS (Lara et al., 2014). The link between structure and persistence length highlights a potential link between the molecular structure of fibrils, their resistance to mechanical stress and therefore cytotoxicity.

To conclude, this work has developed a comparative method for the assessment of amyloid fibrils resistance to mechanical stress.   The results of this analysis has shown the importance of imaging conditions on determining fragmentation rates but more importantly shows that polymer diameter, monomer size and rigidity are unrelated. The rate of amyloid fibril fragmentation for the proteins analysed in this thesis follows the same dependency on length as β2-microglobulin. The agreement of this data with the previously published fragmentation mechanism (Xue and Radford, 2013) adds weight to their mechanism. The most important finding is the observation that the rate of fragmentation for longer particles is influenced by the L_p of the polymers such that more flexible fibrils are

more resistant to mechanical stress than rigid fibrils (Fig. 4.23) and therefore the molecular structure of the polymers (VandenAkker et al., 2011). Whilst there appears to be a correlation of L_p with different fragmentation rates the deviation in values is relatively small (10^{-4} to $10^{-5}s^{-1}$) which can be due to them having the same generic cross-beta sheet structure (Knowles et al., 2007a). It would be interesting to observe if this correlation is amyloid specific or observed in other materials whether they be synthetic or biological for example (Knowles and Buehler, 2011) or the materials highlighted in Fig. 3.34.

The linearity of the relationship between persistence length and fragmentation rate of long particles for the four proteins so far studied by this method suggests that it is possible to predict the fragmentation rate of amyloid fibrils using their calculated persistence length or even their proportion of cross-beta sheet.
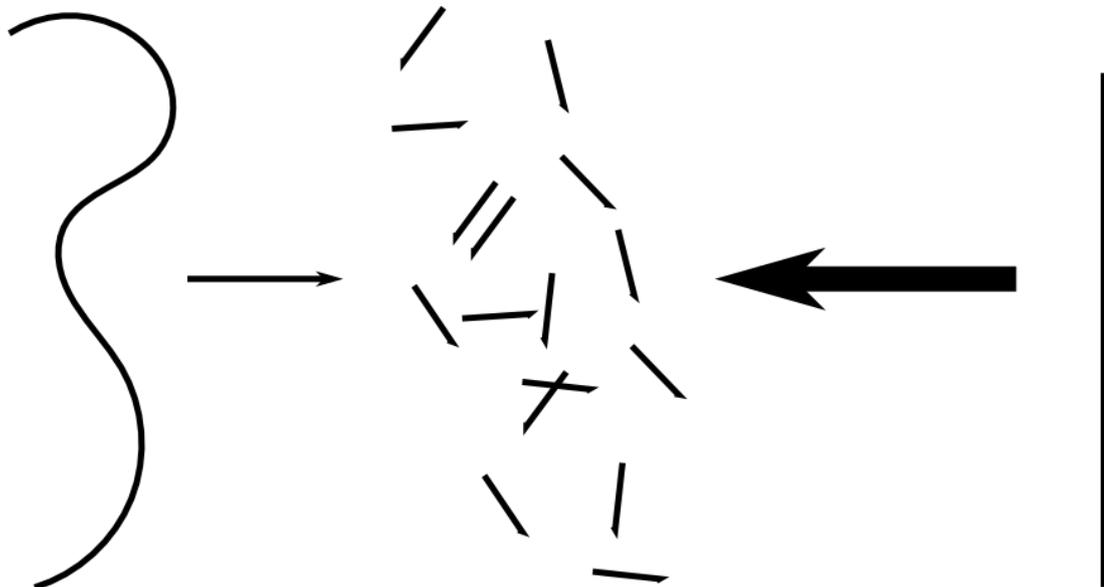
Figure 4.23   Fragmentation Rate of Amyloid Fibrils Dependent on the Persistence Length of the Polymer
Cartoon showing the fragmentation of flexible amyloid fibrils with low Lp values on left and rigid amyloid fibrils with higher L_p values on the rate. The fragmentation rate for the more rigid particles is higher for the more rigid particles as signified by the different sizes of the arrows.

4.4 Future Work

Increasing the number of model systems investigated will allow further confidence to be built in the correlation between L_p and the fragmentation rate constant for longer particles persists with a larger data set. This could also be supplemented by studying the fragmentation of other synthetic and biological materials with different L_p and core structures.

Also of interest would be to examine the the effect of altering fragmentation conditions on the kinetics of breakage to see if the order of resistance is maintained.

The development of this technique and fibril formation conditions for proteins such as Sup35 will hopefully lead to the potential of examining the rate of biologically relevant fragmentation conditions such as the Hsp104 mediated fragmentation of Sup35.

5

Chemical Labelling of Amyloid Fibrils and its Application to Size Determination

5.1   Introduction

Chemical biology is at the interface of chemistry, biology and biochemistry and utilises

synthetic chemistry to understand biological processes, and covers many areas of scientific

research such as bioconjugation, proteomics, affinity enrichment and genetic engineering.

In particular chemical biology has found utility in drug discovery research where

bioconjugation and genetic engineering is used to generate chemically enhanced biologics

such as antibody drug conjugates (ADC) (Agarwal and Bertozzi, 2014) and when applied to

proteomics/affinity enrichment allows the optimisation of small molecule therapeutics by

highlighting drug-target co-localistion, and drug binding sites (Bunnage et al., 2013, 2015).
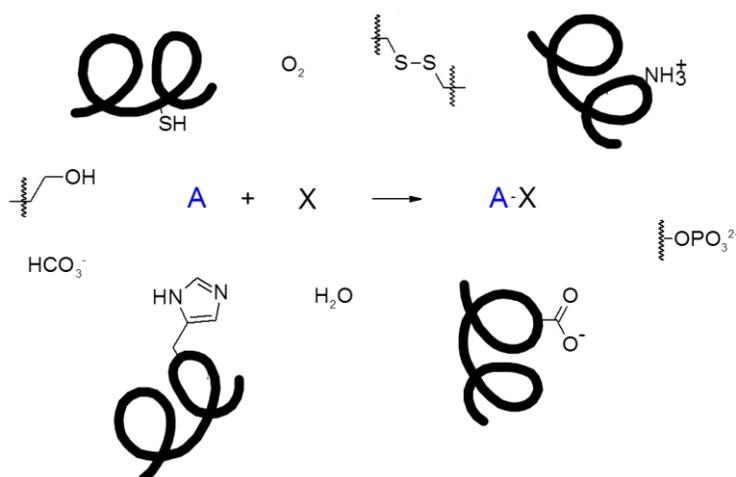
Figure 5.1   Picture Highlighting the Complexity of Bioorthogonal Labelling
For a bioorthogonal reaction the conversion of A and B into a molecule A-B must occur in the
presence of numerous reactive functional groups. Adapted from Sletten and Bertozzi (2009).

The ability to modify biomolecules with different functionality such as fluorescent dyes,

radioactive isotopes, enrichment tags and pharmaceutically active molecules is important

to these endeavours. Bioconjugation allows the modification of biomolecules by reacting

functional groups in the molecule with appropriate functionality in another molecule, for

this to be successful the coupling reactions need to be compatible with the aqueous

environment that biomolecules typically exist in. Traditional labelling strategies utilize the

reactivity of amino acids, for example amide coupling between amines and activated

carboxylic acids and reaction of the highly nucleophilic sulphur in cysteine residues (Chalker et al., 2009). The prevalence of these functional groups in biology makes them useful but is also their main limitation due to a lack of specificity and selectivity.

The absence of specificity in these reactions has led to the development of bioorthogonal chemistries, reactions that are compatible with aqueous environments and react specifically with their coupling partner even in the presence of the other functional groups observed in nature (Fig. 5.1)

These bioorthogonal chemistries are typically based on classic synthetic reactions and commonly follow the rules set out by Kolb *et al* (2001) in their definition of a click reaction. The reactions must be:

- Modular

- Wide in Scope

- High Yield

- Stereo and Regiospecific

- Easily purified from inoffensive by products

Bioorthogonal reactions often utilise functional groups which are unknown in nature for example azides (Fig. 5.2, C) and alkynes (Fig. 5.2, B and D). The reaction of an azide and an alkyne was first published in 1893 by Arthur Michael (Michael, 1893) before it was shown to occur via a [4 + 2] cycloaddition mechanism by Rolf Huisgen in 1963 (Huisgen, 1963). In its original form the reaction does not comply with the parameters required for a click reaction due to the high temperature required to effect the transformation and the generation of two 1,2,4 triazole regioisomers (1,4 and 1,5). The copper(I) catalysed reaction, now synonymous with the click reaction, was published simultaneously in 2002 by Sharpless and Meldal (Rostovtsev et al., 2002; Tornøe et al., 2002) and allowed for the

generation of a single 1,2,4 triazole regioisomer, 1,4, (Fig. 5.2, A) at room temperature in an aqueous environment. This reaction is by far the most common bioorthogonal reaction and is commonly termed a 'click' reaction but has recently been relabelled Copper Azide Alkyne Cycloaddition, CuAAC, to avoid confusion with other reactions.
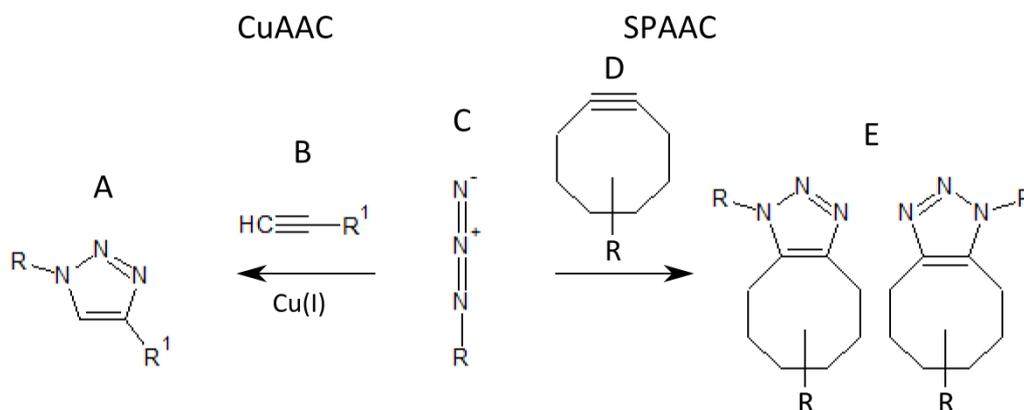
Figure 5.2   The differences between CuAAC and SPAAC
The reaction of an azide with an alkyne in the presence of a copper catalyst at room temperature, CuAAC, gives rise to only the 1,4-substituted 1,2,3-triazole (left). The incorporation of ring strain into the alkyne molecule creates a higher energy substrate which reacts with the azide at room temperature without catalysis but gives a mixture of regioisomers.

Concerns over the toxicity of the catalytic Cu(I) species to biological systems has led to the development of a cycloaddition reaction which does not require catalysis. By incorporating the alkyne component into an eight-membered carbocycle (Fig. 5.2, D), which is higher energy than a straight chain alkyne, allows the original Huisgen cyclisation to occur at room temperature, in water and without catalysis. This Strain Promoted Azide-Alkyne Cycloaddition (SPAAC) has been known since the 1960's (Wittig and Krebs, 1961) but with the current interest in bioorthogonal processes has undergone a renaissance. In this biological context it was first used by Agard et al to label azide modified cell surface glycans (Agard et al., 2004) but until recently (Gröst and Berg, 2015) it has not fulfilled all the requirements of a click reaction as like the original thermally driven reaction it gives rise to a mixture of regioisomers (Fig. 5.2, D), although its ease of use has not stopped its utilisation (Fig. 5.2).

To utilise these bioorthogonal chemistries the appropriate functionality needs to be incorporated into the molecule of interest whether that be *in vivo* or *in vitro*. The methods for incorporation of these unnatural functionalities into biomolecules can be split into 4 main groups: Genetic, Chemical, Expressed Protein Ligation and Enzymatic.

Genetic – rDNA technology (see Fig. 3.1) is generally used to generate recombinant proteins and has enabled, within the confines of natural amino acids, the ability to modify proteins with functionality that aide in their isolation and alter their physical properties.
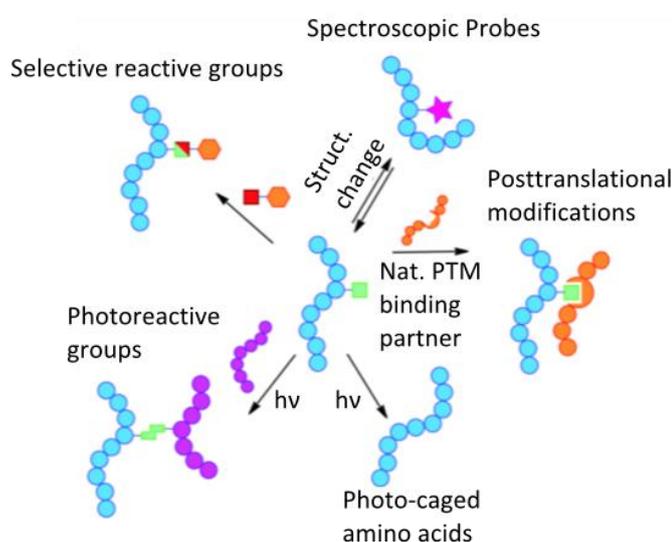
Figure 5.3 **Schematic showing Potential Uses of the Genetic incorporation of Unnatural Amino Acids**
The incorporation of an unnatural amino acid into a protein, see centre picture, can be utilised to probe the reactivity, interactions and activity of proteins. Adapted from (Dumas et al., 2014).

By recoding rare codons, for example AGG (Lee et al., 2015), an amber stop codon, TAG (Chin et al., 2003), or even incorporating a quadruplet codon (Anderson et al., 2004; Neumann et al., 2010) it is possible, by supplementing the cell with the genetic information required for the synthesis of an engineered tRNA/tRNA synthetase pair, to site specifically incorporate unnatural functionality into recombinantly produced proteins (Dumas et al., 2014; Wang et al., 2009a; Xie and Schultz, 2005) (Fig. 5.3).

Chemical – As discussed previously in this chapter (5.1) the unspecific and unselective reaction of the functionality in biomolecules can give rise to a mixture of labelling products. By genetic incorporation or utilisation of a pre-existing highly reactive and low abundance amino acid (Pe'er et al., 2004), such as cysteine, it is possible to achieve more selectivity in labelling procedures. Cysteine for example reacts quickly with many electrophilic reagents which other cellular functionality react slowly with or not at all for example maleimides and α-iodoacetyl functional groups. It is possible to use this reactivity to generate materials which incorporate unnatural functionality by utilising compounds that contain traditional reactive partner and unnatural functionality, so called adaptor reagents, (Chalker et al., 2009; Thompson et al., 2014). When utilising cysteine mutants this process is essentially site selective but suffers from the propensity of thiol groups to oxidise to form a disulphide bond which is no longer reactive towards electrophiles.

Expressed Protein Ligation – The formation of a peptide (amide) bond during translation involves the reaction of the primary amine of the amino acid in the A site of the ribosome with the thioester linkage betwen the peptide chain and the tRNA in the P site of the ribosome. The thioester acts as an activated ester which allows the amine to attack the electrophilic carbon centre of the thioester. The process of protein splicing also utilises the formation of thioester linkages to combine two protein fragments together with excision of an intein. During splicing a cysteine residue at the N terminus of the intein attacks the neighbouring amide bond in the protein to give a thioester which then exchanges with a thiol in the second extein before the intein is eliminated and acyl groups shifts from sulphur to nitrogen, S-N acyl shift, yeilding the spliced protein (Xu and Perler, 1996). Expressed Protein Ligation (EPL) utilises an intein fusion which is not followed by a second cysteine containing extein to promote its elimination from the protein. Disruption of the splicing enables the process to be commandeered by exogenous modified thiols and allows for a

site specific modification of the protein at either terminus (Jagadish et al., 2013; Muir et al., 1998).

Enzymatic – The modification of proteins in vivo is central to many cellular processes from regulation (phosphorylation) to protein decomposition (ubiquitinylation) but the appendage of these molecules to the protein is catalysed by the action of enzymes such as kinases (phosphorylation) and ubiqutin transferases (ubiqutin). By hijacking this process and utilising modified substrates containing unnatural functional groups it is possible to incorporate unnatural functionality, in a site specific manner, enabling further modification. Examples of enzymes whose substrates have been modified in this way include Sortase (Schmohl and Schwarzer, 2014; Tsukiji and Nagamune, 2009), Farnesyl transferase (Rashidian et al., 2010, 2013a), Acetyl transferase (Yang et al., 2013) and glycosidases (Rashidian et al., 2013b; Wu et al., 2015).

Utilisation of these technologies has the potential for unlocking the secrets of biomolecules but also for the incorporation of functionality and properties which make them more useful to the pharmaceutical and nanotechnology industry.

5.2 Chapter Aims

The aim of this chapter is to develop a flexible, ideally unselective, chemical labelling strategy capable of attaching varied functional groups onto amyloid fibrils pre- and post-fibril assembly and then generate particle size information from DOSY NMR experimentation and image the fluorescently labelled particles.

The modification of fibrils, pre and post assembly, with altered chemical functionality allows numerous applications. This chapter will aim to develop strategies for analysis of particle size by different methods such as NMR and fluorescence microscopy. The development of solution based methods for size determination, such as NMR, is important

to supplement imaging methods like AFM where surface deposition may introduce error or not be possible as observed with the fragmentation of insulin B chain (4.2.1.2).

The large number of ^{1}H nuclei present in protein molecules makes their associated NMR spectra very complicated; this complexity can be reduced by the attachment of NMR active nuclei which are less prevalent in nature, such as ^{19}F. By attaching a ^{19}F nuclei to a protein, and utilising pulsed field gradient (DOSY) NMR it is possible to calculate the hydrodynamic radius of the material offering a further method of calculating particle size distributions.

The covalent attachment of fluorescent dyes to fibrils allows for the localisation of fibril aggregates when transferred to biological systems as well as offering an alternative method of detection.

5.3 Chemical Labelling Strategy

To enable the flexible modification of fibrils pre/post assembly and include a linker unit, a two step chemical labelling strategy was envisaged which would first utilise reactive centres contained within all proteins such as amines (lysine) and thiols (cysteine) before allowing selective modification with bioorthogonal chemistry to incorporate the final functionality. The initial labelling strategy for protein monomers is dependent on the amino acid residues contained within the protein sequence, for example the presence of reduced cysteines. Fibril assembly after this first modification would generate fibrils activated toward secondary modification whereas modification pre-assembly would give fully functionalised fibrils after polymerisation.

The popularity of bioorthogonal reactions has increased the toolbox available for the second, bioorthogonal, labelling step (Shih et al., 2014) but by far the most popular choice of bioorthogonal transformation is the [3+2] cycoladdition between an azide and an alkyne, the click reaction. The popularity of this reaction has caused a surge in the amount of

commercial reagents that are available containing either of the prerequisite functional groups and makes the click reaction an obvious choice for the second labelling step.

As the main aim of this work was the development of a labelling strategy for the incorporation of ^{19}F atoms for detection by NMR it was imperative that the ^{19}F be visible to experiment. The slow tumbling of large molecules in the time frame of NMR spectroscopy causes line broadening and eventual disappearance of signal, even for molecules far smaller than amyloid fibrils. Although signal decreases with slower tumbling proteins certain regions of protein exhibiting more conformational motility than the bulk of the protein show enhanced signal (Baldwin et al., 2008; Platt et al., 2009). To mimic this effect the first and second labelling step must incorporate a linker which confers conformational motility to the ^{19}F nuclei from the amyloid core of the protein.

Modified fibrils would then be subjected to AFM analysis, NMR analysis and fluorescent microscopy, with attention being paid to successful incorporation of the fluorescent monomers into the fibril architecture to confirm labelling integrity.

5.4   Results

5.4.1   Conjugation of Lysozyme

The commercial availability of lysozyme offers the advantage of being able to modify the protein easily in large quantities and as such was an ideal model system for chemical labelling.

The amino acid sequence of lysozyme shows the presence of six lysine residues (red) and eight cysteine residues (green) (Fig. 5.4). Initially a promiscuous labelling strategy which could label all amyloidogenic proteins was sought. Due to the relative abundance of lysine compared to cysteine in proteins and that the cysteine residues contained within lysozyme

are part of disulfide bonds, a lysine modification strategy was selected for the initial

modification.

```
KVFGRCELAA AMKRHGLDNY RGYSLGNWVC AAKFESNFNT QATNRNTDGS
TDYGILQINS RWWCNDGRTP GSRNLCNIPC SALLSSDITA SVNCAKKIVS
DGNGMNAWVA WRNRCKGTDV QAWIRGCRL
```

Figure 5.4 **The Amino Acid Sequence of Lysozyme with Reactive Side Chains Labelled**

The amino acid sequence of lysozyme with the side-chain amines (lysine) marked in red and the side-chain thiols (Cysteine) marked in green. The eight cysteine residues all form part of the secondary structure of the protein forming four disulfide bonds: C6 – C127, C30 – C115, C76 – C94, C64 – C80

5.4.1.1 Amide Formation – Amine plus NHS ester

The most common method of amine modification is by conversion to an amide through

reaction with an activated carboxylic acid. To incorporate conformational motility and

introduce a click coupling partner onto the lysozyme molecule, a polyethyleneglycol

reagent with an azide group at one terminus and an N-hydroxysuccinimide (NHS) activated

carboxylic acid at the other (Fig.5.5, 1, Compound ii, Pierce PI26131).

Treatment of lysozyme, dissolved in pH7 phosphate buffered saline, with the acylating

reagent gives rise to a distribution of acylation products ranging from a single modification

to multiple additions (Fig. 5.5, B). Incorporation of too many groups onto the protein

structure resulted in the formation of less soluble material in subsequent steps. The use of

1.1 to 1.3 molar equivalents of the acylating reagent leaves some residual unmodified

protein but yields mostly monoacylated material with small amounts of bis-modified

material as determined by ES(MS) (Fig. 5.5, B).

The incorporation of azide functionality onto lysozyme molecules for click labelling still

leaves the choice of both CuAAC and SPAAC click derivatives as potential labelling methods.

The conversion of lysozyme monomer to fibrillar material requires heating in pH2 HCl,

conditions which are unsuitable to many of the commercial cyclooctyne reagents required

for SPAAC labelling leaving CuAAC as the best method for secondary labelling in this

instance.
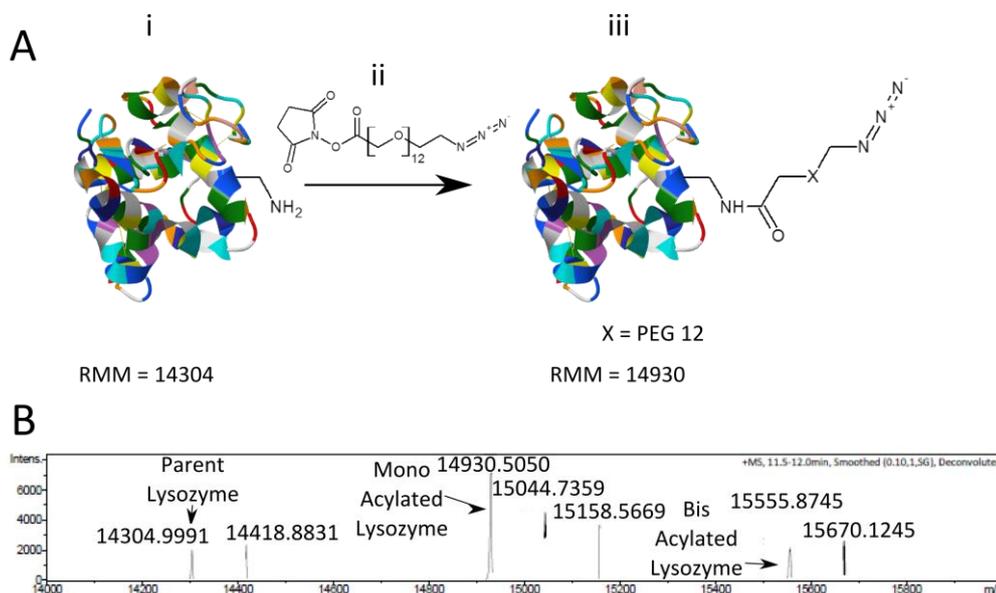
Figure 5.5   Reaction Scheme Showing Modification of Lysozyme with NHS PEG N_3 and MS Analysis of Product

A - Cartoon scheme showing an example modification of one lysine residue in the lysozyme, PDB 4RLM, structure (i) with an NHS activated PEG azide reagent (ii, Pierce PI26131) to give the azide modified product (iii). B – Deconvolution of ES(MS) spectra of lysozyme modified with 1.3eq of acylating reagent in pH7 phosphate buffer. The deconvoluted spectra shows m/z, where z=1 and therefore each value is a molecular weight of a product in solution.

5.4.1.2   Fluorescent Labelling of Lysozyme

CuAAC conversion requires the insertion of catalytic Cu(I) species into the acetylene C-H

bond before reaction with the azide functionality (McKay and Finn, 2014). Typical reaction

conditions to effect the transformation utilise a Cu(II) salt which is reduced *in situ* by

sodium ascorbate or triscarboxyethylphosphine and the incorporation of a tristriazole

ligand which stabilises the Cu(I) species (McKay and Finn, 2014; Presolski et al., 2011). To

investigate the conditions required for successful CuAAC, a fluorescent alkyne synthesised

according to (Xu and Ye, 2011) was utilised (Fig. 5.6, A, Compound ii). Successful

incorporation of the fluorescein moiety could not be confirmed by a decrease in

electrophoretic mobility on Tris-Tricine SDS PAGE (Fig. 5.6, B) but fluorescent scanning of the gel shows the attachment of fluorescent functionality to the protein (Fig. 5.6, C).
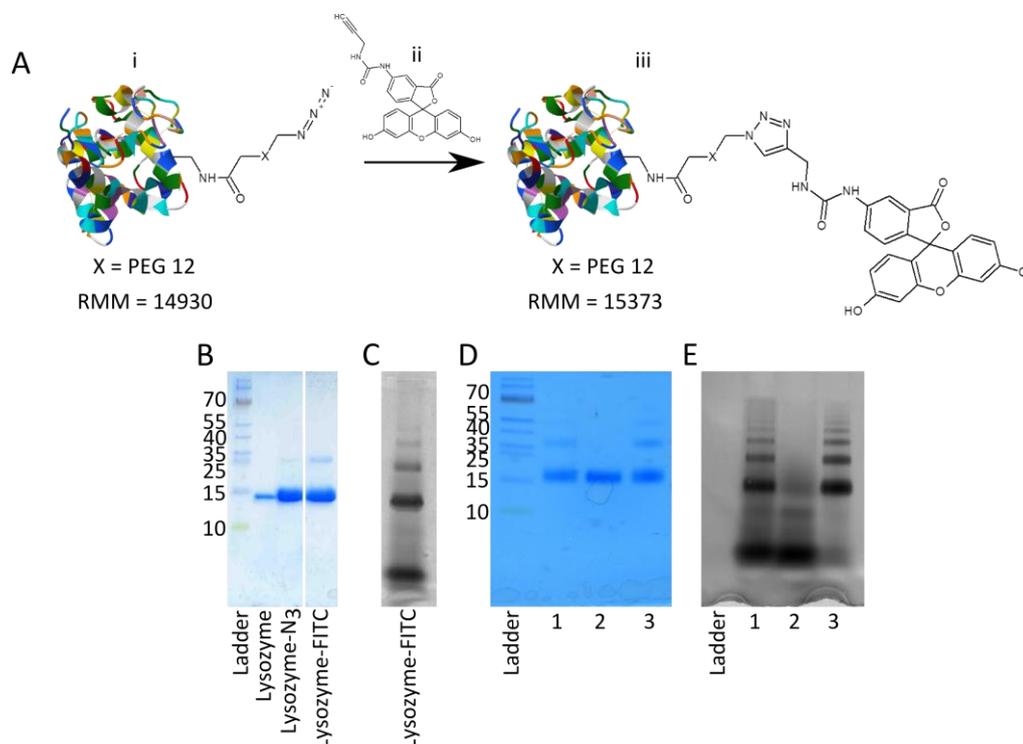
Figure 5.6   **Reaction Scheme of Click Reaction with FITC Alkyne and SDS PAGE Analysis**

A - Cartoon scheme showing modification of azido lysozyme with the FITC alkyne reagent. B – SDS PAGE Tris-Tricine gel coomassie stained showing modification of lysozyme and then click reaction with FITC alkyne. C – Fluorescence analysis of the same gel. D – SDS PAGE Tris Tricine gel which is coomassie stained. 1 Unpurified click reaction with FITC alkyne, 2 – Lysozyme-N$_3$ and FITC alkyne in the absence of Cu(I), 3 – Purification of click reaction with FITC alkyne by ethanol precipitation. E – As D but with the gel is fluorescently scanned.

Fluorescent scanning also shows the aggregation of lysozyme into dimeric, trimeric species.

Confirmation that the dye labelling was via a Cu(I) catalysed process and not a non-specific interaction was confirmed by combining the lysozyme azide and FITC alkyne in the absence of copper (Fig. 5.6, D and E, lane 2). The modified lysozyme material was purified from excess FITC alkyne by ethanol precipitation. Electrospray mass spectrometry analysis did not give the correct mass ion for the product but in fact gave a fragment ion, 14984Da, where the fluorescein isothiocyanate residue had been eliminated from the molecule, leaving the triazolyl methylamine attached to the protein.

The inclusion of lysozyme FITC conjugates into fibrillar species was achieved by incorporating the fluorescent material into a fibril formation with the unmodified lysozyme monomer in a ratio of 1:12, such that the total protein concentration was 699µM.
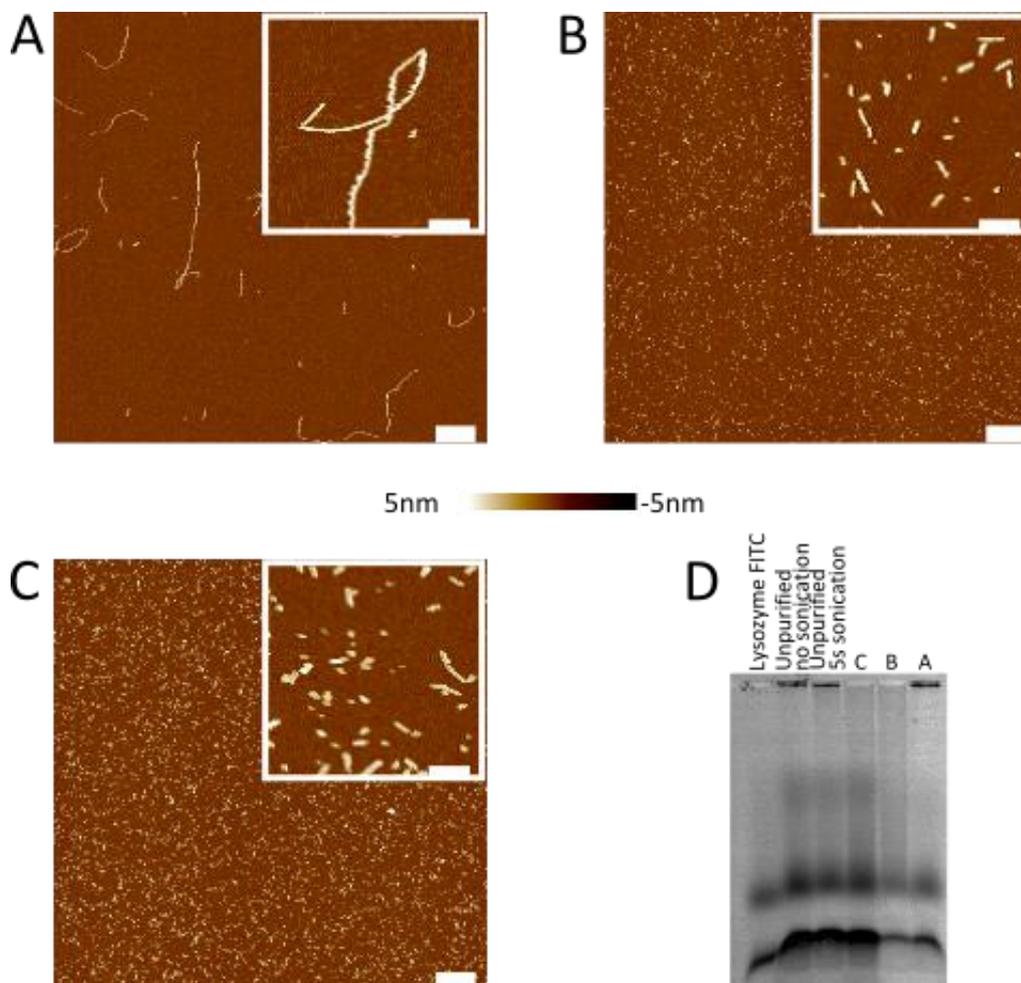
Figure 5.7   AFM Analysis of Fluorescent Fibrils
A – 10µm x 10µm AFM image at 1024 x 1024 pixels resolution with 1µm scale bar and an insert of 1µm x 1µm magnified x 5 with a 200nm scale bar showing purified unfragmented fluorescent fibrils. The height scale of the image is set between -5 and 5nm. B - 10µm x 10µm AFM image at 1024 x 1024 pixels resolution with 1µm scale bar and an insert of 1µm x 1µm magnified x 5 with a 200nm scale bar showing purified fluorescent fibrils sonicated at 20% amplitude for 1min with 5 second bursts. The height scale of the image is set between -5 and 5nm. C – 10µm x 10µm AFM image at 1024 x 1024 pixels resolution with 1µm scale bar and an insert of 1µm x 1µm magnified x 5 with a 200nm scale bar showing unpurified fluorescent fibrils sonicated at 20% amplitude for 1min with 5 second bursts. The height scale of the image is set between -5 and 5nm.   D – Semi-denaturing detergent agarose gel electrophesis of A, B and C compared to lysozyme FITC conjugate and unpurified material, imaged by fluorescent scanning.

Utilising conditions shown to polymerise unmodified lysozyme, heating to 60°C for 6 days in acidic medium, the mixture was converted to its amyloid form. Confirmation that the conditions had led to the formation of fibrillar material was achieved by AFM analysis of the reaction (Fig. 5.7, A). AFM analysis of the reaction confirms the presence of fibrils from this reaction but not of the fluorescence of the fibrils. To confirm incorporation of the fluorescent material into the fibrillar species, semi-denaturing detergent agarose gel electrophoresis (SDD-AGE) was undertaken, showing that not all material was incorporated into the fibrillar structure, even after purification. However, the fluorescence of material which did not enter the gel fully is evidence of incorporation of fluorescent material into large aggregates (Fig. 5.7, D lane unpurified no sonication). As the fibrillar aggregates (Fig. 5.7, A) are too large to enter the gel, the materials were subjected to sonication for 1 minute to give smaller fibril fragments (Fig. 5.7, C). The sonicated particles were then able to enter the gel and gave rise to a smear on the SDD-AGE (Fig. 5.7, D, lane C).

To clean the material for SDD-AGE analysis, the fibrillar material was purified by centrifugation of the sample at 75k rpm for 15 mins, removal of the supernatant and then resuspension of the pellet in fresh pH2 HCl. This method was an effective means of purifying the fibrillar material from smaller molecules although modified monomeric material could still be observed in these samples (Fig. 5.7, D lane A). AFM analysis showed that purification did not alter the parent fibrils (Fig. 5.7, A) and that the sonication conditions also afforded the same materials (Fig. 5.7, B) as those that weren't purified (Fig. 5.7, C). SDD AGE analysis of the purified unfragmented fibrils did not enter the gel (Fig. 5.7, D, lane A) and sonication allowed entry to the gel and gave a smear through the gel (Fig. 5.7, D, lane B)

The confirmation of the incorporation of the fluorescent material into the fibrillar species gave confidence in the process so that it could be transferred to an NMR active labelling strategy.

5.4.1.3   [19]F Labelling of Lysozyme for NMR Studies

NMR has become a standard biochemical tool due to the prevalence of NMR active elements such as hydrogen, carbon, nitrogen and phosporous in the structure of biolmolecules. The hydrogen nucleus is the only one of these elements whose naturally abundant isotope, ^{1}H, is NMR active. As all protons in a molecule give rise to a signal the complexity of a 1D ^{1}H spectrum for a biomolecule is extreme. This complexity can be reduced by combining experiments with less abundant nuclei such as ^{15}N and ^{13}C after the material is produced in an isotopically enriched form. Utilising less abundant nuclei increases the cost and complexity of sample preparation but provides valuable structural information by simplifying the spectra and showing molecule connectivity. Fluorine has a single isotope (^{19}F) which is NMR active and not commonly found in biological systems making it an ideal nuclei for simplifying spectra whilst, depending on method, still giving structural information (Marsh and Suzuki, 2014). The attachment of conformationally flexible protein domains (Baldwin et al., 2008) or caps (Platt et al., 2009) to proteins has allowed NMR to be used to monitor the diffusion of amyloid fibrils. The diffusion of molecules in solution is inherently linked to their size, where smaller molecules diffuse faster than larger ones. The rate of diffusion is therefore a characteristic of a molecule and is termed the Diffusion coefficient, D. The relationship between size/hydrodynamic radius (R_h) and the diffusion coefficient of spherical materials is described by the Stokes-Einstein equation (Equation 5.1).

Diffusion-ordered spectroscopy is a solution NMR technique which utilises a pulsed field gradient, Fig. 5.8, to separate components of a mixture according to their diffusion

coefficient (Morris, 2007). As the nuclei of molecules act as tiny magnets when they are placed into a magnetic field they will first align with the magnetic field and then move towards the magnet. By subjecting the sample to increasing magnetic field strength smaller materials will move towards the magnet faster than larger materials. This difference in diffusion means that the intensity of the peaks associated with the materials will decrease at different speeds.
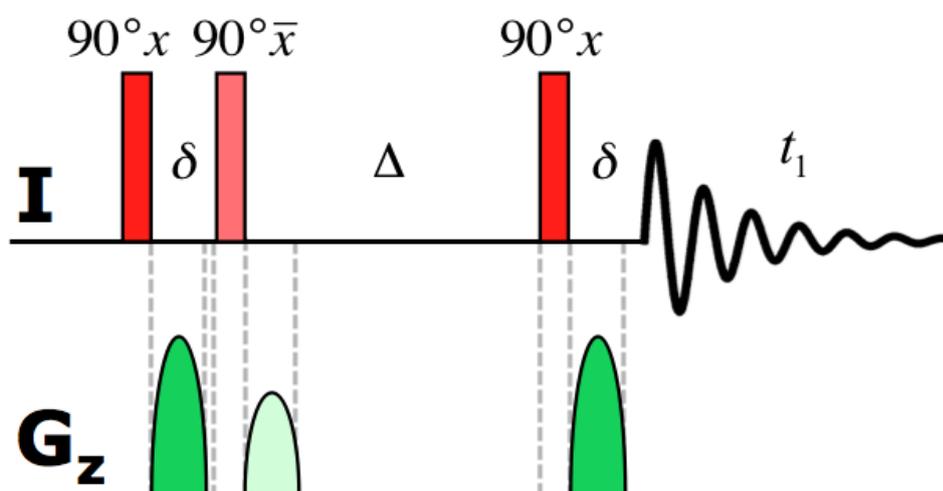
Figure 5.8   Example of a stegp1s1d DOSY NMR pulse sequence
A magnetised sample is excited by 90° radiofrequency bursts (Top) and also subjected to magnetic field bursts (Bottom) of different powers for defined duration of time, (δ). The delay between the magnetic field gradient bursts (Δ) allows diffusion of molecules. Smaller molecules diffuse faster than large molecules which means that the signal intensity for large molecules is maintained through the gradient whereas the signal for small diminishes. http://www.chemie.uni-hamburg.de/nmr/insensitive/tutorial/en.lproj/stegp1f.html

The decrease in observed signal intensity during a pulsed field gradient NMR experiment is related to the diffusion coefficient, D, by the Stejskal-Tanner equation (Equation 5.2) (Stejskal and Tanner, 1965).

Equation 5.1   Stokes-Einstein Equation
Where D is the Diffusion coefficient, κ = Boltzmann constant, T = Absolute Temperature (K), η = Viscosity and R_h = Hydrodynamic radius.

$$D = \frac{\kappa T}{6\pi\eta R_h}$$

Equation 5.2   Stejskal-Tanner Equation

Where I = Intensity, I_o = Total Intensity, γ = gyromagnetic ratio of the nuclei studied, δ = the pulse duration, Δ = the interval between pulses, g = Magnetic field gradient strength and D = the Diffusion Coefficient.

$$I = I_0 \times e^{\left(-\gamma^2\delta^2 g^2 D\left(\Delta - \frac{\delta}{3}\right)\right)}$$

The decay in NMR signal, I, as magnetic field gradient increases can be calculated and plotted against the magnetic field gradient strength, g, using a simplified Stejskal-Tanner equation (Equation 5.3).

Equation 5.3   Simplified Stejskal-Tanner Equation

Where c = a constant for the experimental conditions and C is the value extracted from fitting the Stejskal-Tanner equation to the acquired data.

$$I = I_0 \times e^{(-Cg^2)}$$
$$C = (\gamma^2\delta^2(\Delta - \frac{\delta}{3})) \times D$$

Equation 5.3 allows for the deconvolution of the constant C using internal standards of known diffusion coefficient, D. With an internal standard which decays to give a known diffusion coefficient such as dioxane, $D_{Dioxane}$, it is possible to calculate the constant c (C=c*D) and therefore enables characterisation of the diffusion coefficient of the other materials in the sample. Incorporation of a second reference material containing both hydrogen and fluorine atoms, for example trifluoroethanol (TFE), allows the use of $D_{Dioxane}$ to calculate the diffusion coefficient of unknown fluorinated materials.

The incorporation of a ^{19}F containing group allows the diffusion of fibrils to be monitored using far simpler spectra than that obtained using ^{1}H NMR. To maximise the signal observed in the ^{19}F NMR a commercial label that contained six electronically equivalent fluorine atoms, 1-ethynyl-3,5-bis(trifluoromethyl)benzene - $(CF_3)_2$ alkyne (Fig. 5.9, 1, Compound B) was used.
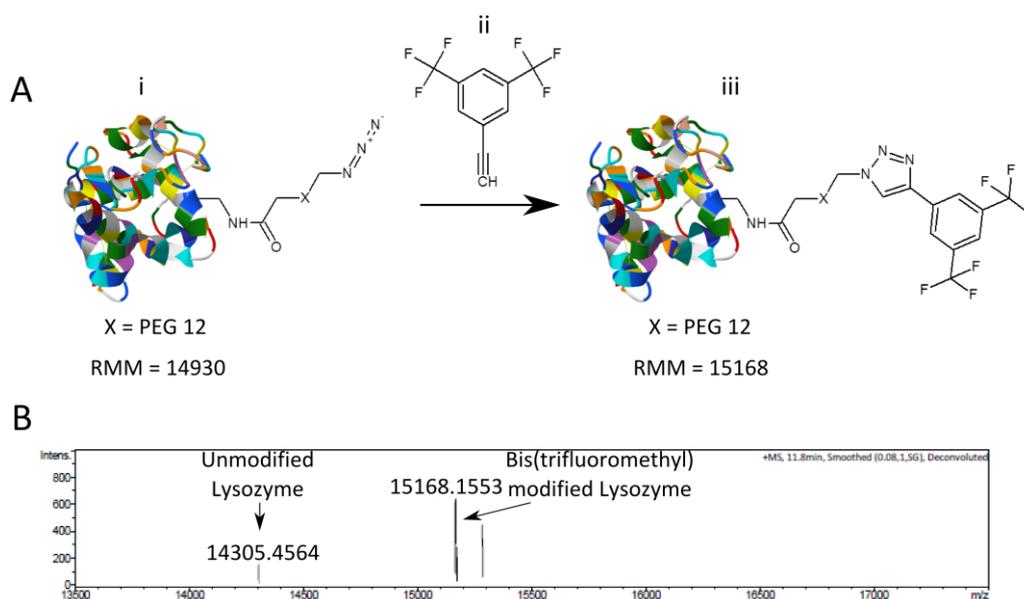
Figure 5.9   Click Modification of Lysozyme Azide with 1-ethynyl-3,5-bis(trifluoromethyl)benzene

A – Cartoon scheme showing the modification of lysozyme azide (i) with 1-ethynyl-3,5-bis(trifluoromethyl)benzene (ii) to give an ^{19}F NMR active lysozyme (iii). B – Electrospray analysis of the conversion showing some unmodified lysozyme remaining but with the replacement of the azide lysozyme with the bis(trifluoromethyl) compound.

By using the reaction conditions previously identified for successful modification of lysozyme azide with fluorescein alkyne the $(CF_3)_2$ alkyne was substituted into the reaction (Fig. 5.9, A). Confirmation of ligation was achieved by ES(MS) (Fig. 5.9, B) as no change in band position was observed by SDS PAGE analysis. The modified protein was converted into a fibrillar state using conditions previously confirmed using the fluorescent analogues (5.4.1.2). A ratio of modified to unmodified monomer of 1:5 was employed in the fibril forming experiment. An overall protein concentration of 699µM was heated to 60°C in pH2 HCl for 6 days. The successful formation of amyloid fibrils was determined by AFM analysis and showed formation of similar fibrils to the unmodified fibrils (Fig. 5.9).

To maximise the potential for successful DOSY analysis of amyloid fibrils, the size of the parent fibril was reduced by subjecting the fibril solution to fragmentation by stirring for 45 hours at 1000rpm using an IKA Squid and confirmed by AFM analysis (Fig. 5.10). The same fragmentation was undertaken for unlabelled fibrils for comparison (Fig. 5.10).
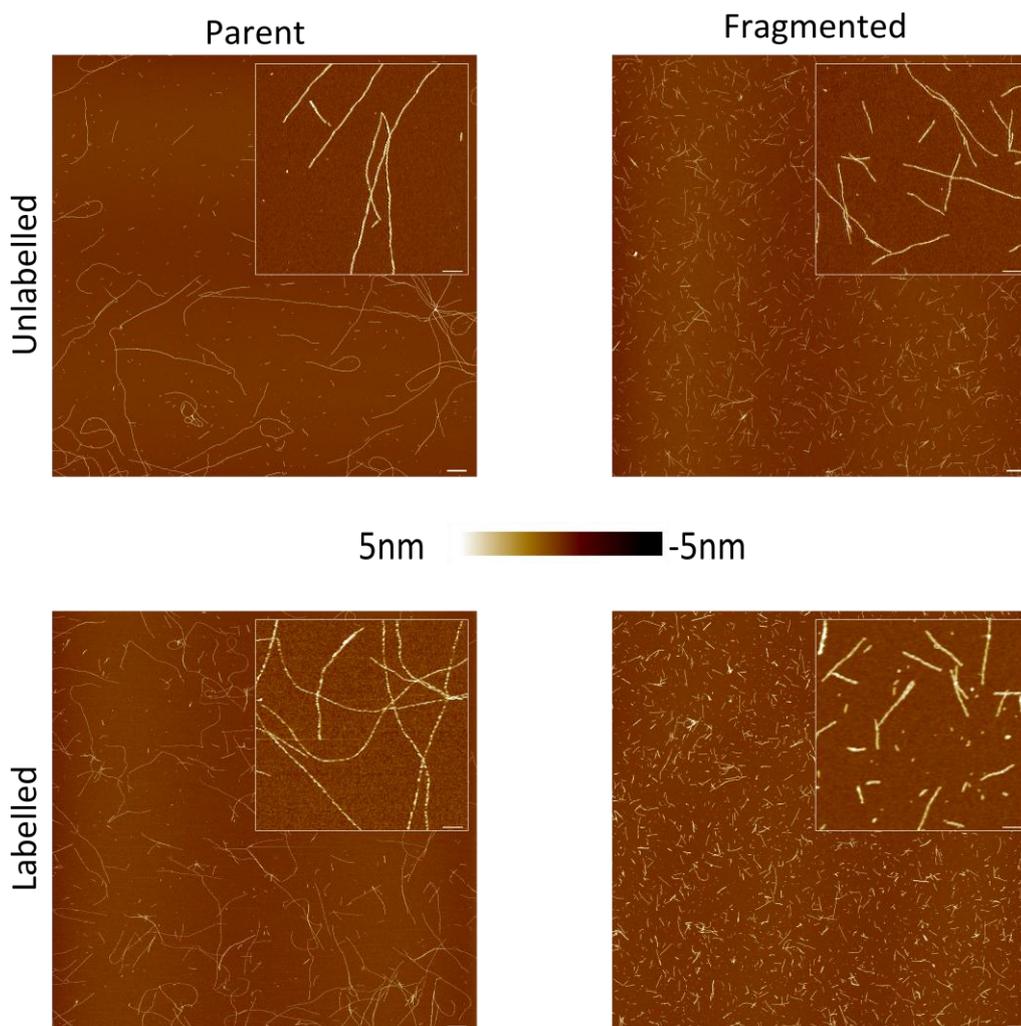
Figure 5.10   AFM Analysis of [19]F Labelled Fibrils and Unlabelled Fibrils
20μm x 20μm scans at 2048 x 2048 pixel resolution with 1μm scale bar and a 2μm[2] insert magnified x5 with 200nm scale bar. The height scales for the images are set between -5 and 5nm. Top – Unlabelled fibrils: Left – Parent fibrils, Right – Fragmented fibrils. Bottom – [19]F labelled fibrils: Left – Parent fibrils, Right – Fragmented fibrils.

The fragmented fibrils were then treated with D_2O (for locking of the NMR signal), dioxane (^{1}H internal standard) and TFE (^{19}F internal standard) before NMR analysis. ^{1}H NMR spectra of the fibrillar and monomeric samples show a change in the peak distribution across the spectra and therefore the structure of the protein contained in the sample (Fig. 5.11). Comparison of the unmodified lysozyme fibrils with modified lysozyme fibrils shows no change although this is not surprising as the modified material only makes up 20% of the total sample.
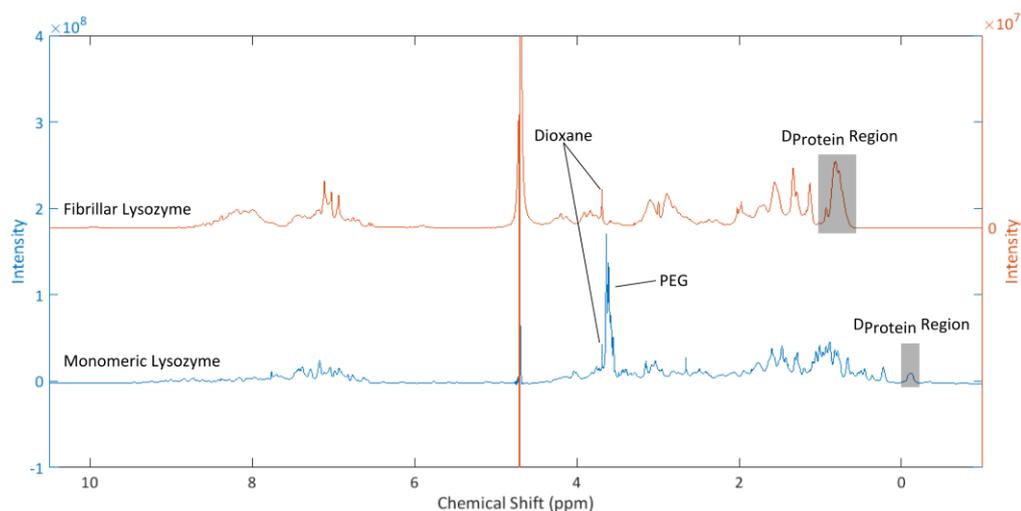
Figure 5.11 ^{1}H NMR Comparison of Modified Fibrillar and Monomeric Lysozyme

The Lower spectra shows the proton spectra for modified monomeric lysozyme which has been non-specifically modified with NHS-PEG N$_3$ (516μM) and the upper spectra shows the modified monomer incorporated into a fibril formation with unmodified fibrils (699μM). The D$_{Protein}$ region marked in both spectra are the regions which were analysed during the magnetic field gradient.
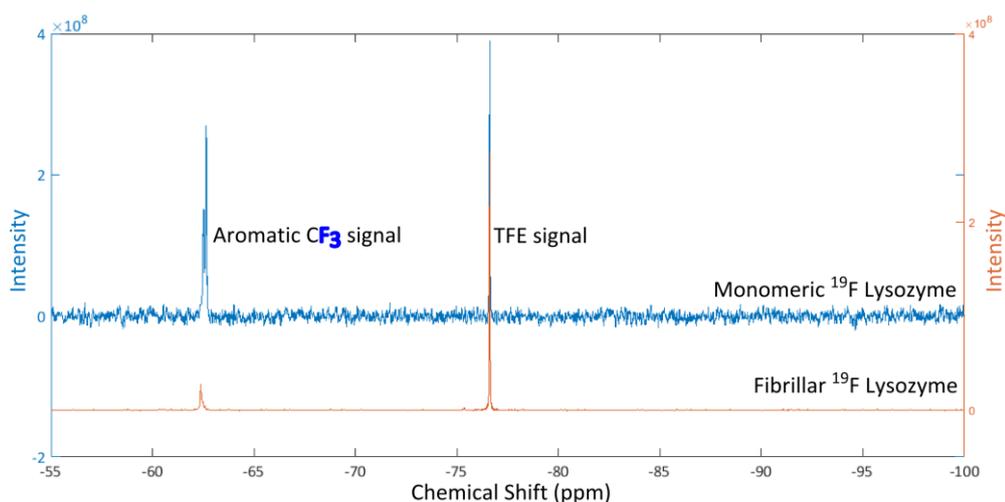
Figure 5.12 ^{19}F NMR Analysis of Modified Monomeric and Fibrillar Lysozyme

The higher spectra shows the monomeric ^{19}F labelled lysozyme where as the lower spectra is of fibrils containing 20% labelled monomer.

It is important that the concentration of the internal reference standard is sufficient for the DOSY analysis. At a concentration similar to the protein concentration the dioxane (3.75ppm) (Fig. 5.11) signal is visible and the decay measurable but the TFE signal (3.88ppm) is completely obscured. Due to the sample signal intensity observed in the ^{19}F spectra (Fig. 5.12) increasing the TFE concentration was not feasible. Replacement of the

TFE with a different internal standard with ^{1}H signals in a region not normally occupied by protein protons, 4,4,4-trifluorobut-2-enol (TFBenol), avoided the problem of ^{1}H peak visibility (Fig. 5.13, higher).
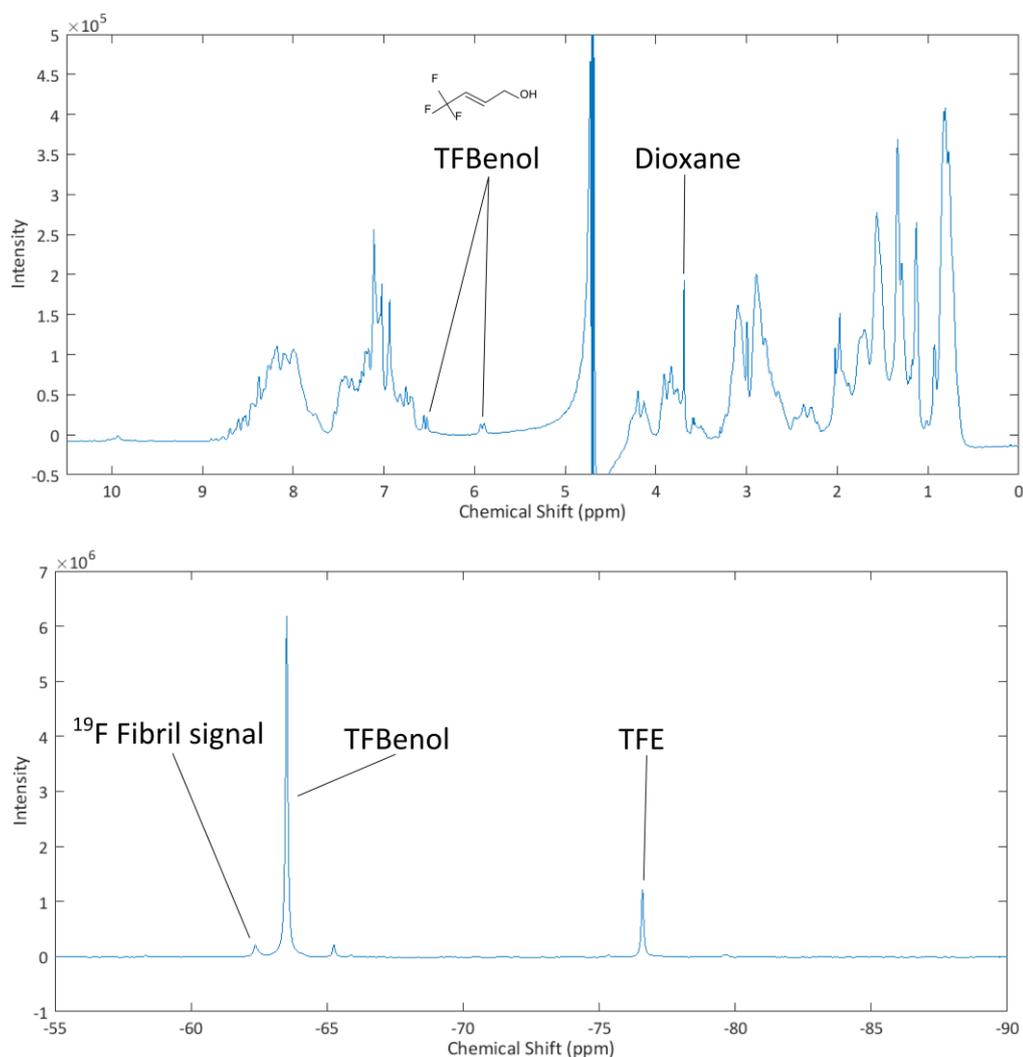
Figure 5.13   [1]H and [19]F Spectra of Labelled Lysozyme Fibrils with TFBenol Reference Sample

[19]F labelled fibrils (699µM) containing TFBenol, TFE and dioxane references were analysed by NMR spectroscopy Top – ^{1}H NMR spectrum. Bottom – ^{19}F NMR spectrum.

The ^{19}F spectra for these materials show peaks in the appropriate region for an aromatic

CF3 species (reference trifluorotoluene: -63.72ppm) (Fig. 5.12 and 5.13, lower).

Figure 5.14 Decay of Lysozyme [19]F Signal In Magnetic Field Gradient

A – A plot of I/I_o against Magnetic Field Gradient Strength showing the decay of protein signal and internal standards: dioxane, TFBenol and TFE.

B – A plot of $\ln(I/I_o)/\alpha^2\beta$ against G^2 showing the decay of protein signal and internal standards: dioxane, TFBenol and TFE

1 – Decay of ^{19}F signal of 1 in 5 labelled Lysozyme fibrils containing TFBenol and TFE.

2 – Decay of ^{19}F signal of 1 in 5 labelled Lysozyme fibrils containing TFE.

3 – Decay of ^{19}F signal of monomeric modified Lysozyme containing TFE.

Figure 5.15 Decay of Lysozyme ^{1}H Signal In Magnetic Field Gradient

A – A plot of I/I_o against Magnetic Field Gradient Strength showing the decay of protein signal and internal standards: dioxane, TFBenol and TFE.

B – A plot of $\ln(I/I_o)/\alpha^2\beta$ against G^2 showing the decay of protein signal and internal standards: dioxane, TFBenol and TFE

1 – Decay of ^{1}H signal of 1 in 5 labelled Lysozyme fibrils containing TFBenol and dioxane.

2 – Decay of the ^{1}H signal of modified containing Lysozyme monomer containing dioxane.

3 – Decay of ^{1}H signal of unmodified Lysozyme fibrils containing dioxane.

**Table 5.1   Table Comparing Extracted Values of C with Derived Diffusion Coefficient Values $D_{Particle}$*

^{1}H					Int. Ref.	$\ln(I/I_o)/\alpha^2\beta$	Data points
Monomer	I_o	C	DELTA	delta	$D_{Particle}$	$D_{Particle}$	
Dioxane	1.03E+09	8.849	50ms	2ms	1.16E-09	1.03E-10	7 Points
TFBenol	-	-	50ms	2ms	-		
TFE	-	-	50ms	2ms	-		
Protein (Monomer)	1.13E+07	4.974	50ms	2ms	6.52E-10	3.58E-10	
PEG	7.23E+07	16.44	50ms	2ms	8.06E-10	4.21E-10	
Labelled Fibrils							
Dioxane	4.46E+06	40.22	50ms	5ms	1.16E-09	4.32E-10	7 Points
TFBenol	1.44E+06	68.27	50ms	5ms	1.97E-09	6.32E-10	10 Points
TFE	-	-	50ms	5ms	-		
Protein (Fibrils)	8.76E+07	8.2	50ms	5ms	2.37E-10	9.21E-11	
Unlabelled Fibrils							
Dioxane	1.06E+07	37.51	50ms	5ms	1.16E-09	4.20E-10	7 Points
TFBenol	-	-	50ms	5ms	-		
TFE	-	-	50ms	5ms	-		
Protein (Fibrils)	2.73E+08	7.381	50ms	5ms	2.28E-10	8.25E-11	
^{19}F					Int. Ref.	$\ln(I/I_o)/\alpha^2\beta$	
Monomer	I_o	C	DELTA	delta	$D_{Particle}$	$D_{Particle}$	
Dioxane	-	-	100ms	5.4ms	-		
TFBenol	-	-	100ms	5.4ms	-		
TFE	8.48E+05	33.74	100ms	5.4ms	2.32E-09	2.41E-10	11 Points
Protein (Monomer)	1.49E+06	5.043	100ms	5.4ms	3.19E-10	2.90E-11	
Labelled Fibrils TFBenol							
Dioxane	-	-	100ms	5.4ms	-		
TFBenol	1.81E+08	35.34	100ms	5.4ms	1.97E-09	1.72E-10	
TFE	3.91E+07	41.56	100ms	5.4ms	2.32E-09	2.55E-10	12 Points
Protein (Fibrils)	1.02E+07	5.043	100ms	5.4ms	2.81E-10	2.68E-11	
Labelled Fibrils no TFBenol							
Dioxane	-	-	100ms	5.4ms	-		
TFBenol	-	-	100ms	5.4ms	-		
TFE	4.33E+07	41.44	100ms	5.4ms	2.32E-09	2.42E-10	11 Points
Protein (Fibrils)	9.79E+06	4.341	100ms	5.4ms	2.24E-10	2.36E-11	

*Table showing the C values extracted from the fitting of the Tanners Stejkland equation to the experimental data. By comparison with the known diffusion coefficient of dioxane, $D_{Dioxane} = 1.16e^{-9}$, the Diffusion coefficient of the unknown samples were calculated. Where the ^{1}H signal of the TFE signal is obscured the diffusion coefficient calculated by comparison to TFBenol was used to estimate D_{TFE}, see red text, and used to calculate the diffusion coefficient of the modified protein samples.

The TFE fluorine signal (-76.618ppm) and TFBenol reference in both ^{1}H and ^{19}F and ([1]H

6.575-6.513ppm 5.958-5.875ppm and [19]F -63.549ppm) material in the sample is obvious

(Fig. 5.12 and 5.13, lower) as is the decrease in signal intensity and change in peak shape

observed on conversion from monomeric protein to fibrillar aggregate (Fig. 5.12).

The lower intensity observed in the fibrillar sample is explained by a decrease in

conformational motility as the polymeric species is formed. The TFBenol reference gives

rise to ^{1}H peaks that are discrete from protein peaks but the ^{19}F reference peak is very close

to the sample signal due to the CF_3 group being attached to an SP2 carbon centre like in the

label.

Due to the lack of baseline resolution between the internal references and the protein

signals the value of C for dioxane was a source of error. An absence of baseline resolution

also obscured the ^{1}H TFE signal, to calculate diffusion coefficients in samples containing TFE

the D_{TFE} value calculated by comparison with the $D_{TFBenol}$ value in the ^{19}F fibril sample was

used to estimate values in other experiments (table 1, red text). To avoid this problem an

alternative method of calculating the diffusion coefficient by plotting $\left(\ln\frac{I}{I_o}\right)/\alpha^2\beta$ against

G^2, where $\alpha = \delta*\gamma$ and $\beta = \Delta-(\delta/3)$, gives a straight line with a gradient of $-D_{particle}$. The signal

observed for the reference samples is lost by 50% gradient strength therefore the line

fitted to the decay is skewed by the later data. The line of best fit is therefore only

calculated for the portion of data which gives rise to a decrease in signal (sample column of

Table 5.1).

The absolute method of analysis of ^{19}F and ^{1}H data appears to give rise to values correlating

to larger species than the comparative method in all cases see table 5.1. Whilst there

appears to be a difference in size between monomer and fibrils in the ^{1}H data this is less

apparent in ^{19}F spectra as adjudged by $D_{protein}$ values, [1]H – 3.58×10^{-10} and [19]F – 2.9×10^{-11}.

The difference between the fibrillar and the monomer samples diffusion coefficient is small and not in line with the significant change in size associated with fibril assembly.

As discussed in chapter 3 the formation of fibrillar materials from lysozyme at pH2 and 60°C is a hydrolytic process in which the peptide backbone of the protein is broken to give smaller peptide fragments. The hydrolysis of the labelled monomer during fibril formation means that the label is potentially lost from the protein during aggregation and this liberated material is free diffuse throughout the solution unrestricted by the amyloid core.

Due to the intensity of the observable ^{19}F signal the acquisition time associated with each gradient point are extremely long meaning that a full ^{19}F experiment takes about twenty four hours and is too long to be a practical in these circumstances.

The hydrolytic mechanism of fibril formation as well as the long acquisition times associated with lysozyme requires the use of another model system with alternative characteristics.

5.4.2 Conjugation of Sup35NM

Proteins which convert to fibrils at pH7 suffer from the complication that their production and purification is often in competition with their aggregation. As discussed in chapter 3 the purification of Sup35 and the truncated form Sup35NM is achieved under denaturing conditions. Due to the protection from aggregation that guanidine provides post purification and the higher expression yield, Sup35NM was an ideal substrate for chemical labelling experiments. Another aspect which makes Sup35 an ideal substrate for this experiment is the structure of the fibrillar materials. It has previously been shown that the N domain is responsible for fibril formation and that the flexible M and structured C domains are pointed away from the core into solvent (Baxa et al., 2011) and therefore potentially containing conformational motility not observed in the fibril core.

5.4.2.1 Site specific labelling – Sup35NM model system

To take advantage of the conformational motility proposed to exist in the M domain of

Sup35NM site specific modification of the C terminus of the protein was performed by

placing a cysteine residue at the end of the polypeptide.

To increase conformational motility, an extra linker between the M domain and the

cysteine was incorporated based on the amino acid sequence, GSGGSGGSG, shown to

exhibit conformational motility when attached to β_2-microglobulin (Platt et al., 2009) (Fig.

5. 16).

```
MSDSNQGNNQQNYQQYSQNGNQQQGNNRYQGYQAYNAQAQPAGGYYQNYQGYSGYQQGGY
QQYNPDAGYQQQYNPQGGYQQYNPQGGYQQQFNPQGGRGNYKNFNYNNNLQGYQAGFQPQ
SQGMSLNDFQKQQKQAAPKPKKTLKLVSSSGIKLANATKKVGTKPAESDKKEEEKSAETK
EPTKEPTKVEEPVKKEEKPVQTEEKTEEKSELPKVEDLKISESTHNTNNANVTSADALIK
EQEEEVDDEVVNDGSGGSGGSGC
```

Figure 5.16   Amino Acid Sequence of Modified Sup35NM Protein, Sup35NM LCys

The Sup35NM amino acid sequence is supplemented by the addition of a linker previously shown to exhibit conformation flexibility, GSGGSGGSG, highlighted in red and the addition of a single cysteine residue, text in purple, to the C terminus.
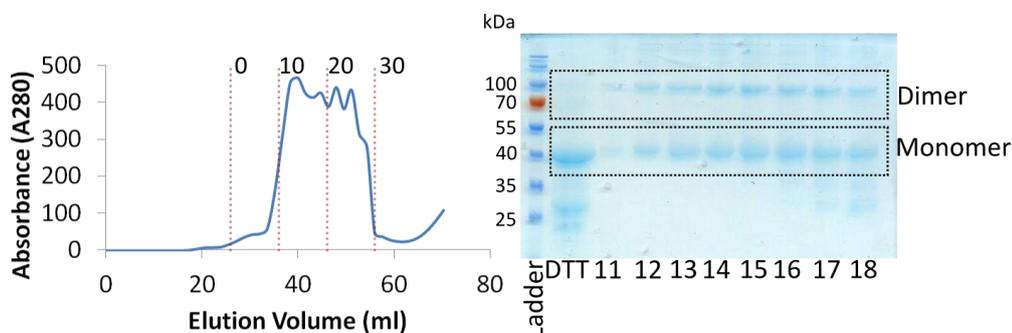
Figure 5.17   Purification Profile of Sup35NM LCys

Left – Äkta elution profile for Sup35NM LCys. The denaturing purification of all Sup35 related materials gives similar profiles with the left handside of the peak containing the purest protein.   Right – SDS PAGE analysis of the fractions eluted from the column with the monomeric protein and dimeric protein highlighted. The DTT sample is that of material eluted from the NiNTA column and then treated with DTT showing that the higher molecular weight material is reducable.

Appendage of the desired linker and cysteine to the Sup35NM sequence to give Sup35NM

LCys, was achieved by PCR amplification of the Sup35NM sequence with primers encoding

the tag. The DNA containing this sequence was then digested and ligated into an empty

pET15b vector before transformation into Top10 *E.coli* strain for plasmid replication. The

vector was then transformed into BL21 [DE3] pLysS *E.coli* for protein expression induced by

the addition of 1mM IPTG for four hours.

The expression and purification of Sup35NM LCys under denaturing conditions gave rise to

the formation of a product with approximately the correct size, 31257Da as seen on SDS

PAGE, with a by product roughly twice the size that would not be removed by further size

exclusion chromatography (Fig. 5.17). This dimeric product is formed through disulphide

formation and could be reduced to monomer by treatment with DTT.

5.4.2.2 Sup35NM LCys Azide Incorporation

RMM = 31257.3   RMM = 31515.6
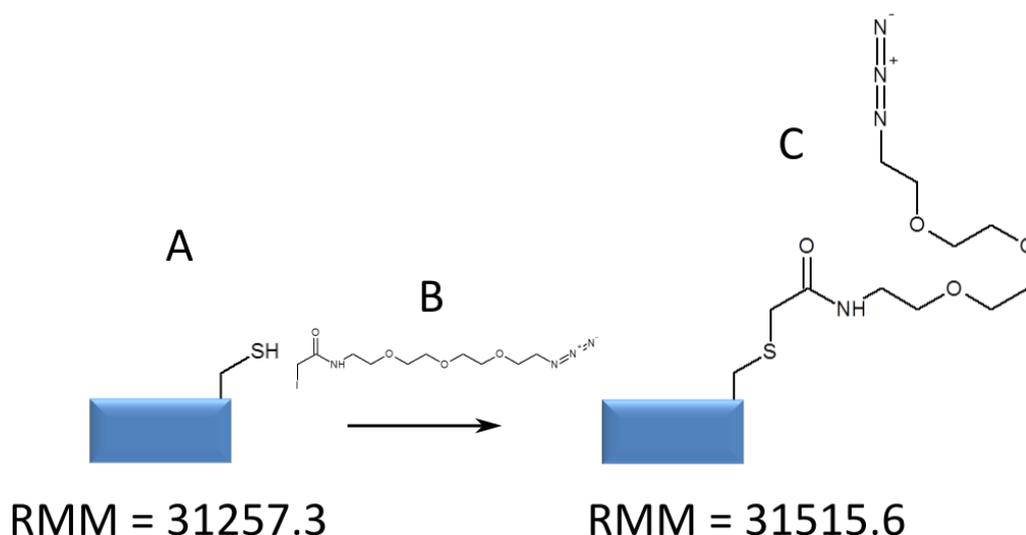
Figure 5.18   Reaction Scheme Showing the Reaction of the Cysteine residue with the Iodoacetamide Reagent
Reaction of the C terminal cysteine residue in Sup35NM LCys protein (Blue square) (A) with an Iodoacetamide derivative containing an azide (B) which incorporates the azide onto the protein (C).

Commonly used reducing agents such as dithiothreitol (DTT) and

triscarboxyethylphosphine (TCEP) are highly nucleophilic and require removal before

subsequent treatment with an electrophile. To avoid gel filtration and dilution post

reduction a resin immobilised phosphine reagent, TCEP, was used to reduce the impurity

prior to conjugation reaction.

There are two common conjugation chemistries utilised for thiol modification: thiol reaction with α-iodo acetate derivatives or addition to double bonds in conjugation with a carbonyl, such as maleimides. α-Iodo acetate reagents where chosen as the coupling partner due to their common usage under denaturing conditions for the capping of cysteines in proteomic experiments whilst maleimides propensity to ring open in the presence of some nucleophiles suggested some stability issues.

The azido functionalised α-iodoacetamide (Fig. 5.18, Compound B) used to modify the cysteine residue contained a small PEG unit to further increase conformational motility (Fig. 5.18) and was synthesised from the appropriate azido amine and N-hydroxysuccinimide activated α-iodo acetic acid.

Residual alkylating agent was removed by Äkta purification using a HiLoad^{TM} 16/600 Superdex^{TM} 200pg column in 6M guanidine based buffer which also removed some residual lower molecular weight impurities (Fig. 5.19) and the modification was confirmed by ES(MS), 31258Da.
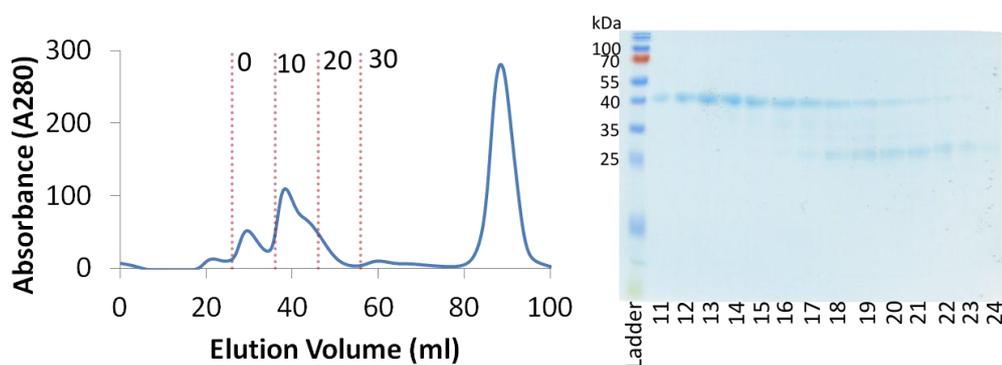
Figure 5.19   Purification of Alkylated Sup35NM LCys with Iodoacetyl N$_3$
Right - HiLoad^{TM} 16/600 Superdex^{TM} 200pg column Äkta purification profile of the alkylation reaction of Sup35NM LCys with Iodoacetyl azide reagent. Left - SDS PAGE analysis of the fractions taken from the Äkta purification showing removal of lower molecular weight impurities and the absence of dimer from the mixture. Fractions 11-15 were combined for further reaction.

5.4.2.3 Sup35NM LCys N$_3$ Modification Through Click Chemistry

Modification of the Sup35NM LCys N$_3$ was possible through CuAAC reaction similar to the

modification of lysozyme but due to the milder aggregation conditions of Sup35NM the

commercial reagents for strain promoted click chemistry (SPAAC) are also possible reactive

partners.
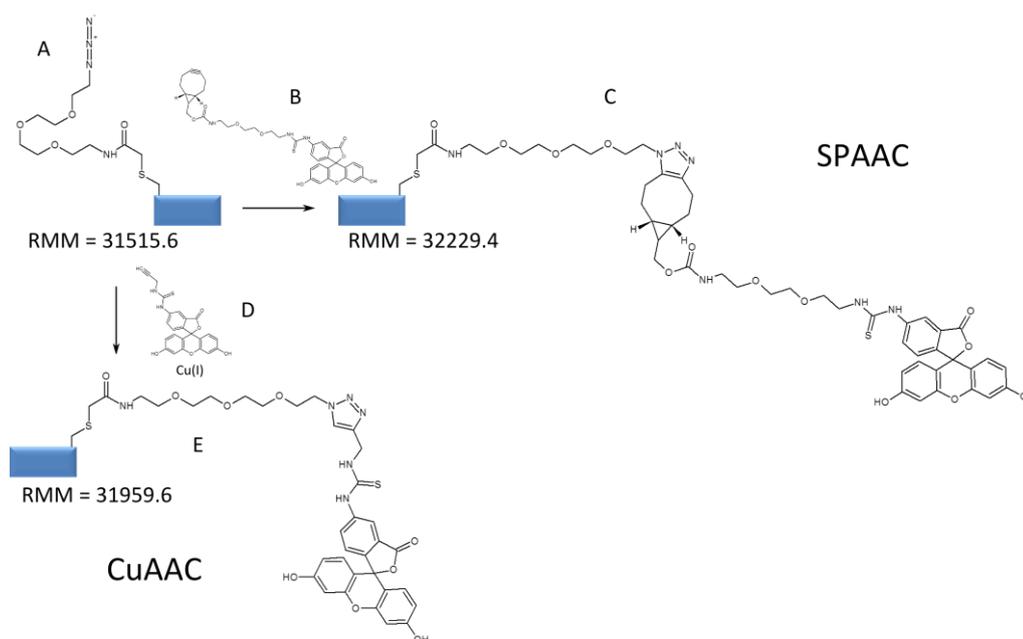
Figure 5.20   Scheme Showing the Potential Fluorescent Modification of Sup35NM LCys N$_3$ through SPAAC and CuAAC Chemistry
Reaction scheme showing the reaction of azide modified Sup35NM LCys (A) with a FITC modified with a strained alkyne (SPAAC) (B) or a terminal alkyne (CuAAC) (C) to yield fluorescent protein products.
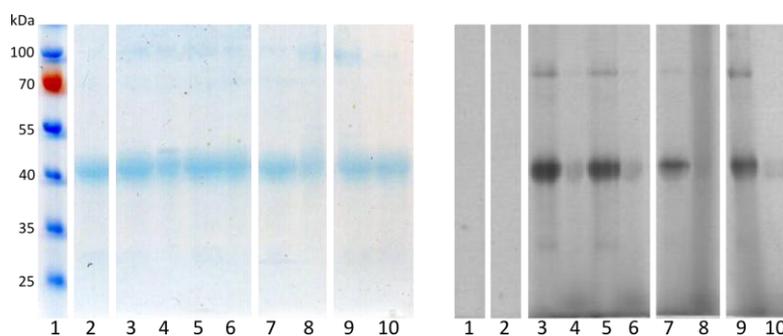
Figure 5.21   Comparison of Sup35NM LCys N$_3$ Reactivity Towards CuAAC and SPAAC Reactions.
Left – Coomassie stain of gel showing the modification of Sup35NM LCys N$_3$ with FITC Alkyne or BCN FITC for different amounts of time and reagent. Right – Fluorescent scan of the same SDS PAGE showing reaction of Sup35NM LCys N$_3$ with fluorescent reagents. Lane 1 = Ladder, 2 = Sup35NM LCys N$_3$, 3 = 10eq of BCN FITC for 1 hour, 4 = 10eq of FITC alkyne for 1 hour, 5 = 10eq of BCN FITC for 2 hours, 6 = 10eq of FITC alkyne for 2 hours, 7 = 10eq of BCN FITC for 4 hours, 8 = 10eq of FITC alkyne for 4 hours, 9 = 5eq of BCN FITC for 4 hour and 10 = 5eq of FITC alkyne for 4 hours

As with lysozyme, fluorescent click reagents were used to monitor the success of protein modification and subsequent fibril formation (Fig. 5.20). Reaction of BCN FITC (SPAAC), generated by reaction of BCN amine with FITC, (Fig. 5.20, Compound B) and FITC alkyne (CuAAC) (Fig. 5.20, Compound D) with Sup35NM LCys N_3 under the same conditions used for lysozyme were monitored over time by SDS PAGE. Coomassie staining of the gel shows no change in electrophoretic mobility for all reactions but the fluorescence scan shows that even after 1 hour the SPAAC conditions are far superior to the CuAAC conditions (Fig. 5.21). The ratio of observed signal in the SPAAC reactions to the CuAAC reactions shows that the SPAAC reactions were % more effective than the CuAAC reactions.

5.4.2.4 Sup35NM LCys N3 BCN FITC Assembly into Amyloid Fibrils

The conversion of the Sup35NM LCys N_3 BCN FITC conjugate to the amyloid state was initiated by buffer exchange into 20mM sodium phosphate buffer, 50mM sodium chloride at pH7 to remove the guanidine and any excess small molecule label from the SPAAC reaction. The assembly conditions were identical to those used for unmodified Sup35NM, 10µM protein concentration at 30$^\text{o}$C and the assembly kinetics were monitored by ThT fluorescence (Fig. 5.22).

The lag phase for unseeded polymerisation was about 10 hours similar to that observed with unmodified Sup35NM (See 3.2.1.2.2) and ^{19}F modified Sup35NM (Fig. 5.27). The increase in fluorescence intensity observed over the duration of the fibril formation was four times higher than baseline compared with unmodified Sup35NM and ^{19}F modified Sup35NM where the increase observed was 8 fold. The difference in signal is due to the emission of bound ThT occurring at 482nm which is in the absorbance range of the FITC dye conjugated to the protein causing the emission of light to be red shifted, lowering the observed signal.
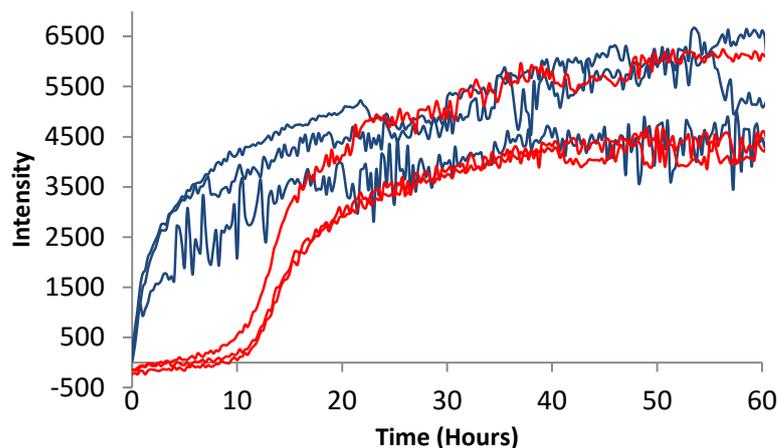
Figure 5.22   Kinetics of Sup35NM LCys N$_3$ BCN FITC Aggregation
10µM Sup35NM LCys N$_3$ BCN FITC was treated with ThT and the increase in fluorescence monitored. 3 replicates were run of each experiment and then blank corrected. Red samples were unseeded polymerisation and blue samples were seeded polymerisation with sonicated 1% Sup35NM fibrils.
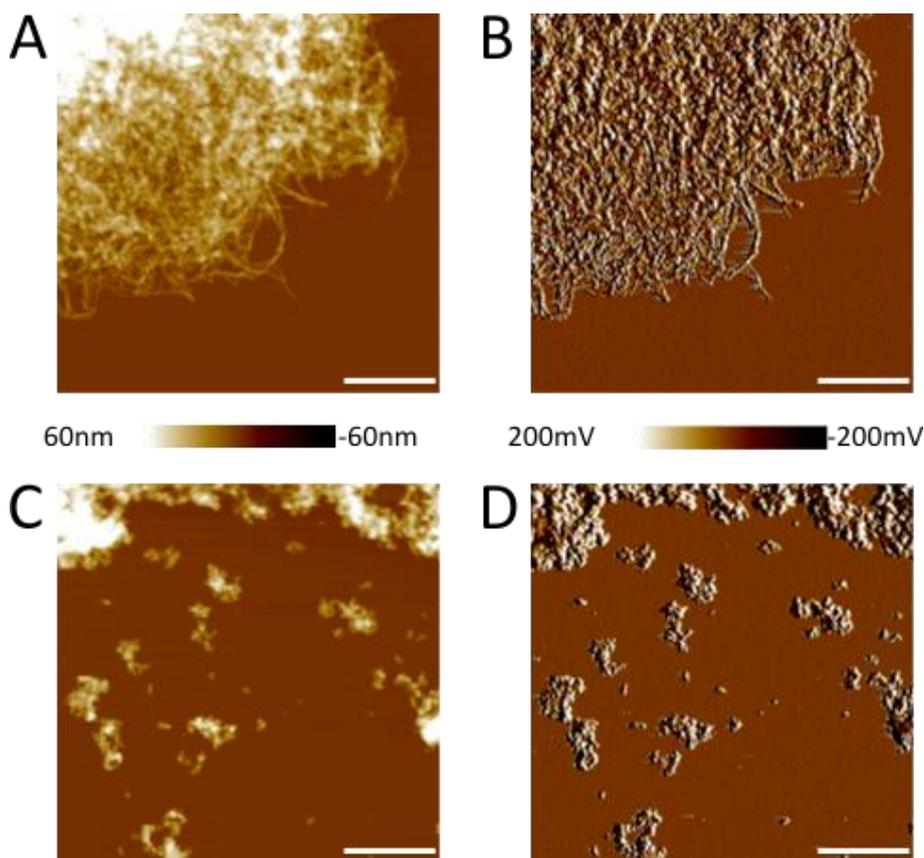
Figure 5.23   AFM Analysis of Sup35NM LCys N$_3$ BCN FITC Aggregates
AFM images all of 2 x 2µm areas at 204.8 x 204.8 pixel resolution with a 500nm scale bar. The height scale for the images is set between -60 and 60nm and the peak force error is set between -200mV and 200mV. A – Height channel information of unsonicated aggregates. B – Peak force channel of unsonicated aggregates. C – Height channel of aggregates which had been sonicated for 4 mins with 5 second bursts in an ice bath. D – Peak force information of the same sample as C.

AFM analysis of the polymers formed in the assembly reaction shows a large mass of fibrils (Fig. 5.23, A and B) similar to that observed when unmodified Sup35NM and Sup35NMC aggregate (see 3.2.1.2.2). The fragmentation pattern observed (Fig. 5.23, C and D) also follows that observed with the unmodified Sup35NM fibrils (see 4.2.1.1, Fig. 4.1)

The residual monomer concentration in the assembly reaction was measured by centrifugation at 75krpm for 15 mins and loading of the supernatant onto and SDS PAGE gel. To confirm that monomeric protein wasn't pelleted under these conditions two serial dilutions were run, one of which had been centrifuged (Fig. 5.24, lanes 2-6) and the other not (Fig. 5.24, lanes 7-9). Coomassie stain (Fig. 5.24, left) showed no residual monomer in the centrifuged fibril sample (Fig. 5.24, lane 10) but fluorescent scan shows no monomeric Sup35NM LCys N_3 BCN FITC remained in solution but the formation of lower molecular weight fluorescent material not present in the starting material. This suggests that after fibril assembly the unprotected region of the monomer is susceptible to hydrolysis, liberating smaller fluorescent materials.
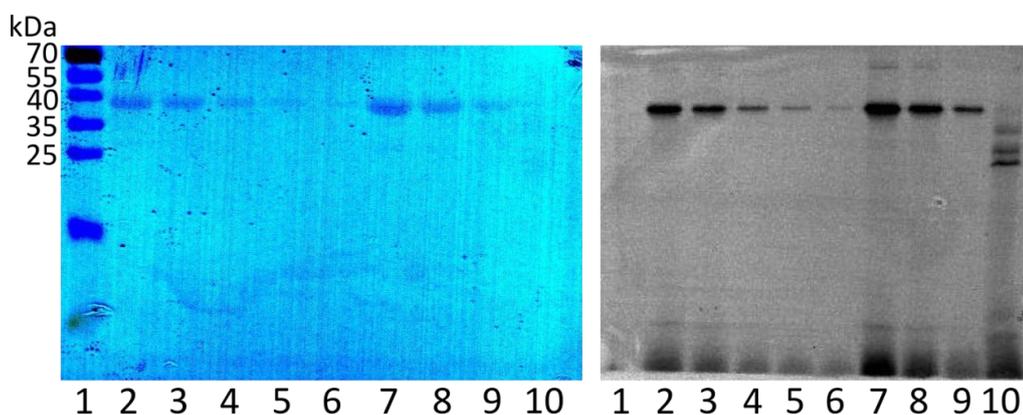
Figure 5.24   Sup35NM LCys N_3 Assembly Residual Monomer Assay
Fibril solution (10µM) were centrifuged at 75krpm and then load onto an SDS PAGE, lane 10. A centrifuged monomer serial dilution was also loaded: 7µM (lane 2), 3.5µM (lane 3), 1.75µM (lane 4), 0.875µM (lane 5) and 0.44µM (lane 6). An uncentrifuged serial dilution was loaded: 7µM (lane 7), 3.5µM (lane 8) and 1.875µM (lane 9).

Although some hydrolysis of the protein post fibril formation was observed the bulk of fluorescence was still attached to the fibrillar material, the combined concentration of the lower bands is equal to approximately 5µM protein concentration meaning 5µM remains attached to the protein. Confirmation of the retention of fluorescence post assembly was achieved by confocal imaging of fluorescent fibrils deposited onto a mica surface attached directly to a microscope slide and shows the large fluorescent clumps with a similar size to the aggregates observed on the mica surface by the optical microscope (Fig. 5.25). Comparison of the AFM images (Fig. 5.23) of the large aggregates observed in the AFM images showed similar characteristics around the edges of the clumps and possible fibrillar architecture.
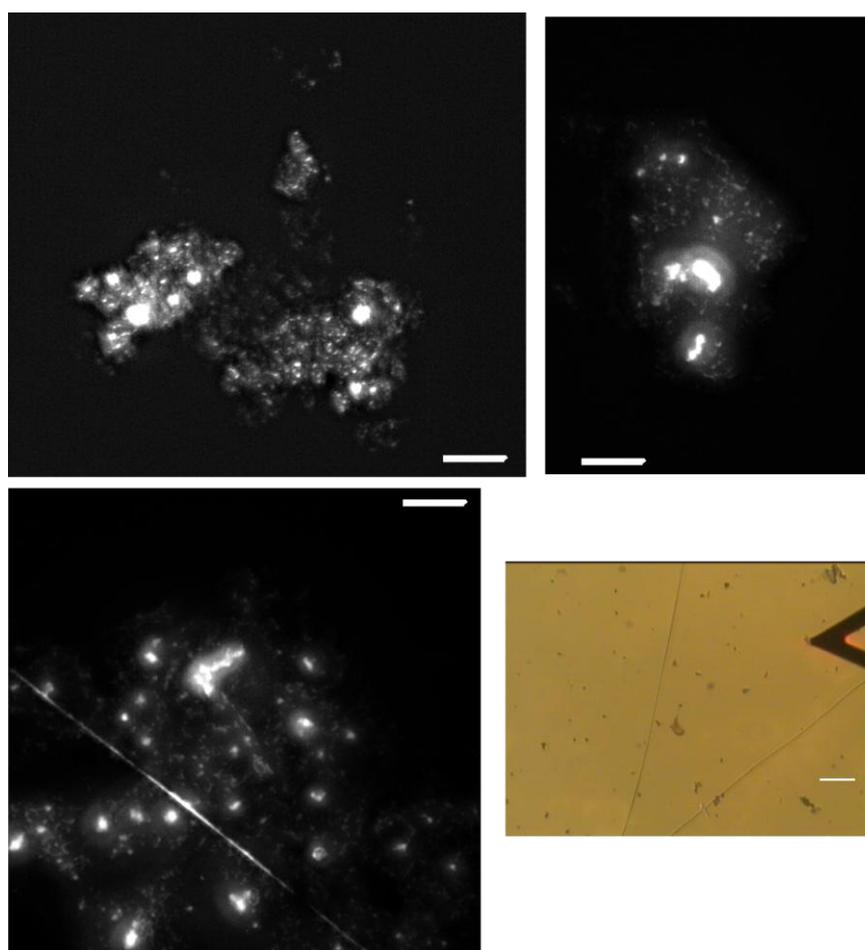
Figure 5.25   Confocal Microscopy of Sup35NM LCys N$_3$ BCN FITC Fibrils
3 confocal images with 10µm scale bars showing the prescence of large fluorescent clumps. In the bottom right a representative image of Sup35NM fibril clumps on the mica surface using the optical microscope, scale bar 1mm.

5.4.2.5 ^{19}F Labelling of Sup35NM LCys N$_3$

The successful fluorescent labelling strategy and fibril incorporation allowed implementation of the method for incorporation of ^{19}F substituents. To incorporate the same electronically equivalent trifluoromethyl groups, used in CuAAC labelling, into a reagent applicable for SPAAC chemistry 3,5-Bis(trifluoromethyl)phenyl isothiocyanate was reacted with BCN amine to furnish an appropriately modified reaction partner, BCN (CF$_3$)$_2$ (Fig. 5.26, Compound B).
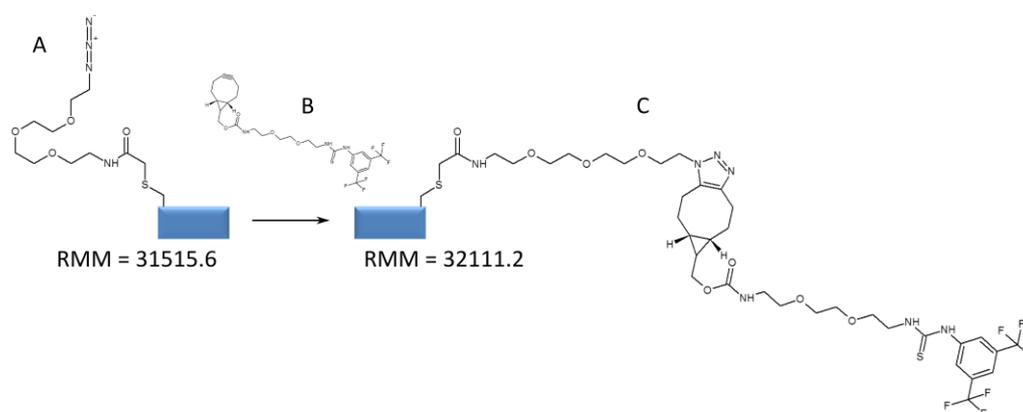
Figure 5.26   Reaction Scheme Showing the Modification of Sup35NM LCys N$_3$ with BCN (CF$_3$)$_2$

Scheme showing the conversion of Sup35NM LCys N$_3$ (blue box) (A) with a BCN reagent containing two electronically equivalent CF3 groups (B) to give a ^{19}F NMR active protein (C).
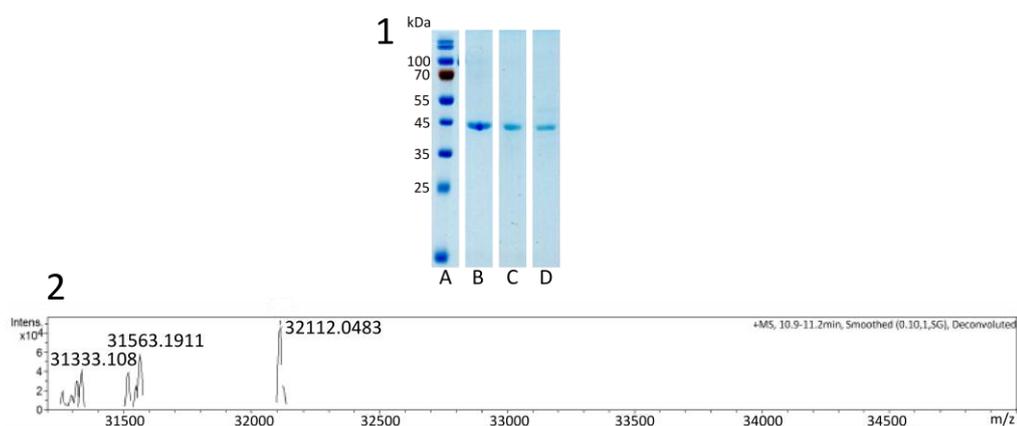
Figure 5.27   SDS PAGE Analysis of the Conversion of Sup35NM LCys to Sup35NM LCys N$_3$ BCN (CF$_3$)$_2$

1 - Coomassie stain of SDS PAGE showing the conversion of Sup35NM LCys (B) to Sup35NM LCys N$_3$ (C) and then finally modified with fluorinated reagents to give Sup35NM LCys N$_3$ BCN (CF$_3$)$_2$ (D) with a ladder as reference material (A). 2 – Deconvoluted ES(MS) spectra (CF$_3$)$_2$ modified Sup35NM showing the presence of the desired modification but also some residual starting material.

Using the same conditions used for BCN FITC attachment the BCN $(CF_3)_2$ was reacted with Sup35NM LCys N_3 (Fig. 5.26). Due to the relatively small changes in size associated with the modification no change in electrophoretic mobility was associated with modification (Fig. 5.27, Left) but successful attachment of the label was confirmed using ES(MS) (Fig. 5.27, 2), although there is some of the starting material remaining. Comparison of the ion intensity, as the starting material and the product are very similar, suggests that the reaction has gone approximately 50% to completion.

5.4.2.6 Assembly of Sup35NM LCys N3 BCN $(CF_3)_2$ into Amyloid Fibrils

Assembly of the ^{19}F labelled Sup35NM was implemented using the same conditions utilised for unmodified Sup35NM for FITC labelled Sup35NM with the kinetics monitored by ThT fluorescence (Fig. 5.28) and showed a similar profile to previous Sup35NM and the FITC derivative.
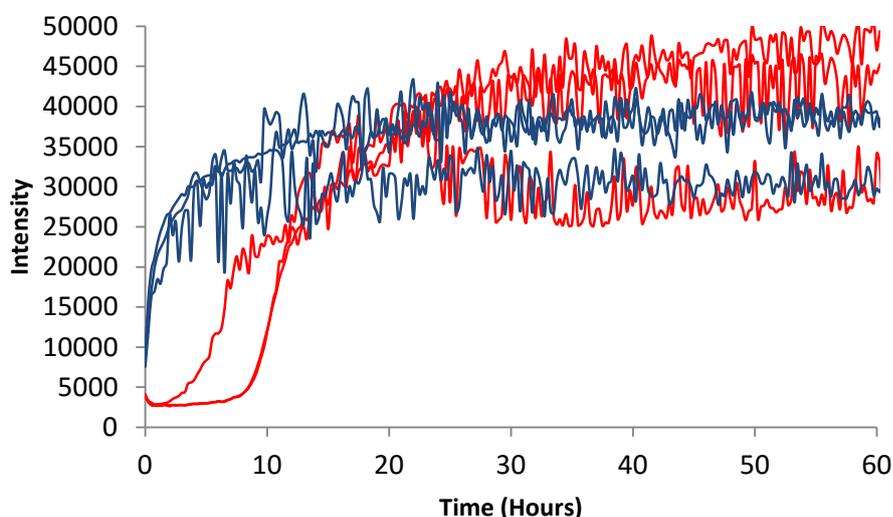
Figure 5.28   Kinetic Analysis of Sup35NM LCys N_3 BCN $(CF_3)_2$ Fibril Formation
10μM Sup35NM LCys N_3 BCN $(CF_3)_2$ was treated with ThT and the increase in fluorescence monitored. 3 replicates were run of each experiment where the yellow samples contained no protein, red samples were unseeded and blue samples were seeded with sonicated 1% Sup35NM fibrils.

AFM analysis of the material generated in the assembly reactions shows the formation of large masses of fibrils (Fig. 5.28, A and B) comparable with what is observed during the

polymerisation of unmodified Sup35NM (see 3.2.1.2.2) and Sup35NM LCys N_3 BCN FITC

(Fig. 5.22). When subjected to sonication for 4 mins on ice in 5 second bursts the large

aggregate of fibrils were broken in to small nicely dispersed particles (Fig. 5.29, C and D)

which are different to what is observed from Sup35NM and Sup35NM FITC fragmentation

and more similar to what is observed with other proteins.
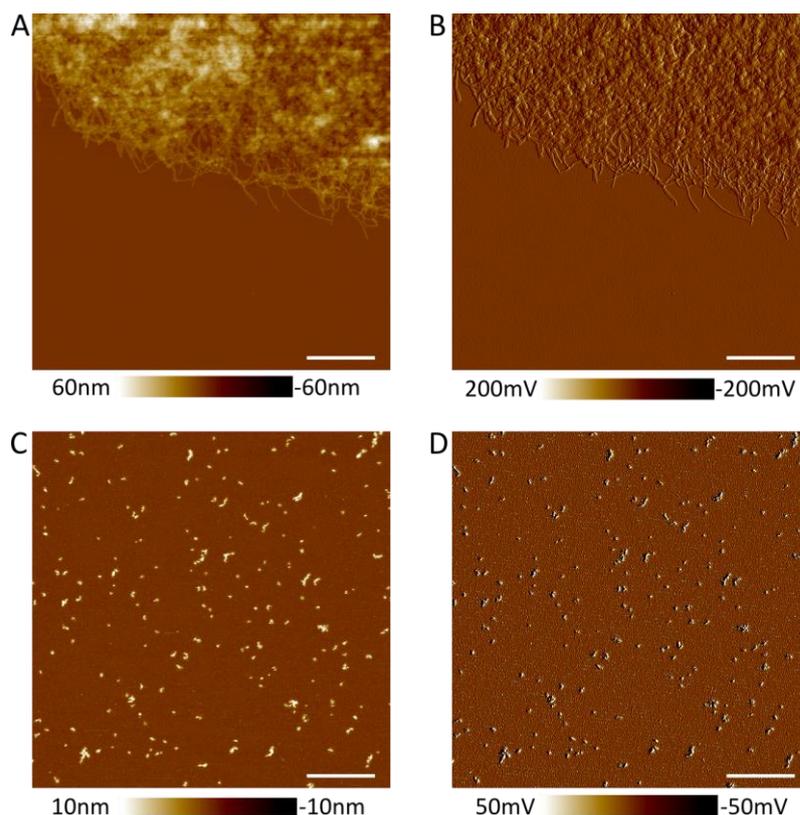
Figure 5.29   AFM Analysis of Sup35NM LCys N_3 BCN $(CF_3)_2$
AFM images all of 5 x 5µm areas at 512 x 512 pixel resolution with a 1µm scale bar. A –
Height channel information of unsonicated aggregates. The height scale is set between -60
and 60nm. B – Peak force channel of unsonicated aggregates. The peak force error scale is
set between -200 and 200mV. C – Height channel of aggregates which had been sonicated for
4 mins with 5 second bursts in an ice bath. The height scale is set between -10 and 10nm. D
– Peak force information of the same sample as C. The peak force error scale is set between -
50 and 50mV.

5.4.2.7   NMR Analysis of ^{19}F Labelled Sup35NM

Analysis of the ^{19}F NMR spectrum of the fragmented fibrillar samples gave no evidence of

any signal other than the reference materials. ^{1}H NMR analysis though did show the

presence of a large number of signals which due to the conditions should be associated

with fibrils, although the potential hydrolysis observed in Fig 5.24 could also explain this.

Subjecting the ^{1}H NMR signal to a magnetic field gradient strength showed a decay in signal (Fig. 5.30, A) and by comparison of the decay of the region between 0.5 and 1.0ppm to the dioxane signal decay gives a diffusion coefficient for the particle (Fig. 5.30, B) of $5.94 \times 10^{-10} m^2 s^{-1}$ (Table 5.2). Calculation of $D_{Protein}$ by absolute method (Fig. 5.30, C) gave a value of $1.053 \times 10^{-10} m^2 s^{-1}$ and $D_{Dioxane}$ as $3.171 \times 10^{-10} m^2 s^{-1}$.
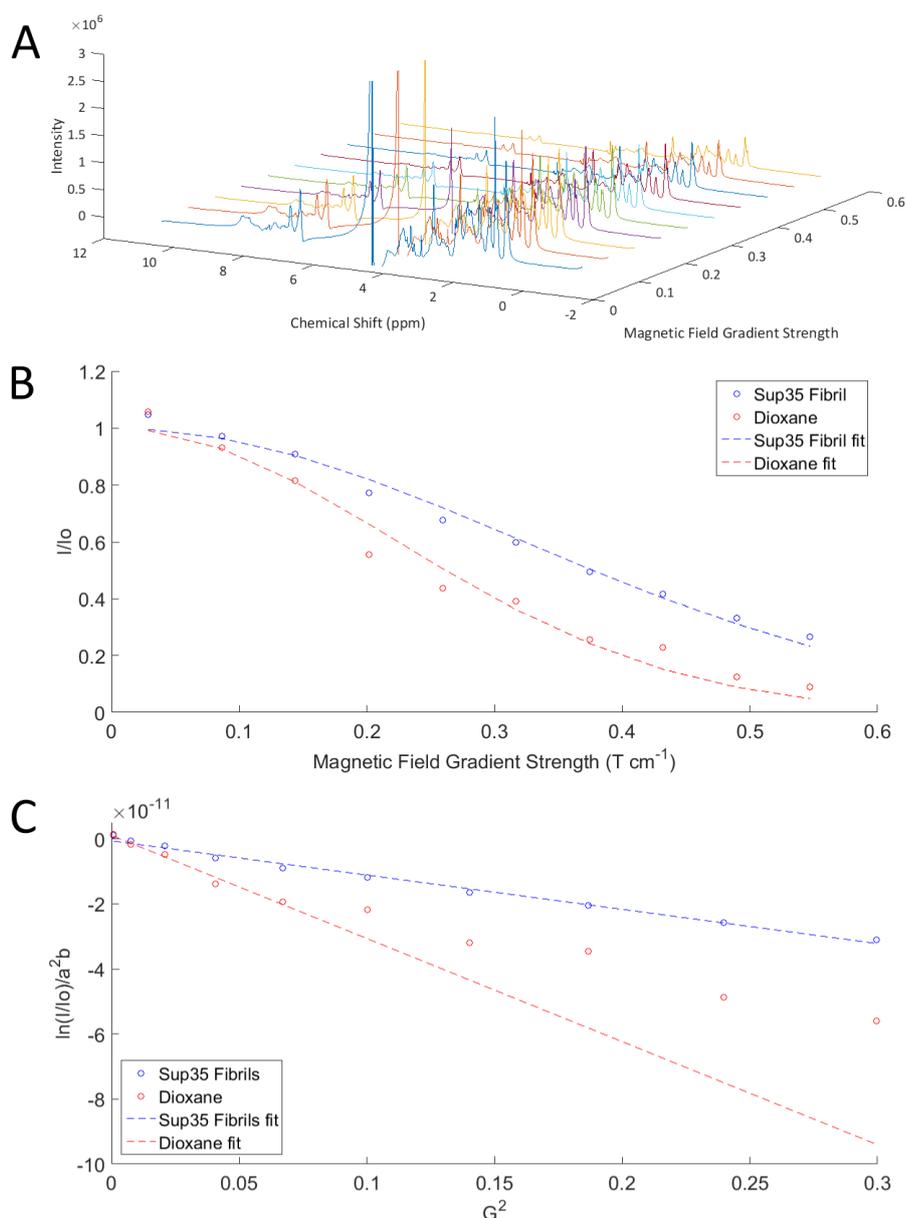
Figure 5.30 ^{1}**H DOSY NMR Analysis of Sup35NM LCys N$_3$ BCN (CF$_3$)$_2$ Fibrils**
A – Decay of ^{1}H proton NMR spectra of Sup35NM LCys N$_3$ BCN (CF$_3$)$_2$ fibrils (10μM) as magnetic field strength increases. B – Plot of the decay of of the region between 0.7 and 1ppm using a plot of I/I$_o$ against G. C – As B but plotting $\ln(I/I_o)/\alpha^2\beta$ vs G^2.

As for the lysozyme model the comparative method for Sup35 is likely to be underestimated due to the lack of baseline resolution for the internal standards. No monomer comparison was available due to the ability of Sup35 to polymerise under ambient conditions.

Table 5.2 **Calculation of the Diffusion Coefficient Associated with the Visible ^{1}H Species**

1H							
					Int. Ref.	$\ln(I/I_o)/\alpha 2\beta$	
Sample	I_o	C	DELTA	delta	$D_{Particle}$	$D_{Particle}$	Data Points
Dioxane	1.06E+07	10.12	50ms	3.5ms	1.16E-09	3.17E-10	5
Protein (Fibril)	2.81E+08	4.875	50ms	3.5ms	5.94E-10	1.05E-10	

5.5 Discussion and Conclusion

The modification of amyloidogenic proteins using bioorthogonal reactions enables the addition of a large range of functionality to materials and in the cases investigated here does not interfere with the assembly of monomeric proteins into their fibrillar state.

The understanding of which protein region is susceptible to cross-beta sheet formation is imperative for a site selective modification strategy where wrong placement could potentially inhibit polymerisation. In the case of Sup35NM the C domain and the C terminus of the M domain had previously been shown to be remote from the cross-beta sheet and therefore were an ideal target for modification and formed fibrils similarly to the unmodified monomer. However, when information regarding the active fibril forming region is unknown the promiscuous approach still offers potential as a method due to the randomness of labelling minimising any potential blocking of monomer interactions as observed by the successful modification of lysozyme fibrils.

The success of the workflow post fibril assembly is dependent on the choice of model system and the potential use of the fibrillar material. Fluorescence labelling of fibrils has utility *in vivo* and is a very good method for confirmation of reaction with azide modified

protein monomer but it also gives a degree of confidence in fibril modification that could not be achieved using other methods. The generation of fluorescent fibrils was observed for both lysozyme and Sup35NM where the attachment of the fluorophore and continued attachment post assembly gave rise to fluorescent fibrils (Fig. 5.7 and 5.24). Due to the hydrolytic mechanism associated with conversion of lysozyme monomer to fibrils under high temperature and acidic conditions (Lara et al., 2014), the degree of label lost from the fibril core during assembly could not be controlled or determined. Comparison of the labelled monomer to the fibrillar form pre and post purification by SDD AGE show the liberation of fluorescent material (Fig. 5.7 D). After extended heating the fibrils derived from lysozyme contain the ILQINS amino acid sequence (Fig. 5.31) but conversion to fibrils is not delayed until this peptide is eventually formed. As hydrolysis occurs the lysozyme structure is loosened such that amyloidogenic regions are exposed and allowed to interact. The progression towards ILQINS fibrils is therefore a gradual process where peptides are liberated from the unprotected non cross-beta sheet core, as shown by the different size peptides contained in the fibril core (Lara et al., 2014). The result of this is that some modification remains attached to the fibrils even when a hydrolytic mechanism is used but the confidence in labelling and the concentration of said label is impossible to measure.

```
KVFGRCELAA AMKRHGLDNY RGYSLGNWVC AAKFESNFNT QATNRNTDGS
TDYGILQINS RWWCNDGRTP GSRNLCNIPC SALLSSDITA SVNCAKKIVS
DGNGMNAWVA WRNRCKGTDV QAWIRGCRL
```

Figure 5.31   The Amino Acid Sequence of Lysozyme with Reactive Side Chains and Amyloid Sequence labelled.

The amino acid sequence of Lysozyme with the side-chain amines (lysine) marked in red and the side-chain thiols (cysteine) marked in green. The eight cysteine residues all form part of the secondary structure of the protein forming four disulfide bonds: C6 – C127, C30 – C115, C76 – C94, C64 – C80. The amino acid sequence shown to form the amyloid core of pH2 heat formed lysozyme fibrils (Lara et al., 2014) is highlighted in cyan.

Although the much milder conditions associated with Sup35NM aggregation should limit hydrolysis, the SDS PAGE based residual monomer assay (Fig. 5.24) showed evidence of post assembly hydrolytic process where large regions of the C domain were lost from the

fibrils but confocal imaging of the large aggregates deposited on mica show attachment of fluorophore to fibrils (Fig. 5.25). Unfortunately confocal microscopy of fluorescent lysozyme fibrils proved difficult even though SDD AGE confirmed fluorescent dye attachment to the fibrils.

Despite the success of the chemical labelling strategies the conversion of this into a robust solution based method for the assessment of particle size of amyloid fibrils by ^{19}F NMR has, so far, proved elusive. The choice of protein model system is imperative. The addition of a ^{19}F containing molecule to the C terminus of Sup35NM with a further conformationally flexible linker was expected to give the requisite flexibility for NMR visibility but no signal was observed. The lack of signal could be due to the C-terminus of the protein M domain not being completely exposed to solvent, the linker groups interacting with the protein or the hydrophobic label group burying itself into the protein structure and therefore not tumbling independently of the protein. The presence of ^{1}H signal in the Sup35NM samples suggests that some regions of the fibrils do exhibit conformational freedom, with $D_{Protein}$ = 1.053 x $10^{-10}m^2s^-1$. Replacement of the term $6\pi\eta R_h$ in the Stokes-Einstein equation (Equation 5.1) with f_T, the friction coefficient of translation, utilised by Baldwin et al (Baldwin et al., 2008) to calculate lengths from a diffusion coefficient for a fibrillar sample by analogy with a rod (Equation 5.4) it is possible to calculate an expected diffusion coefficient for a sample imaged by AFM (Fig. 5.32, A, B, C).

Equation 5.4   Calculation of Friction Coefficient For a Rod of Dimension L and r
Where η = viscosity, L = fibril length and r = fibril radius (height)

$$f_T = 3\pi\eta L \left(\frac{1}{ln(L/r) - 0.3} \right)^5$$

Incorporation of the calculated $D_{Protein}$ value into the Stejskal-Tanner equation allows a predicted decay curve to be calculated (Fig. 31, D). Surprisingly, there is close correlation between the decay observed experimentally by NMR and the decay predicted by

conversion of mean fibril length into diffusion coefficient. Comparison of this data with a prediction of the monomer derived from a predicted R_h (Fig. 31, D) shows that the decay is very similar to the fibril prediction and experimental data.
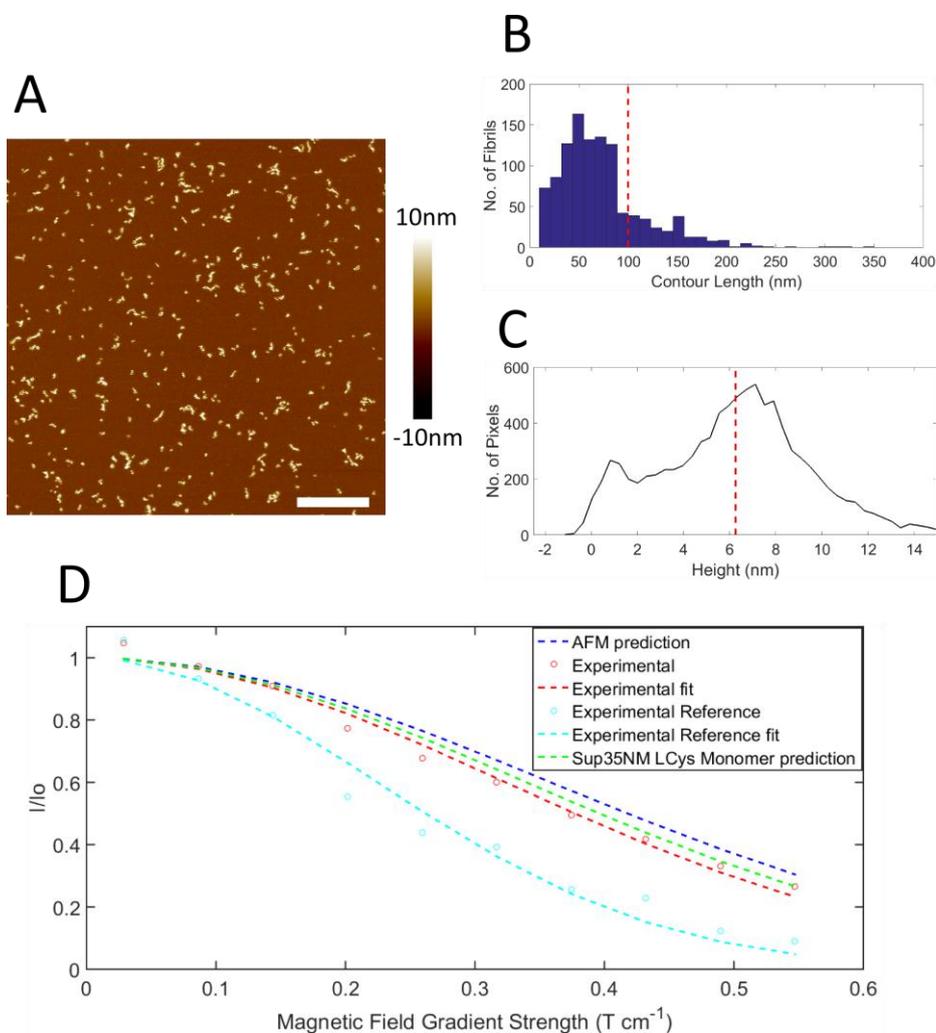
Figure 5.32   Prediction of NMR Signal Decay for Sup35NM Fibrils

A – A 5µm x 5µm image at 512pixel x 512 pixel resolution with 1µm scale bar of fragmented Sup35NM LCys N_3 BCN $(CF_3)_2$. B – Fibril length distribution calculated by MatLab analysis of the full 10µm x 10µm image observed in A. C – Fibril Height distribution calculated by MatLab analysis of the full 10µm x 10µm image observed in A. D – A plot of I/I_o against G showing the experimental decay of of labelled fibrils (red), the prediction derived from AFM particle size analysis (blue), the predicted decay of monomeric Sup35NMLCys (from the Stokes-Einstein equation and $R_h=4.75N^{0.29}$ where N = number of residues) (green) and the internal reference material (cyan).

The similarity of the experimental data to both the monomer and the predicted translational motion of the fibrils make it difficult to assess if the ^{1}H signal observed in the fibril solution is due to monomer or fibril. The residual monomer assay for the fluorescent

fibrils (Fig. 5.24) shows that the percentage conversion to fibrils is 100% which would suggests that the diffusion is related to fibrils but could also be due to the smaller protein liberated from the protein. The absence of ^{19}F signal suggests that the fluorinated signal is not tumbling in a time frame conducive to NMR analysis. The observation of ^{1}H signal is potentially explained by regions of protein having enhanced conformational motility to even the C terminus of the protein where the ^{19}F group is attached. Promiscuous labelling of the protein may well allow incorporation of ^{19}F groups onto flexible regions of the protein.

Due to the lack of confidence in the consistency of modified lysozyme fibrils the observed ^{19}F signal was placed in doubt, the value of $D_{Protein}$ extracted from the ^{19}F spectra is not dissimilar to the monomeric species (Table 1). Using the mean height (Fig. 5.33, C) and fibril length (Fig. 5.33, B) of fragmented lysozyme fibrils derived from AFM (Fig. 5.33, A) in conjunction with the Baldwin (Baldwin et al., 2008) equation for determination of f_T a diffusion coefficient of $1.905 \times 10^{-12} m^2 s^{-1}$ would be expected, significantly different to what is observed. The signal decay of the prediction shows significant difference between the observed and expected decay.

The incorporation of internal standards into the DOSY experiment allows for simpler estimation of the diffusion coefficient of the other components of the sample but the choice of reference is important. Utilisation of TFBenol as a reference sample gives ^{1}H signals with complete baseline resolution with most biomolecules and is therefore an ideal reference. TFE and dioxane suffer from the drawback of not having any ^{1}H baseline resolution due to overlap with amino acid side chain signals.

The strategy of utilising ^{19}F spectra for DOSY analysis does much simplify the analysis of the samples in terms of phasing, baseline correction and integration so is of benefit to the analysis. The fact that all parts of the protein contain protons means that it is far more

likely to observe conformationally flexible regions and suggests that a promiscuous labelling strategy may well be better for achieving flexibility of the ^{19}F label.
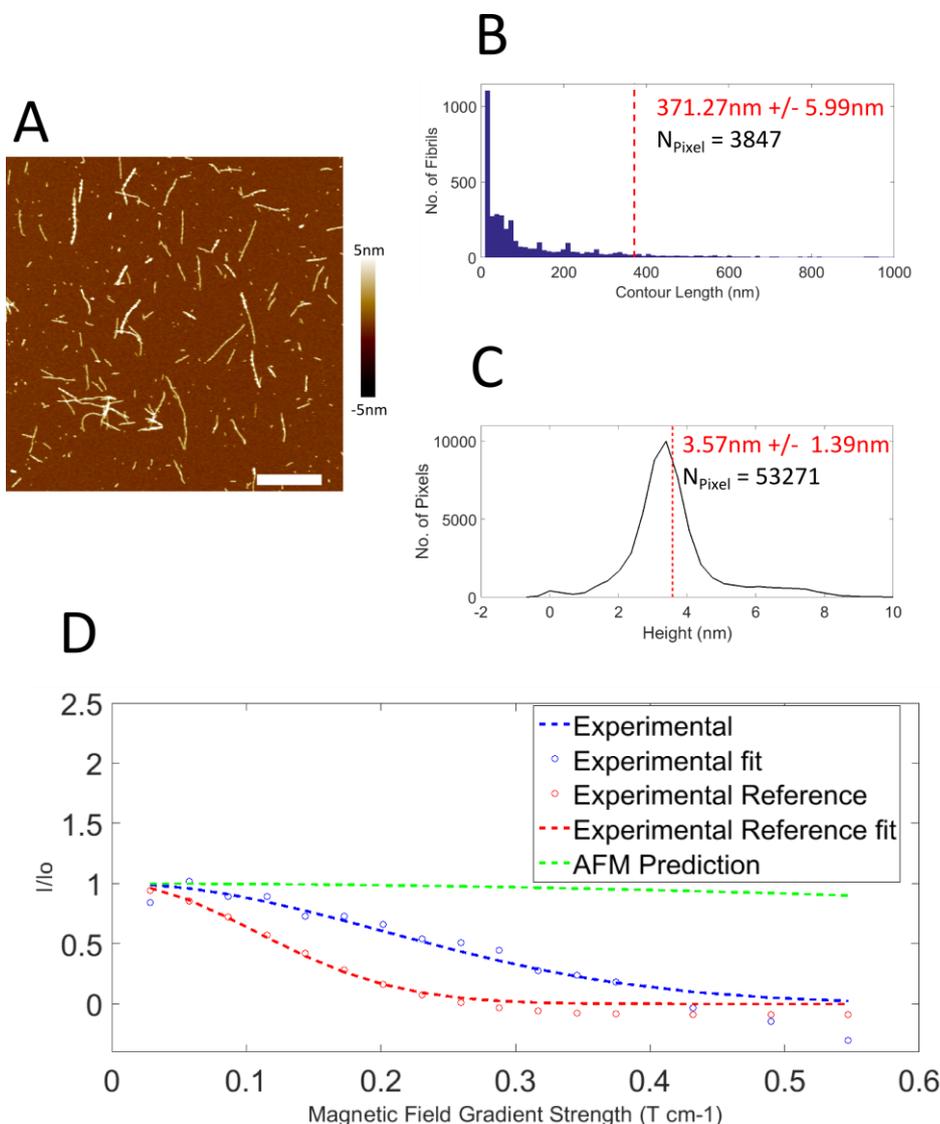
Figure 5.33   Prediction of NMR Decay for Lysozyme Fibrils

A – A 5µm x 5µm image at 512pixel x 512 pixel resolution with 1µm scale bar of fragmented lysozyme $(CF_3)_2$ fibrils. B – Fibril length distribution calculated by MatLab analysis of the full 10µm x 10µm image observed in A. C – Fibril Height distribution calculated by MatLab analysis of the full 20µm x 20µm image observed in A. D – A plot of I/I_o against G showing the experimental decay of of labelled fibrils, the prediction derived from the AFM and the internal reference material.

To conclude the chemical labelling strategy for the modification of amyloid fibrils is successful and enables the flexible functionalisation of monomers. The fluorescent analogues of these materials aide visualisation and in the case of Sup35NM have potential

as *in vivo* probes. The development of the labelling procedure to allow the determination of particle size by NMR requires further work particulary in finding a model fibril system which is suitable for analysis.

5.6   Further work

The labelling of amyloid fibrils could be aided by the utilisation of amber codon suppression technology or by the implementation of expressed protein ligation or enzymatic modifications. Further investigation of promiscuous labeeling strategies on pH7 forming fibrils is also required.

The transformation of *in vitro* generated fluorescent Sup35NM into yeast cells would allow for the identification of the localisation of these particles post transformation.

To continue the development of ^{19}F DOSY NMR as a tool for in solution size determination the promiscuous labelling of Sup35NM would potentially allow incorporation of ^{19}F label onto the conformationally flexible domain visable in the ^{1}H spectra. Alternatively the use of α-synuclein as a model system would give access to fibrils with lower propensity to aggregate whilst also maintaining the advantage of pH7 fibril assembly. Investigation of other linker units which would be more directional and not liable to collapse onto the protein structure such as a polyproline rod (Sato et al., 2007) may give more flexibility and therefore NMR signal. In order to aid the accurate determination of diffusion coefficient by comparison with internal reference samples characterisation of the R_h of TFBenol would allow the use of only one reference sample with baseline resolution in both ^{1}H and ^{19}F experiments. Alternatively use of TFBenol or other ^{19}F references with ^{1}H baseline resolution in conjunction with reference materials with defined diffusion coefficients that have distinct ^{1}H NMR signals such as tetramethylsilane (TMS) (Cabrita and Berger, 2001) or the water soluble alternative 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) (Horst and

Horwich, 2011; Wang et al., 2014) would allow more accurate diffusion coefficient calculation.

6

Discussion and Conclusions

6.1   Summary

Whilst the ability of proteins to adopt the correct protein fold is imperative to their biological function, the conversion to amyloid fibrils is not only associated with disease (Knowles et al., 2014). Amyloid formation also has positive impacts such as acting as an epigenetic switch allowing for conversion from one phenotype to another and allowing adaptation to the environment (Table 1.1) (Prusiner, 1982; Prusiner et al., 1983). Therefore to consider amyloid as a purely negative phenomenon to be cured is wrong. The remarkable physical properties such as strength and stiffness/flexibility (Fig. 1.21) associated with amyloid structures make them an ideal starting point for the development of nanomaterials for many different purposes.

Currently only some proteins are known to convert into their amyloid form (Table 1.1) but the assumption that all proteins under the appropriate conditions will convert to their amyloid form makes an understanding of what is required to induce amyloid formation important (Dobson, 1999). Typically conversion to amyloid involves the misfolding of monomeric proteins exposing regions of protein that are susceptible to dimerisation which leads to the formation of oligomers and then fibrils.

Whilst a link between aggregation and disease is observed, the identity of the toxic species associated with disease progression is currently contested. It is apparent from current knowledge of the kinetics of fibril formation that fragmentation is a key process in the formation of amyloid fibrils but the breakage of fibrils has other implications biologically. It has been observed that fibril fragmentation is a key process in the transfer of information from one cell to another in prion inheritence (Byrne et al., 2009). Recent developments have also highlighted fragmentation as a potential mechanism for eliciting toxicity in amyloid associated disease (Xue et al., 2009b).

In this study we have investigated the expression of amyloidogenic proteins and their subsequent assembly into amyloid fibrils. These materials have enabled the assessment of the different fibrils resistance to mechanical stress, as a measure of their propensity to fragment, using an AFM based imaging protocol (Xue and Radford, 2013). In parallel we have developed chemical labelling strategies to modify amyloid fibrils in a general way but with specific focus on developing an NMR based method for the determination of particle size in solution.

6.2   Production and *In Vitro* Assembly of Amyloidogenic Proteins

For proteins currently known to convert to amyloid fibrils the conditions required for their production and purification vary considerably. The purification of the material is dependent on the protein's propensity to undergo conversion to the amyloid form for example Aβ can been generated by bacterial expression or solid phase synthesis but is notoriously difficult to reproducibly generate fibrils, even with identical chemical structures (Finder and Glockshuber, 2007). The aggregation state post purification is the determining factor for the observed lack of reproducibility. This has led to the use of many techniques to disaggregate Aβ monomers prior to polymerisation (Lee et al., 2007; Teplow, 2006) but these methods still rely on an appropriate method of synthesis and purification (Chemuru et al., 2014; Finder et al., 2010).

Where the conversion from monomer to fibrillar form requires initiation by forcibly unfolding the protein, like heating in acidic medium, the isolation of the protein is simplified. As discussed in chapter 3 the conversion of lysozyme, β-lactoglobulin and insulin B chain are initiated by heating in pH2 this combined with their wide commercial availability as monomeric protein suggests that their isolation is relatively easy.

For proteins, like Aβ, which convert to their amyloid form at pH7 and physiological/room temperature, the purification of pure monomeric protein precursor is complicated by concomitant aggregation during expression and purification.   The development of conditions, observed in chapter 3, for the purification of α-synuclein, Sup35NM and Sup35NMC exemplifies the different strategies that need to be employed even for proteins with similar aggregation conditions. α-Synuclein is a well investigated protein (Cappai et al., 2005b; Smith et al., 2008; Wu et al., 2009) due to its association with Parkinson's disease but aside from this its high expression yield and ease of purification under native conditions make it an ideal substrate for the analysis of amyloid.   In contrast to α-synuclein purification Sup35NM, Sup35NM LCys and Sup35NMC form aggregates during expression which can be disaggregated by denaturing conditions.   The intrinsically disordered nature of Sup35NM means that on removal of the denaturant its ability to convert to amyloid fibrils is not impaired.   Whilst the N and M domains are intrinsically disordered the C terminal domain of Sup35NMC is structured and although it isn't explicitly involved in amyloid formation its unfolded state give rise to amorphous aggregates in a faster time frame than amyloid formation, when the denaturant is removed.

As well as highlighting the differences in purification of amyloidogenic proteins Chapter 3 also shows the different conditions required to assemble protein monomers into fibrils. The reaction schematic (Fig. 1.15) shows that the initial step in fibril formation is an unfolding event which yields molecules which are activated towards cross-beta sheet formation.   Initially the assumption was that the harsher conditions of heating the protein in pH2 was required to unfold lysozyme, β-lactoglobulin and insulin B chain, which unlike α-synuclein and Sup35 whose amyloid forming regions are disordered, are folded.   Chapter 3 highlights that simple misfolding is not the only process occurring in these acidic polymerisation and in fact amide hydrolysis is an important factor which is corroborated by

Sup35NM

Sup35NMC

10nm ▬▬▬ -10nm   10nm ▬▬▬ -10nm

α-Synuclein

Insulin B chain

10nm ▬▬▬ -10nm   10nm ▬▬▬ -10nm

Length = 1862.1nm +/- 56.02nm   Length = 5937.30nm +/- 488.04nm
Height = 6.94nm +/- 2.42nm   Height = 10.74nm +/- 21.05nm
Lp = 1855.52nm   Lp = 148nm

β-Lactoglobulin

Lysozyme

5nm ▬▬▬ -5nm   5nm ▬▬▬ -5nm

Length = 1117.8nm +/- 23.97nm   Length = 2757.4nm +/- 93.38nm
Height = 2.49nm +/- 1.18nm   Height = 3.38nm +/- 1.91nm
Lp = 5118.91nm   Lp = 872nm
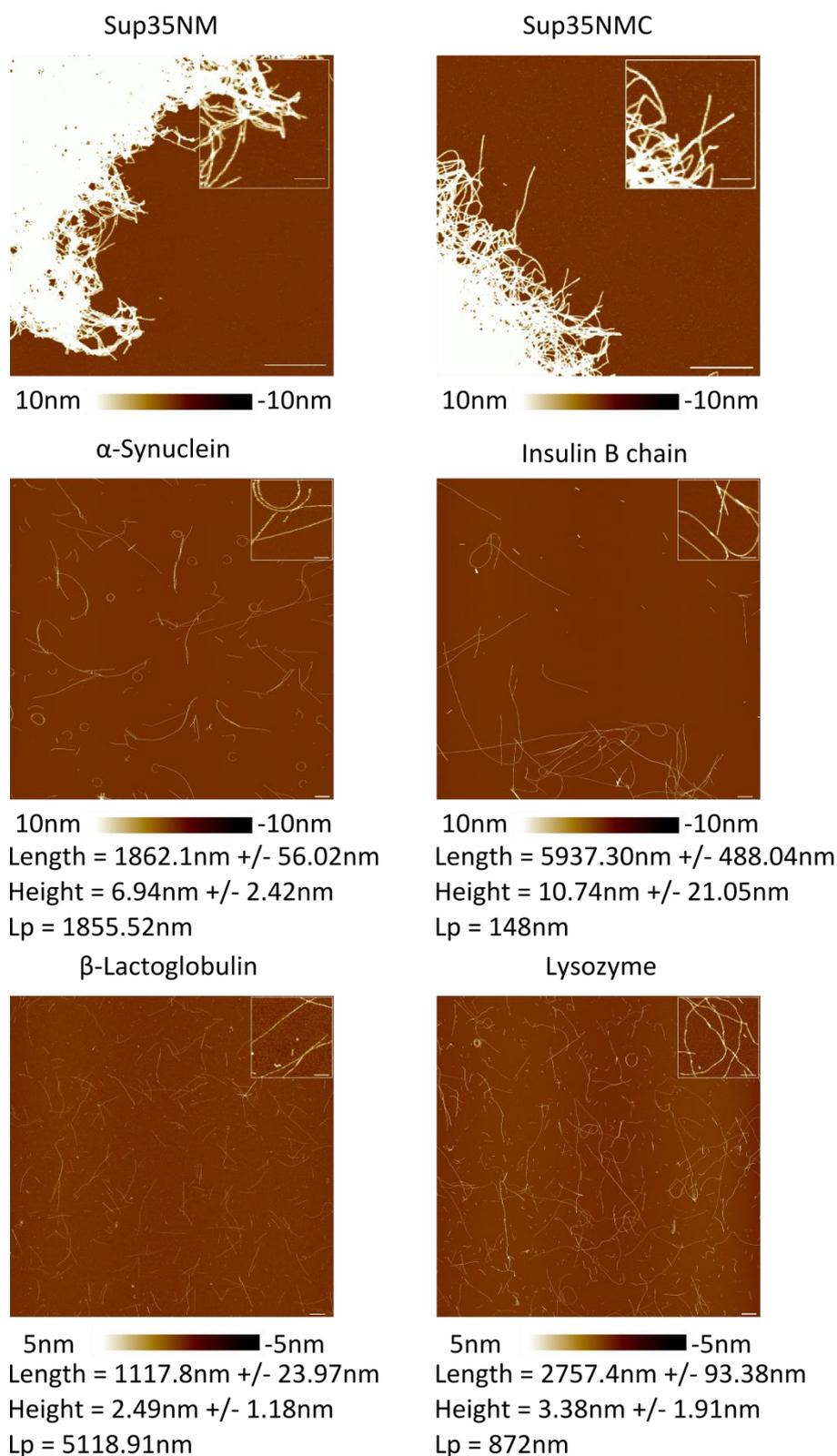
Figure 6.1   Comparison of *in vitro* formed amyloid fibrils AFM images and their properties.

AFM images of Sup35NMC, Sup35NM, α-synuclein, insulin B chain, β-lactoglobulin and lysosyme fibrils deposited onto mica and imaged at a resolution of 1pixel = 9.5656nm. The Sup35NMC, Sup35NM, α-synuclein, insulin B chain height scales are set between -10 and 10nm. β-Lactoglobulin and lysozyme have a height scale of between -5 and 5nm. Also quoted are mean contour length, height and persistence length values.

recent literature (Akkermans et al., 2008; Frare et al., 2004, 2006; Lara et al., 2011a, 2013, 2014; Piejko et al., 2015).

The heated acidic conditions employed for these proteins make amide hydrolysis unsurprising but hydrolysis has also been observed in proteins under neutral pH α-synuclein for example has exhibited autoproteolytic activity which liberates more amyloidogenic peptides under ambient conditions. This raises the question of the role that hydrolysis plays in the formation of amyloid fibrils generally and although no evidence of this process was observed for α-synuclein in this work the conversion of fluorescently labelled Sup35NM monomer to fibrils, in chapter 5, does show evidence of hydrolysis. Comparison by SDS PAGE of fluorescent monomer and fibrils showed the presence of a large amount of a slightly smaller material which suggests that it has been excised from the cross-beta sheet core. Whilst there is no evidence as to whether the hydrolysis occurred pre or post assembly it does highlight that the substrate protein monomer and the monomer unit of the cross-beta sheet core do not necessarily share the same amino acid sequence. The hydrolytic events associated with these proteins suggests further investigation on the 'misfolding' associated with the lag phase of fibril assembly.

The fibrils derived from the *in vitro* assembly reactions show the varied structures and properties associated with fibrils despite the same core structure (Fig. 6.1), despite in some cases the similar size of their starting monomer. Calculation of the persistence length for each species shows that there is no relationship between the rigidity of the polymer and its width a fact which seems at odds with the work of Knowles *et al* (Knowles et al., 2007b) who suggest that the core structures exhibit similar flexibility but it is the extraneous sidechain interactions which decrease the stiffness. Our findings suggest that L_p is not correlated with fibril width or initial monomer size which is in agreement with the work of vandenAkker *et al* (Schleeger et al., 2013; VandenAkker et al., 2011), who show a

correlation with the size of the cross-beta sheet forming domain (Fig. 6.1). The stiffness data described in this work is based on many more particles, tens of thousands, than previously attempted, Knowles *et al* 50 particles per protein, see Table 3.2, and allows greater confidence in comparison of these amyloid structures with other polymers.

6.3   Protein Conjugation and Analysis

The strength and tunable rigidity inherent to amyloid make it interesting as a nanomaterial but the ability to chemically modify fibrils offers the advantage of being able to combine these properties with non-biological functionality. There have been some examples of fibril modification previously using traditional labelling approaches (Scheibel et al., 2003) and by reappropriation of amino acid sidechain functionality into a role it is not commonly used for (Li et al., 2014). The combination of traditional labelling and state of the art bioconjugation techniques allows a completely flexible synthesis of fibrils with the potential to use one modified monomer to access a range of modified fibril products pre and post assembly. These modified materials would allow access to conducting polymers as well as allowing incorporation of radioactive tracers and therapeutic agents for tissue distribution.

Due to the complexities associated with fibril formation in both acidic and neutral pH media labelling strategies are different for both. Chapter 5 shows the development of labelling strategies which overcome the constraints of both conditions. The hydrolytic mechanism of acidic formation requires promiscuous labelling to give the highest probability of the label being attached to the fibrils post assembly. This method also has the benefit of being able to modify all amyloidogenic proteins and requires no prior knowledge of the fibril forming region to be employed. For pH 7 formed fibrils we have utilised a denatured protein with the addition of a cysteine residue to allow site selective modification. This method whilst more time consuming gives a more homogeneous monomer set and allows for utilisation of prior knowledge regarding the organisation of

the fibril to be utilised. Due to the nature, size, of the particles being generated one of the

few methods suitable for assessment of whether the modification was attached to the fibril

post assembly was by fluorescent labelling. For the cases of both lysozyme and Sup35NM

fluorescent labelling allowed confirmation of modification but also adds to the evidence of

hydrolytic events occurring during assembly (6.2).
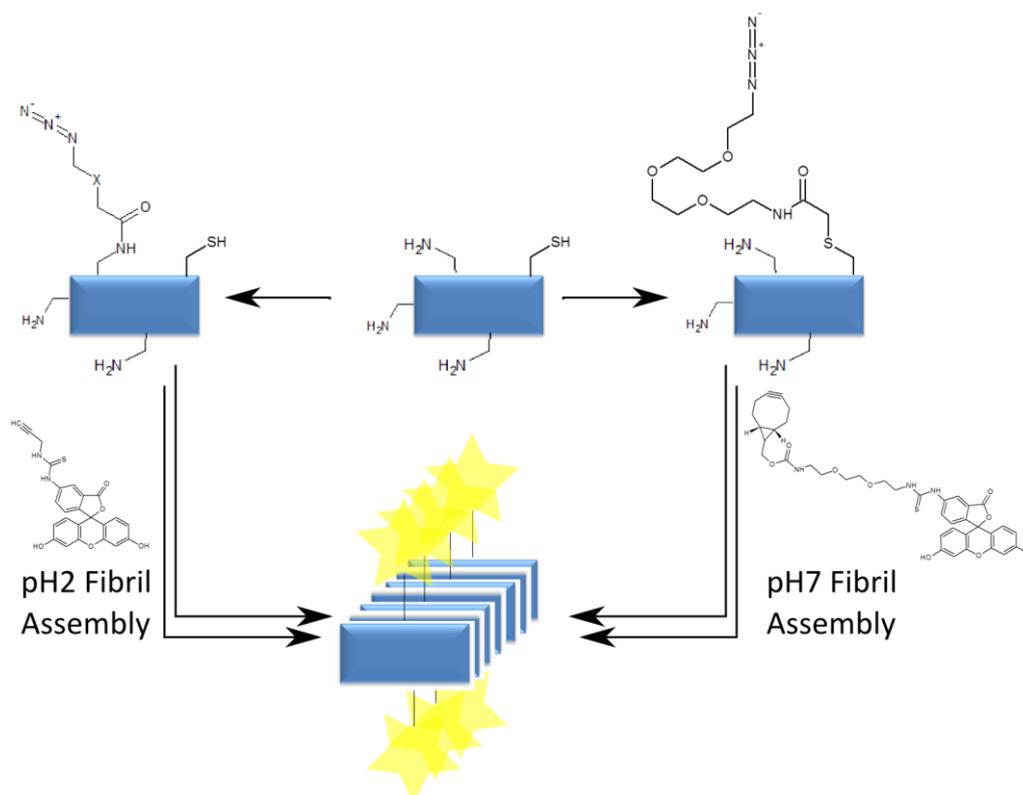
Figure 6.2 – Reaction Scheme Highlighting the Different Methods of Formation of Fluorescent Protein Monomer.
A schematic representing lysine/cysteine containing proteins being modified by the incorporation of an azide moiety through traditional labelling chemistry. To the right, reaction of the thiol functional group with an iodo acetyl reagent followed by SPAAC to give the fluorescent product which is then polymerised in pH7 conditions. Proceeding left, reaction of an amine functional group with an N-hydroxysuccinimde reagent followed by CuAAC reaction to give the fluorescent product which is polymerised under pH2 conditions.

Although the deposition of fibrils onto a surface for particle size analysis is incredibly

powerful its limitations were evident from the attempted analysis of the fragmentation of

insulin B chain in chapter 4. To circumvent these issues a solution based method for

particle size analysis was developed which uses a labelling strategy developed through fluorescent labelling to incorporate ^{19}F nuclei onto proteins which when probed gives rise to simplified NMR spectra. ^{1}H DOSY NMR of fusion proteins containing flexible domains (Baldwin et al., 2008) and genetically encoded conformationally flexible caps (Platt et al., 2009) has previously been used to estimate particle diffusion and therefore size. The complex spectra associated with ^{1}H spectra complicates these methods but the chemical labelling strategy to incorporate ^{19}F nuclei developed in chapter 5. Whilst the labelling strategies employed to date have not yielded a working particle size analysis by ^{19}F NMR DOSY the data from the ^{1}H DOSY analysis of Sup35NM suggests that this is mainly due to choice of model system and labelling strategy employed.

6.4   Fragmentation

The further development of the AFM based methodology, first described by Xue *et al* (Xue et al., 2009a), for analysing the resistance to mechanical stress of amyloid fibrils has shown that a key component of the analysis is the generation of samples suitable for AFM imaging. Insulin B-chain and Sup35 both highlight potential limitations of the method. The aggregation of Sup35 fibrils into microscopic "clumps" which are difficult to untangle preclude them from analysis whilst the potential alternative fragmentation mechanism of insulin B-chain fibrils prevents monitoring the change in amyloid fibril particle size distribution over time.

Despite the limitations observed, the method was successful for the remaining three model systems and the data captured followed the previously analysed β2m samples in following a two phase fragmentation where the fragmentation of longer fibrils occurs faster than shorter fibrils. Monitoring the change in fragmentation rate as a function of fibril length for all fibrils shows that their rates of fragmentation do not change that much until a point at around 250-500nm where the fragmentation rate constant lowers dramatically (Fig. 4.22).

Previous work has shown that the L_p of amyloid fibrils is linked with the cross-beta sheet structure which runs along the fibril backbone (Knowles et al., 2007a; VandenAkker et al., 2011). This work highlights the previously undescribed relationship of L_p with the fibrils propensity to fragment when subjected to mechanical stress. Comparison of the fragmentation rate constant for particles 3µm in length for all particles with their L_p shows the emergence of an apparent linear correlation between the stiffness of the polymer and their resistance to mechanical stress. Whilst this linearity needs to be further tested by the expansion of the data set it does raise the possibility of predicting an amyloid fibrils resistance to mechanical stress by calculation of its persistence length.

6.5   Concluding Remarks

The prevalence of amyloid in disease and its potential applications as a nanomaterial make the development of reproducible *in vitro* polymerisation of amyloid fibrils great importance and allows for their physical properties to be fully investigated and understood.

By using a range of proteins either commercially available or produced through bacterial expression we have investigated different strategies for the purification of monomers, the conditions required to initiate their conversion to amyloid fibrils, and their mechanical and structural properties.

The development of robust, reproducible methods for fibril assembly has allowed the investigation of the resistance to mechanical stress of amyloid fibrils. This work has generated quantitative data from 345403 fibrils across three amyloid forming proteins and has enabled a hitherto novel comparison of fibrils propensity to fragment and their mechanical stabilities. This comparison highlights a connection between the fragmentation rate of long particles with their stiffness. This study also suggest the fragmentation of all fibrillar materials may be expected to break in a similar length dependent manner to that described by W.F. Xue & Radford (2013) when subjected to mechanical agitation. This work

has highlighted that whilst most fibrillar materials tend to break symmetrically, there may be exceptions to this, for example insulin B-chain whose fragmentation generates large volumes of small material rather than giving a gradual reduction in size that is observed for lysozyme, α-synuclein, β-lactoglobulin and $\beta_2$microglobulin. In addition, this work has also shown the importance of hydrolysis on the conversion of proteins to amyloid fibrils, *in vitro*, even for proteins not exposed to low pH and elevated temperatures and poses questions regarding the primary amino acid sequences observed in mature fibril monomers.

The chemical conjugation strategy presented here in chapter 5 has allowed the development of a flexible chemical labelling strategy capable of modifying all amyloid fibrils with varied functionality, with particular focus on chemical biology strategies for in-solution particle size determination (^{19}F DOSY NMR). Whilst the NMR size determination has so far been unsuccessful it has provided valuable insights into the requirements for conversion of this process into a successful method, such as choice of linker and label to different amyloid models, which will be able to supplement AFM mediated size determination. The labelling strategy used also allowed development of alternative fluorescently labelled fibrils whilst offering advantages for future visualisation of *in vitro* fibrils *in vivo* has also provided evidence that there is potentially a hydrolytic component to the conversion of proteins to fibrils even under mild, ambient conditions and offers a way of monitoring this process.

Bibliography

Abels, J.A., Moreno-Herrero, F., van der Heijden, T., Dekker, C., and Dekker, N.H. (2005). Single-Molecule Measurements of the Persistence Length of Double-Stranded RNA. Biophys. J. *88*, 2737–2744.

Adamcik, J., and Mezzenga, R. (2012). Study of amyloid fibrils via atomic force microscopy. Curr. Opin. Colloid Interface Sci. *17*, 369–376.

Adamcik, J., Jung, J.-M., Flakowski, J., De Los Rios, P., Dietler, G., and Mezzenga, R. (2010). Understanding amyloid aggregation by statistical analysis of atomic force microscopy images. Nat. Nanotechnol. *5*, 423–428.

Agard, N.J., Prescher, J. a., and Bertozzi, C.R. (2004). A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. J. Am. Chem. Soc. *126*, 15046–15047.

Agarwal, P., and Bertozzi, C.R. (2014). Site-specific antibody-drug conjugates : the nexus of bioorthogonal chemistry , protein engineering , and drug development. Bioconjug. Chem. *26*, 176–192.

Aguzzi, A., and O'Connor, T. (2010). Protein aggregation diseases: pathogenicity and therapeutic perspectives. Nat. Rev. Drug Discov. *9*, 237–248.

Akkermans, C., Venema, P., van der Goot, A.J., Gruppen, H., Bakx, E.J., Boom, R.M., and van der Linden, E. (2008). Peptides are building blocks of heat-induced fibrillar protein aggregates of beta-lactoglobulin formed at pH 2. Biomacromolecules *9*, 1474–1479.

Al-Hilaly, Y.K., Williams, T.L., Stewart-Parker, M., Ford, L., Skaria, E., Cole, M., Bucher, W.G., Morris, K.L., Sada, A.A., Thorpe, J.R., et al. (2013). A central role for dityrosine crosslinking of Amyloid-β in Alzheimer's disease. Acta Neuropathol. Commun. *1*, 83.

Anderson, J.C., Wu, N., Santoro, S.W., Lakshman, V., King, D.S., and Schultz, P.G. (2004). An expanded genetic code with a functional quadruplet codon. Proc. Natl. Acad. Sci. U. S. A. *101*, 7566–7571.

Askarieh, G., Hedhammar, M., Nordling, K., Saenz, A., Casals, C., Rising, A., Johansson, J., and Knight, S.D. (2010). Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. Nature *465*, 236–238.

Atwood, C.S., Perry, G., Zeng, H., Kato, Y., Jones, W.D., Ling, K., Huang, X., Moir, R.D., Wang, D., Sayre, L.M., et al. (2004). Copper Mediates Dityrosine Cross-Linking of Alzheimer's Amyloid-B. Biochemistry *43*, 560–568.

Bagni, C., and Tassone, F. (2012). Fragile X syndrome: causes, diagnosis, mechanisms, and therapeutics. J. Clin. Invest. *122*, 4314–4322.

Bai, S., Wang, W., and Dybowski, C. (2010). Solid State NMR Spectroscopy. Anal. Chem. *82*, 4917–4924.

Baldwin, A.J., Anthony-Cahill, S.J., Knowles, T.P.J., Lippens, G., Christodoulou, J., Barker, P.D., and Dobson, C.M. (2008). Measurement of amyloid fibril length distributions by inclusion of rotational motion in solution NMR diffusion measurements. Angew. Chem. Int. Ed. Engl. *47*, 3385–3387.

Barth, A. (2007). Infrared spectroscopy of proteins. Biochim. Biophys. Acta *1767*, 1073–1101.

Baxa, U., Keller, P.W., Cheng, N., Wall, J.S., and Steven, A.C. (2011). In Sup35p filaments (the [PSI+] prion), the globular C-terminal domains are widely offset from the amyloid fibril backbone. Mol. Microbiol. *79*, 523–532.

Bergström, J., Gustavsson, Å., Hellman, U., Sletten, K., Murphy, C.L., Weiss, D.T., Solomon, A., Olofsson, B.O., and Westermark, P. (2005). Amyloid deposits in transthyretin-derived amyloidosis: Cleaved transthyretin is associated with distinct amyloid morphology. J. Pathol. *206*, 224–232.

Bhak, G., Lee, S., Park, J.W., Cho, S., and Paik, S.R. (2010). Amyloid hydrogel derived from curly protein fibrils of alpha-synuclein. Biomaterials *31*, 5986–5995.

Biancalana, M., and Koide, S. (2010). Molecular mechanism of Thioflavin-T binding to amyloid fibrils. Biochim. Biophys. Acta *1804*, 1405–1412.

Blake, C.C., Koenig, D.F., Mair, G. a, North, a C., Phillips, D.C., and Sarma, V.R. (1965). Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. Nature *206*, 757–761.

Bolisetty, S., Adamcik, J., and Mezzenga, R. (2011). Snapshots of fibrillation and aggregation kinetics in multistranded amyloid β-lactoglobulin fibrils. Soft Matter *7*, 493–499.

Bonar, L., Cohen, A.S., and Skinner, M.M. (1969). Characterization of the Amyloid Fibril as a Cross-β Protein. Proc. Soc. Exp. Biol. Med. *131*, 1373–1375.

Booth, D.R., Sunde, M., Bellotti, V., Robinson, C. V., Hutchinson, W.L., Fraser, P.E., Hawkins, P.N., Dobson, C.M., Radford, S.E., Blake, C.C.F., et al. (1997). Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. Nature *385*, 787–793.

Brange, J., Langkjaer, L., Havelund, S., and Volund, A. (1992). Chemical stability of insulin. 1. Hydrolytic degradation during storage of pharmaceutical preparations. Pharm. Res. *9*, 715–726.

Bunnage, M.E., Chekler, E.L.P., and Jones, L.H. (2013). Target validation using chemical probes. Nat. Chem. Biol. *9*, 195–199.

Bunnage, M.E., Gilbert, A.M., Jones, L.H., and Hett, E.C. (2015). Know your target, know your molecule. Nat. Chem. Biol. *11*, 368–372.

Byrne, L.J., Cole, D.J., Cox, B.S., Ridout, M.S., Morgan, B.J.T., and Tuite, M.F. (2009). The number and transmission of [PSI] prion seeds (Propagons) in the yeast Saccharomyces cerevisiae. PLoS One *4*, e4670.

Cabrita, E.J., and Berger, S. (2001). DOSY studies of hydrogen bond association: tetramethylsilane as a reference compound for diffusion studies. Magn. Reson. Chem. *39*, S142–S148.

Cappai, R., Leck, S.-L., Tew, D.J., Williamson, N. a, Smith, D.P., Galatis, D., Sharples, R. a, Curtain, C.C., Ali, F.E., Cherny, R. a, et al. (2005a). Dopamine promotes alpha-synuclein aggregation into SDS-resistant soluble oligomers via a distinct folding pathway. FASEB J. *19*, 1377–1379.

Chalker, J.M., Bernardes, G.J.L., Lin, Y. a., and Davis, B.G. (2009). Chemical modification of proteins at cysteine: Opportunities in chemistry and biology. Chem. - An Asian J. *4*, 630–640.

Chemuru, S., Kodali, R., and Wetzel, R. (2014). Improved chemical synthesis of hydrophobic Aβ peptides using addition of C-terminal lysines later removed by carboxypeptidase B. Biopolymers *102*, 206–221.

Chin, J.W., Cropp, T.A., Anderson, J.C., Mukherji, M., Zhang, Z., and Schultz, P.G. (2003). An expanded eukaryotic genetic code. Science *301*, 964–967.

Chiti, F., and Dobson, C.M. (2006). Protein misfolding, functional amyloid, and human disease. Annu. Rev. Biochem. *75*, 333–366.

Chiti, F., Webster, P., Taddei, N., Clark, a, Stefani, M., Ramponi, G., and Dobson, C.M. (1999). Designing conditions for in vitro formation of amyloid protofilaments and fibrils. Proc. Natl. Acad. Sci. U. S. A. *96*, 3590–3594.

Conway, K.A., Harper, J.D., and Lansbury, P.T. (1998). Accelerated in vitro fibril formation by a mutant α-synuclein linked to early-onset Parkinson disease. Nat. Med. *4*, 1318–1320.

Cox, B.S. (1965). Psi, A cytoplasmic suppressor of super-suppressor in yeast. Heredity (Edinb). *20*, 505–521.

Cox, B., Ness, F., and Tuite, M. (2003). Analysis of the generation and segregation of propagons: entities that propagate the [PSI+] prion in yeast. Genetics *165*, 23–33.

Crick, F.H.C., Barnett, L., Brenner, S., and Watts-Tobin, R.J. (1961). General Nature of the Genetic Code for Proteins. Nature *192*, 1227–1232.

Croguennec, T., Mollé, D., Mehra, R., and Bouhallab, S. (2004). Spectroscopic characterization of heat-induced nonnative B-Lactoglobulin monomers. Protein Sci. 1340–1346.

Dauer, W., and Przedborski, S. (2003). Parkinson's Disease : Mechanisms and Models. Neuron *39*, 889–909.

DePace, A.H., Santoso, A., Hillner, P., and Weissman, J.S. (1998). A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. Cell *93*, 1241–1252.

Devlin, G.L., Knowles, T.P.J., Squires, A., McCammon, M.G., Gras, S.L., Nilsson, M.R., Robinson, C. V., Dobson, C.M., and MacPhee, C.E. (2006). The Component Polypeptide Chains of Bovine Insulin Nucleate or Inhibit Aggregation of the Parent Protein in a Conformation-dependent Manner. J. Mol. Biol. *360*, 497–509.

Dobson, C.M. (1999). Protein misfolding, evolution and disease. TIBS *0004*, 329–332.

Dobson, C.M. (2003). Protein folding and misfolding. Nature *426*, 884–890.

Duennwald, M.L., Echeverria, A., and Shorter, J. (2012). Small heat shock proteins potentiate amyloid dissolution by protein disaggregases from yeast and humans. PLoS Biol. *10*, e1001346.

Dumas, A., Lercher, L., Spicer, C.D., and Davis, B.G. (2014). Designing logical codon reassignment – expanding the chemistry in biology. Chem. Sci.

Duyao, M., Ambrose, C., Myers, R., Novelletto, A., Persichetti, F., Frontali, M., Folstein, S., Ross, C., Franz, M., Abbott, M., et al. (1993). Trinucleotide repeat length instability and age of onset in Huntington's disease. Nat. Genet. 387–392.

Fahn, S., and Cohen, G. (2007). Principles and practice of movement disorders.

Ferreira, S.T., Vieira, M.N.N., and De Felice, F.G. (2007). Soluble protein oligomers as emerging toxins in Alzheimer's and other amyloid diseases. IUBMB Life *59*, 332–345.

Finder, V.H., and Glockshuber, R. (2007). Amyloid-β Aggregation. Neuro-Degenerative Dis. *4*, 13–27.

Finder, V.H., Vodopivec, I., Nitsch, R.M., and Glockshuber, R. (2010). The Recombinant Amyloid-B Peptide AB1-42 Aggregates Faster and Is More Neurotoxic than Synthetic AB1-42. J. Mol. Biol. *396*, 9–18.

Fitzpatrick, a. W.P., Debelouchina, G.T., Bayro, M.J., Clare, D.K., Caporini, M. a., Bajaj, V.S., Jaroniec, C.P., Wang, L., Ladizhansky, V., Muller, S. a., et al. (2013). Atomic structure and hierarchical assembly of a cross- amyloid fibril. Proc. Natl. Acad. Sci. *110*, 5468–5473.

Foderà, V., Librizzi, F., Groenning, M., Van De Weert, M., and Leone, M. (2008). Secondary nucleation and accessible surface in insulin amyloid fibril formation. J. Phys. Chem. B *112*, 3853–3858.

Frare, E., Polverino De Laureto, P., Zurdo, J., Dobson, C.M., and Fontana, A. (2004). A highly amyloidogenic region of hen lysozyme. J. Mol. Biol. *340*, 1153–1165.

Frare, E., Mossuto, M.F., Polverino de Laureto, P., Dumoulin, M., Dobson, C.M., and Fontana, A. (2006). Identification of the Core Structure of Lysozyme Amyloid Fibrils by Proteolysis. J. Mol. Biol. *361*, 551–561.

Frauenfelder, H., Sligar, S., and Wolynes, P. (1991). The energy landscapes and motions of proteins. Science (80-.). *254*, 1598–1603.

Frottin, F., Martinez, A., Peynot, P., Mitra, S., Holz, R.C., Giglione, C., and Meinnel, T. (2006). The proteomics of N-terminal methionine cleavage. Mol. Cell. Proteomics *5*, 2336–2349.

Genentech (1978). First successful lab production of human insulin announced.

Gittes, F., Mickey, B., Nettleton, J., and Howard, J. (1993). Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape. J. Cell Biol. *120*, 923–934.

Glover, J.R., Kowal, a S., Schirmer, E.C., Patino, M.M., Liu, J.J., and Lindquist, S. (1997). Self-seeded fibers formed by Sup35, the protein determinant of [PSI+], a heritable prion-like factor of S. cerevisiae. Cell *89*, 811–819.

Goodman, M. (2009). Market watch: Sales of biologics to show robust growth through to 2013. Nat. Rev. Drug Discov. *8*, 837.

Gosal, W.S., Clark, A.H., and Ross-Murphy, S.B. (2004). Fibrillar beta-lactoglobulin gels: Part 1. Fibril formation and structure. Biomacromolecules *5*, 2408–2419.

Greenfield, N.J. (2006). Using circular dichroism spectra to estimate protein secondary structure. Nat. Protoc. *1*, 2876–2890.

Gröst, C., and Berg, T. (2015). PYRROC: the first functionalized cycloalkyne that facilitates isomer-free generation of organic molecules by SPAAC. Org. Biomol. Chem. *13*, 3866–3870.

Hamada, D., and Dobson, C.M. (2002). A kinetic study of beta -lactoglobulin amyloid fibril formation promoted by urea. Protein Sci. *11*, 2417–2426.

Hansma, H.G., Revenko, I., Kim, K., and Laney, D.E. (1996). Atomic force microscopy of long and short double-stranded, single-stranded and triple-stranded nucleic acids. Nucleic Acids Res. *24*, 713–720.

Hartl, F., and Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. Science (80-.). *295*, 1852–1859.

Hartl, F.U., and Hayer-Hartl, M. (2009). Converging concepts of protein folding in vitro and in vivo. Nat. Struct. Mol. Biol. *16*, 574–581.

Hatters, D.M., MacRaild, C. a, Daniels, R., Gosal, W.S., Thomson, N.H., Jones, J. a, Davis, J.J., MacPhee, C.E., Dobson, C.M., and Howlett, G.J. (2003). The circularization of amyloid fibrils formed by apolipoprotein C-II. Biophys. J. *85*, 3979–3990.

Heegaard, N.H.H., Jørgensen, T.J.D., Rozlosnik, N., Corlin, D.B., Pedersen, J.S., Tempesta, A.G., Roepstorff, P., Bauer, R., and Nissen, M.H. (2005). Unfolding, aggregation, and seeded amyloid formation of lysine-58-cleaved beta 2-microglobulin. Biochemistry *44*, 4397–4407.

Heise, H. (2008). Solid-state NMR spectroscopy of amyloid proteins. ChemBioChem *9*, 179–189.

Hill, T. (1983). Length dependence of rate constants for end-to-end association and dissociation of equilibrium linear aggregates. Biophys. J. *44*, 285–288.

Hong, D., and Fink, A.L. (2005). Independent Heterologous Fibrillation of Insulin and Its B-Chain Peptide. Biochemistry *44*, 16701–16709.

Horst, R., and Horwich, A.L. (2011). Translational Diffusion of Macromolecular Assemblies Measured. J. Am. Chem. Soc. *133*, 16354–16357.

Huisgen, R. (1963). Kinetics and Mechanism of 1,3-Dipolar Cycloadditions. Angew. Chemie Int. Ed. English *2*, 633–645.

Iijima, S. (1991). Helical microtubules of graphitic carbon. Nature *354*, 56–58.

Ikenoue, T., Lee, Y.-H., Kardos, J., Saiki, M., Yagi, H., Kawata, Y., and Goto, Y. (2014). Cold Denaturation of Alpha-Synuclein Amyloid Fibrils. Angew. Chem. Int. Ed. Engl. *53*, 7799–7804.

Itakura, K., Hirose, T., Crea, R., Riggs, A.D., Heyneker, H.L., Bolivar, F., and Boyer, H.W. (1977). Expression in Escherichia coli of a chemically synthesized gene for the hormone somatostatin. Science *198*, 1056–1063.

Jackson, D.A., Symons, R.H., and Berg, P. (1972). Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. *69*, 2904–2909.

Jagadish, K., Borra, R., Lacey, V., Majumder, S., Shekhtman, A., Wang, L., and Camarero, J. a (2013). Expression of fluorescent cyclotides using protein trans-splicing for easy monitoring of cyclotide-protein interactions. Angew. Chem. Int. Ed. Engl. *52*, 3126–3131.

Jakhria, T., Hellewell, A.L., Porter, M.Y., Jackson, M.P., Tipping, K.W., Xue, W.-F., Radford, S.E., and Hewitt, E.W. (2014). β2-Microglobulin Amyloid Fibrils Are Nanoparticles That Disrupt Lysosomal Membrane Protein Trafficking and Inhibit Protein Degradation by Lysosomes. J. Biol. Chem. *289*, 35781–35794.

Jaskolski, M., Dauter, Z., and Wlodawer, A. (2014). A brief history of macromolecular crystallography, illustrated by a family tree and its nobel fruits. FEBS J. *281*, 3985–4009.

Johnson, L.N. (1998). The early history of lysozyme. Nat. Struct. Biol. *5*, 942–944.

Kabani, M., Cosnier, B., Bousset, L., Rousset, J.-P., Melki, R., and Fabret, C. (2011). A mutation within the C-terminal domain of Sup35p that affects [PSI+] prion propagation. Mol. Microbiol. *81*, 640–658.

Kashchiev, D., Cabriolu, R., and Auer, S. (2013). Confounding the paradigm: peculiarities of amyloid fibril nucleation. J. Am. Chem. Soc. *135*, 1531–1539.

Kayed, R., and Lasagna-Reeves, C.A. (2013). Molecular mechanisms of amyloid oligomers toxicity. J. Alzheimer's Dis. *33*, S67–S78.

Kayed, R., Head, E., Thompson, J.L., McIntire, T.M., Milton, S.C., Cotman, C.W., and Glabe, C.G. (2003). Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. Science *300*, 486–489.

Kellenberger, E. (2004). The evolution of molecular biology. EMBO Rep. *5*, 546–549.

Kheterpal, I., Williams, A., Murphy, C., Bledsoe, B., and Wetzel, R. (2001). Structural Features of the Aβ Amyloid Fibril Elucidated by Limited Proteolysis. Biochemistry *40*, 11757–11767.

Khurana, R., Coleman, C., Ionescu-Zanetti, C., Carter, S. a, Krishna, V., Grover, R.K., Roy, R., and Singh, S. (2005). Mechanism of thioflavin T binding to amyloid fibrils. J. Struct. Biol. *151*, 229–238.

Kim, M.-S., Pinto, S.M., Getnet, D., Nirujogi, R.S., Manda, S.S., Chaerkady, R., Madugundu, A.K., Kelkar, D.S., Isserlin, R., Jain, S., et al. (2014). A draft map of the human proteome. Nature *509*, 575–581.

King, C.-Y., and Diaz-Avalos, R. (2004). Protein-only transmission of three yeast prion strains. Nature *428*, 319–323.

Knowles, T.P.J., and Buehler, M.J. (2011). Nanomechanics of functional and pathological amyloid materials. Nat. Nanotechnol. *6*, 469–479.

Knowles, T.P., Fitzpatrick, A.W., Meehan, S., Mott, H.R., Vendruscolo, M., Dobson, C.M., and Welland, M.E. (2007a). Role of intermolecular forces in defining material properties of protein nanofibrils. Science *318*, 1900–1903.

Knowles, T.P., Fitzpatrick, A.W., Meehan, S., Mott, H.R., Vendruscolo, M., Dobson, C.M., and Welland, M.E. (2007b). Role of intermolecular forces in defining material properties of protein nanofibrils. Science *318*, 1900–1903.

Knowles, T.P.J., Vendruscolo, M., and Dobson, C.M. (2014). The amyloid state and its association with protein misfolding diseases. Nat. Rev. Mol. Cell Biol. *15*, 384–396.

Kok, W.M., Cottam, J.M., Ciccotosto, G.D., Miles, L. a., Karas, J. a., Scanlon, D.B., Roberts, B.R., Parker, M.W., Cappai, R., Barnham, K.J., et al. (2013). Synthetic dityrosine-linked β-amyloid dimers form stable, soluble, neurotoxic oligomers. Chem. Sci. *4*, 4449.

Kolb, H.C., Finn, M.G., and Sharpless, K.B. (2001). Click Chemistry: Diverse Chemical Function from a Few Good Reactions. Angew. Chemie Int. Ed. *40*, 2004–2021.

Krebs, M.R., Wilkins, D.K., Chung, E.W., Pitkeathly, M.C., Chamberlain, a K., Zurdo, J., Robinson, C. V, and Dobson, C.M. (2000). Formation and seeding of amyloid fibrils from wild-type hen lysozyme and a peptide fragment from the beta-domain. J. Mol. Biol. *300*, 541–549.

Krebs, M.R.H., Bromley, E.H.C., and Donald, a M. (2005). The binding of thioflavin-T to amyloid fibrils: localisation and implications. J. Struct. Biol. *149*, 30–37.

Krebs, M.R.H., Devlin, G.L., and Donald, A.M. (2009). Amyloid fibril-like structure underlies the aggregate structure across the pH range for beta-lactoglobulin. Biophys. J. *96*, 5013–5019.

Krzewska, J., and Melki, R. (2006). Molecular chaperones and the assembly of the prion Sup35p, an in vitro study. EMBO J. *25*, 822–833.

Krzewska, J., Tanaka, M., Burston, S.G., and Melki, R. (2007). Biochemical and functional analysis of the assembly of full-length Sup35p and its prion-forming domain. J. Biol. Chem. *282*, 1679–1686.

Lara, C., Adamcik, J., Jordens, S., and Mezzenga, R. (2011a). General self-assembly mechanism converting hydrolyzed globular proteins into giant multistranded amyloid ribbons. Biomacromolecules *12*, 1868–1875.

Lara, C., Usov, I., Adamcik, J., and Mezzenga, R. (2011b). Sub-persistence-length complex scaling behavior in lysozyme amyloid fibrils. Phys. Rev. Lett. *107*, 238101.

Lara, C., Handschin, S., and Mezzenga, R. (2013). Towards lysozyme nanotube and 3D hybrid self-assembly. Nanoscale *5*, 7197–7201.

Lara, C., Reynolds, N.P., Berryman, J.T., Xu, A., Zhang, A., and Mezzenga, R. (2014). ILQINS hexapeptide, identified in lysozyme left-handed helical ribbons and nanotubes, forms right-handed helical ribbons and crystals. J. Am. Chem. Soc.

Lecker, S.H., Goldberg, A.L., and Mitch, W.E. (2006). Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. J. Am. Soc. Nephrol. *17*, 1807–1819.

Lee, B.S., Shin, S., Jeon, J.Y., Jang, K., Lee, B.Y., Choi, S., and Yoo, T.H. (2015). Incorporation of Unnatural Amino Acids in Response to the AGG Codon. ACS Chem. Biol. *10*, 1648–1653.

Lee, S., Fernandez, E.J., and Good, T.A. (2007). Role of aggregation conditions in structure , stability , and toxicity of intermediates in the Ab fibril formation pathway. Protein Sci. *16*, 723–732.

Li, D., Furukawa, H., Deng, H., Liu, C., Yaghi, O.M., and Eisenberg, D.S. (2014). Designed amyloid fibers as materials for selective carbon dioxide capture. Proc. Natl. Acad. Sci. U. S. A. *111*, 191–196.

Li, J., Uversky, V.N., and Fink, A.L. (2001). Effect of Familial Parkinson's Disease Point Mutations A30P and A53T on the Structural Properties, Aggregation, and Fibrillation of Human alpha-Synuclein. Biochemistry *40*, 11604–11613.

Lian, H.-Y., Jiang, Y., Zhang, H., Jones, G.W., and Perrett, S. (2006). The yeast prion protein Ure2: structure, function and folding. Biochim. Biophys. Acta *1764*, 535–545.

Linke, R.P., Hampl, H., Lobeck, H., Ritz, E., Bommer, J., Waldherr, R., and Eulitz, M. (1989). Lysine-specific cleavage of beta 2-microglobulin in amyloid deposits associated with hemodialysis. Kidney Int. *36*, 675–681.

Lomenick, B., Olsen, R.W., and Huang, J. (2011). Identification of direct protein targets of small molecules. ACS Chem. Biol. *6*, 34–46.

Lorenzen, N., Cohen, S.I. a, Nielsen, S.B., Herling, T.W., Christiansen, G., Dobson, C.M., Knowles, T.P.J., and Otzen, D. (2012). Role of elongation and secondary pathways in S6 amyloid fibril growth. Biophys. J. *102*, 2167–2175.

Maji, S.K. (2011). Amyloid: A Natural Nanomaterial. Int. J. Nanosci. *10*, 909–918.

Mangione, P.P., Esposito, G., Relini, A., Raimondi, S., Porcari, R., Giorgetti, S., Corazza, A., Fogolari, F., Penco, A., Goto, Y., et al. (2013). Structure, Folding Dynamics, and Amyloidogenesis of D76N β2-Microglobulin: ROLES OF SHEAR FLOW, HYDROPHOBIC SURFACES, AND α-CRYSTALLIN. J. Biol. Chem. *288*, 30917–30930.

Marion, D. (2013). An introduction to biological NMR spectroscopy. Mol. Cell. Proteomics *12*, 3006–3025.

Marsh, E.N.G., and Suzuki, Y. (2014). Using 19F NMR to Probe Biological Interactions of Proteins and Peptides. ACS Chem. Biol. *9*, 1242–1250.

McKay, C.S., and Finn, M.G. (2014). Click Chemistry in Complex Mixtures: Bioorthogonal Bioconjugation. Chem. Biol. *21*, 1075–1101.

McParland, V.J., Kad, N.M., Kalverda, a P., Brown, a, Kirwin-Jones, P., Hunter, M.G., Sunde, M., and Radford, S.E. (2000). Partially unfolded states of beta(2)-microglobulin and amyloid formation in vitro. Biochemistry *39*, 8735–8746.

Merlini, G., and Bellotti, V. (2005). Lysozyme: A paradigmatic molecule for the investigation of protein structure, function and misfolding. Clin. Chim. Acta *357*, 168–172.

Michael, A. (1893). Ueber die Einwirkung von Diazobenzolimid auf Acetylendicarbonsäuremethylester. J. Für Prakt. Chemie *48*, 94–95.

Morris, G.A. (2007). Diffusion-Ordered Spectroscopy. In eMagRes,.

Morris, K.L., and Serpell, L.C. (2012). X-Ray Fibre Diffraction Studies of Amyloid Fibrils. In Amyloid Proteins Methods and Protocols, pp. 121–135.

Mossuto, M.F., Dhulesia, A., Devlin, G., Frare, E., Kumita, J.R., de Laureto, P.P., Dumoulin, M., Fontana, A., Dobson, C.M., and Salvatella, X. (2010). The non-core regions of human lysozyme amyloid fibrils influence cytotoxicity. J. Mol. Biol. *402*, 783–796.

Muir, T.W., Sondhi, D., and Cole, P. a (1998). Expressed protein ligation: a general method for protein engineering. Proc. Natl. Acad. Sci. U. S. A. *95*, 6705–6710.

Naiki, H., Higuchi, K., Hosokawa, M., and Takeda, T. (1989). Fluorometric determination of amyloid fibrils in vitro using the fluorescent dye, thioflavin T1. Anal. Biochem. *177*, 244–249.

Neumann, H., Wang, K., Davis, L., Garcia-Alai, M., and Chin, J.W. (2010). Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. Nature *464*, 441–444.

Nielsen, L., Khurana, R., Coats, a, Frokjaer, S., Brange, J., Vyas, S., Uversky, V.N., and Fink, a L. (2001). Effect of environmental factors on the kinetics of insulin fibril formation: elucidation of the molecular mechanism. Biochemistry *40*, 6036–6046.

Nilsson, M.R., and Dobson, C.M. (2003). Chemical modification of insulin in amyloid fibrils. Protein Sci. *12*, 2637–2641.

Olzscha, H., Schermann, S.M., Woerner, A.C., Pinkert, S., Hecht, M.H., Tartaglia, G.G., Vendruscolo, M., Hayer-Hartl, M., Hartl, F.U., and Vabulas, R.M. (2011). Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. Cell *144*, 67–78.

Pace, C.N., and McGrath, T. (1980). Substrate stabilization of lysozyme to thermal and guanidine hydrochloride denaturation. J. Biol. Chem. *255*, 3862–3865.

Paleologou, K.E., Oueslati, A., Shakked, G., Rospigliosi, C.C., Kim, H.-Y., Lamberto, G.R., Fernandez, C.O., Schmid, A., Chegini, F., Gai, W.P., et al. (2010). Phosphorylation at S87 is enhanced in synucleinopathies, inhibits alpha-synuclein oligomerization, and influences synuclein-membrane interactions. J. Neurosci. *30*, 3184–3198.

Pe'er, I., Felder, C.E., Man, O., Silman, I., Sussman, J.L., and Beckmann, J.S. (2004). Proteomic Signatures: Amino Acid and Oligopeptide Compositions Differentiate among Phyla. Proteins Struct. Funct. Genet. *54*, 20–40.

Pepys, M.B., Hawkins, P.N., Booth, D.R., Vigushin, D.M., Tennent, G. a, Soutar, a K., Totty, N., Nguyen, O., Blake, C.C., and Terry, C.J. (1993). Human lysozyme gene mutations cause hereditary systemic amyloidosis. Nature *362*, 553–557.

Pezza, J. a, Villali, J., Sindi, S.S., and Serio, T.R. (2014). Amyloid-associated activity contributes to the severity and toxicity of a prion phenotype. Nat. Commun. *5*, 4384.

Piejko, M., Dec, R., Babenko, V., Hoang, A., Szewczyk, M., Mak, P., and Dzwolak, W. (2015). Highly Amyloidogenic Two-chain Peptide Fragments Are Released upon Partial Digestion of Insulin with Pepsin. J. Biol. Chem. *290*, 5947–5958.

Pinotsi, D., Buell, A.K., Galvagnion, C., Dobson, C.M., Kaminski Schierle, G.S., and Kaminski, C.F. (2014). Direct observation of heterogeneous amyloid fibril growth kinetics via two-color super-resolution microscopy. Nano Lett. *14*, 339–345.

Platt, G.W., Xue, W.-F., Homans, S.W., and Radford, S.E. (2009). Probing dynamics within amyloid fibrils using a novel capping method. Angew. Chem. Int. Ed. Engl. *48*, 5705–5707.

Presolski, S.I., Hong, V., and Finn, M.G. (2011). Copper-Catalyzed Azide–Alkyne Click Chemistry for Bioconjugation. Curr. Protoc. Chem. Biol. *3*, 153–162.

Prusiner, S.B. (1982). Novel proteinaceous infectious particles cause scrapie. Science *216*, 136–144.

Prusiner, S., McKinley, M., and Bowman, K. (1983). Scrapie prions aggregate to form amyloid-like birefringent rods. Cell *35*, 349–358.

Quist, A., Doudevski, I., Lin, H., Azimova, R., Ng, D., Frangione, B., Kagan, B., Ghiso, J., and Lal, R. (2005). Amyloid ion channels: a common structural link for protein-misfolding disease. Proc. Natl. Acad. Sci. U. S. A. *102*, 10427–10432.

Rashidian, M., Dozier, J.K., Lenevich, S., and Distefano, M.D. (2010). Selective labeling of polypeptides using protein farnesyltransferase via rapid oxime ligation. Chem. Commun. (Camb). *46*, 8998–9000.

Rashidian, M., Mahmoodi, M.M., Shah, R., Dozier, J.K., Wagner, C.R., and Distefano, M.D. (2013a). A highly efficient catalyst for oxime ligation and hydrazone-oxime exchange suitable for bioconjugation. Bioconjug. Chem. *24*, 333–342.

Rashidian, M., Dozier, J.K., and Distefano, M.D. (2013b). Enzymatic labeling of proteins: Techniques and approaches. Bioconjug. Chem.

Rekas, A., Adda, C.G., Andrew Aquilina, J., Barnham, K.J., Sunde, M., Galatis, D., Williamson, N. a., Masters, C.L., Anders, R.F., Robinson, C. V., et al. (2004). Interaction of the molecular chaperone aB-crystallin with a-Synuclein: Effects on amyloid fibril formation and chaperone activity. J. Mol. Biol. *340*, 1167–1183.

Riedel, S. (2005). Edward Jenner and the history of smallpox and vaccination. BUMC Proc. *18*, 21–25.

Rosano, G.L., and Ceccarelli, E. a. (2014). Recombinant protein expression in Escherichia coli: Advances and challenges. Front. Microbiol. *5*, 1–17.

Rostovtsev, V. V, Green, L.G., Fokin, V. V, and Sharpless, K.B. (2002). A Stepwise Huisgen Cycloaddition Process: Copper (I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. Angew. Chemie Int. Ed. *41*, 2596–2599.

Sato, S.I., Kwon, Y., Kamisuki, S., Srivastava, N., Mao, Q., Kawazoe, Y., and Uesugi, M. (2007). Polyproline-rod approach to isolating protein targets of bioactive small molecules: Isolation of a new target of indomethacin. J. Am. Chem. Soc. *129*, 873–880.

Scheibel, T., Kowal, A.S., Bloom, J.D., and Lindquist, S.L. (2001). Bidirectional amyloid fiber growth for a yeast prion determinant. Curr. Biol. *11*, 366–369.

Scheibel, T., Parthasarathy, R., Sawicki, G., Lin, X.-M., Jaeger, H., and Lindquist, S.L. (2003). Conducting nanowires built by controlled self-assembly of amyloid fibers and selective metal deposition. Proc. Natl. Acad. Sci. U. S. A. *100*, 4527–4532.

Scherzinger, E., Lurz, R., Turmaine, M., Mangiarini, L., Hollenbach, B., Hasenbank, R., Bates, G.P., Davies, S.W., Lehrach, H., and Wanker, E.E. (1997). Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. Cell *90*, 549–558.

Schleeger, M., vandenAkker, C.C., Deckert-Gaudig, T., Deckert, V., Velikov, K.P., Koenderink, G., and Bonn, M. (2013). Amyloids: From molecular structure to mechanical properties. Polymer (Guildf). *54*, 2473–2488.

Schmohl, L., and Schwarzer, D. (2014). Sortase-mediated ligations for the site-specific modification of proteins. Curr. Opin. Chem. Biol. *22C*, 122–128.

Schuetz, A., Wasmer, C., Habenstein, B., Verel, R., Greenwald, J., Riek, R., Böckmann, A., and Meier, B.H. (2010). Protocols for the sequential solid-state NMR spectroscopic assignment of a uniformly labeled 25 kDa protein: HET-s(1-227). ChemBioChem *11*, 1543–1551.

Sélo, I., Clément, G., Bernard, H., Chatel, J., Créminon, C., Peltre, G., and Wal, J. (1999). Allergy to bovine beta-lactoglobulin: specificity of human IgE to tryptic peptides. Clin. Exp. Allergy *29*, 1055–1063.

Seo, J.-W., and Lee, K.-J. (2004). Post-translational Modifications and Their Biological Functions: Proteomic Analysis and Systematic Approaches. J. Biochem. Mol. Biol. *37*, 35–44.

Serpell, L.C. (2000). Alzheimer's amyloid fibrils: structure and assembly. Biochim. Biophys. Acta - Proteins Proteomics *1502*, 16–30.

Shi, L., and Ladizhansky, V. (2012). Magic Angle Spinning Solid-State NMR Experiments for Structural Characterization of Proteins. In Intrinsically Disordered Protein Analysis, V.N. Uversky, and A.K. Dunker, eds. (Springer Science), pp. 153–165.

Shih, H.-W., Kamber, D.N., and Prescher, J. a (2014). Building better bioorthogonal reactions. Curr. Opin. Chem. Biol. *21*, 103–111.

Shorter, J., and Lindquist, S. (2004). Hsp104 catalyzes formation and elimination of self-replicating Sup35 prion conformers. Science *304*, 1793–1797.

Sipe, J.D., and Cohen, a S. (2000). Review: history of the amyloid fibril. J. Struct. Biol. *130*, 88–98.

Sletten, E.M., and Bertozzi, C.R. (2009). Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. Angew. Chem. Int. Ed. Engl. *48*, 6974–6998.

Smith, D.P., Tew, D.J., Hill, A.F., Bottomley, S.P., Masters, C.L., Barnham, K.J., and Cappai, R. (2008). Formation of a high affinity lipid-binding intermediate during the early aggregation phase of alpha-synuclein. Biochemistry *47*, 1425–1434.

Smith, J.F., Knowles, T.P.J., Dobson, C.M., Macphee, C.E., and Welland, M.E. (2006). Characterization of the nanoscale properties of individual amyloid fibrils. Proc. Natl. Acad. Sci. U. S. A. *103*, 15806–15811.

Smyth, M.S., and Martin, J.H. (2000). X Ray Crystallography. Mol. Pathol. *53*, 8–14.

Soto, C. (2001). Protein misfolding and disease; protein refolding and therapy. FEBS Lett. *498*, 204–207.

Souza, J.M., Giasson, B.I., Lee, V.M.Y., and Ischiropoulos, H. (2000). Chaperone-like activity of synucleins. FEBS Lett. *474*, 116–119.

Starck, C.S., and Sutherland-Smith, A.J. (2010). Cytotoxic aggregation and amyloid formation by the myostatin precursor protein. PLoS One *5*, e9170.

Stejskal, E.O., and Tanner, J.E. (1965). Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. J. Chem. Phys. *42*, 288.

Stoppini, M., and Bellotti, V. (2015). Systemic Amyloidosis: Lessons from β2-Microglobulin. J. Biol. Chem. *290*, 9951–9958.

Struhl, K. (2008). The hisB463 mutation and expression of a eukaryotic protein in Escherichia coli. Genetics *180*, 709–714.

Strynadka, N.C.J., and James, M.N.G. (1991). Lysozyme revisited: Crystallographic evidence for distortion of an N-acetylmuramic acid residue bound in site D. J. Mol. Biol. *220*, 401–424.

Sweers, K., van der Werf, K., Bennink, M., and Subramaniam, V. (2011). Nanomechanical properties of α-synuclein amyloid fibrils: a comparative study by nanoindentation, harmonic force microscopy, and Peakforce QNM. Nanoscale Res. Lett. *6*, 270.

Tanaka, M. (2010). A protein transformation protocol for introducing yeast prion particles into yeast. In Methods in Enzymology, (Elsevier Inc), pp. 681–693.

Tanaka, M., Chien, P., Naber, N., Cooke, R., and Weissman, J. (2004). Conformational variations in an infectious protein determine prion strain differences. Nature *428*, 323–328.

Tanaka, M., Collins, S.R., Toyama, B.H., and Weissman, J.S. (2006). The physical basis of how prion conformations determine strain phenotypes. Nature *442*, 585–589.

Teplow, D.B. (2006). Preparation of amyloid beta-protein for structural and functional studies. Methods Enzymol. *413*, 20–33.

Thompson, D. a, Evans, E.G.B., Kasza, T., Millhauser, G.L., and Dawson, P.E. (2014). Adapter reagents for protein site specific dye labeling. Biopolymers *102*, 273–279.

Tipping, K.W., Karamanos, T.K., Jakhria, T., Iadanza, M.G., Goodchild, S.C., Tuma, R., Ranson, N. a., Hewitt, E.W., and Radford, S.E. (2015). pH-induced molecular shedding drives the formation of amyloid fibril-derived oligomers. Proc. Natl. Acad. Sci. *112*, 5691–5696.

Tornøe, C.W., Christensen, C., and Meldal, M. (2002). Peptidotriazoles on Solid Phase : [1,2,3] -Triazoles by Regiospecific Copper (I) -Catalyzed 1, 3-Dipolar Cycloadditions of Terminal Alkynes to Azides. J. Org. Chem. *67*, 3057–3064.

True, H.L., and Lindquist, S.L. (2000). A yeast prion provides a mechanism for genetic variation and phenotypic diversity. Nature *407*, 477–483.

True, H.L., Berlin, I., and Lindquist, S.L. (2004). Epigenetic regulation of translation reveals hidden genetic variation to produce complex traits. Nature *431*, 184–187.

Tsukiji, S., and Nagamune, T. (2009). Sortase-mediated ligation: a gift from Gram-positive bacteria to protein engineering. Chembiochem *10*, 787–798.

Tuite, M.F., and Cox, B.S. (2006). The [PSI+] prion of yeast: A problem of inheritance. Methods *39*, 9–22.

Uéda, K., Fukushima, H., Masliah, E., Xia, Y., Iwai, a, Yoshimoto, M., Otero, D. a, Kondo, J., Ihara, Y., and Saitoh, T. (1993). Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. Proc. Natl. Acad. Sci. U. S. A. *90*, 11282–11286.

Uversky, V.N., Li, J., and Fink, A.L. (2001). Metal-triggered structural transformations, aggregation, and fibrillation of human a-synuclein: A possible molecular link between parkinson's disease and heavy metal exposure. J. Biol. Chem. *276*, 44284–44296.

VandenAkker, C.C., Engel, M.F.M., Velikov, K.P., Bonn, M., and Koenderink, G.H. (2011). Morphology and persistence length of amyloid fibrils are correlated to peptide molecular structure. J. Am. Chem. Soc. *133*, 18030–18033.

Vlad, C., Lindner, K., Karreman, C., Schildknecht, S., Leist, M., Tomczyk, N., Rontree, J., Langridge, J., Danzer, K., Ciossek, T., et al. (2011). Autoproteolytic Fragments Are Intermediates in the Oligomerization/Aggregation of the Parkinson's Disease Protein Alpha-Synuclein as Revealed by Ion Mobility Mass Spectrometry. Chembiochem *12*, 2740–2744.

Walter, P., and Ron, D. (2011). The unfolded protein response: from stress pathway to homeostatic regulation. Science (80-.). *334*, 1081–1087.

Wang, C.K., Northfield, S.E., Swedberg, J.E., Harvey, P.J., Mathiowetz, A.M., and David, A. (2014). Translational diffusion of cyclic peptides measured using pulsed-field gradient NMR. J. Phys. Chem. B 1–26.

Wang, Q., Parrish, A.R., and Wang, L. (2009a). Expanding the genetic code for biological studies. Chem. Biol. *16*, 323–336.

Wang, X., Das, T.K., Singh, S.K., and Kumar, S. (2009b). Potential aggregation prone regions in biotherapeutics: A survey of commercial monoclonal antibodies. MAbs *1*, 254–267.

Wickner, R.B. (1994). [URE3] as an altered URE2 protein: evidence for a prion analog in Saccharomyces cerevisiae. Science *264*, 566–569.

Wickner, R.B., Edskes, H.K., Ross, E.D., Pierce, M.M., Baxa, U., Brachmann, A., and Shewmaker, F. (2004). Prion genetics: new rules for a new kind of gene. Annu. Rev. Genet. *38*, 681–707.

Wittig, G., and Krebs, A. (1961). Zur Existenz niedergliedriger Cycloalkine, I. Chem. Ber. *94*, 3260–3275.

Wu, K.P., Weinstock, D.S., Narayanan, C., Levy, R.M., and Baum, J. (2009). Structural Reorganization of α-Synuclein at Low pH Observed by NMR and REMD Simulations. J. Mol. Biol. *391*, 784–796.

Wu, Z.L., Huang, X., Burton, A.J., and Swift, K. a. D. (2015). Glycoprotein labeling with click chemistry (GLCC) and carbohydrate detection. Carbohydr. Res. *412*, 1–6.

Xie, J., and Schultz, P.G. (2005). Adding amino acids to the genetic repertoire. Curr. Opin. Chem. Biol. *9*, 548–554.

Xu, C., and Ye, L. (2011). Clickable molecularly imprinted nanoparticles. Chem. Commun. (Camb). *47*, 6096–6098.

Xu, M., and Lewis, R. V (1990). Structure of a protein superfiber: spider dragline silk. Proc. Natl. Acad. Sci. U. S. A. *87*, 7120–7124.

Xu, M.Q., and Perler, F.B. (1996). The mechanism of protein splicing and its modulation by mutation. EMBO J. *15*, 5146–5153.

Xu, S., Falvey, D. a, and Brandriss, M.C. (1995). Roles of URE2 and GLN3 in the proline utilization pathway in Saccharomyces cerevisiae. Mol. Cell. Biol. *15*, 2321–2330.

Xue, W.-F., and Radford, S.E. (2013). An imaging and systems modeling approach to fibril breakage enables prediction of amyloid behavior. Biophys. J. *105*, 2811–2819.

Xue, W.F., Homans, S.W., and Radford, S.E. (2009a). Amyloid fibril length distribution quantified by atomic force microscopy single-particle image analysis. Protein Eng. Des. Sel. *22*, 489–496.

Xue, W.-F., Homans, S.W., and Radford, S.E. (2008). Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. Proc. Natl. Acad. Sci. U. S. A. *105*, 8926–8931.

Xue, W.-F., Hellewell, A.L., Gosal, W.S., Homans, S.W., Hewitt, E.W., and Radford, S.E. (2009b). Fibril fragmentation enhances amyloid cytotoxicity. J. Biol. Chem. *284*, 34272–34282.

Yang, C., Mi, J., Feng, Y., Ngo, L., Gao, T., Yan, L., and Zheng, Y.G. (2013). Labeling lysine acetyltransferase substrates with engineered enzymes and functionalized cofactor surrogates. J. Am. Chem. Soc. *135*, 7791–7794.

Young, C.L., Britton, Z.T., and Robinson, A.S. (2012). Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. Biotechnol. J. *7*, 620–634.

Yumlu, S., Barany, R., Eriksson, M., and Röcken, C. (2009). Localized insulin-derived amyloidosis in patients with diabetes mellitus: a case report. Hum. Pathol. *40*, 1655–1660.

Appendix 1

PolProp script

```
%Extracts polymer information from the ans file
fib=img.features.polymer;
%generates a value for the no. of polymers in the file
s=size(fib,1);
%A loop which extracts the Contour length, Height, Mode Height and
Mean
%Height from the ans file.
for n=1:s;
    l{n}=fib(n,1).lContour;
    h{n}=fib(n,1).z;
    hparmean{n}=mean(fib(n,1).z)';
    hparmode{n}=mode(fib(n,1).z)';
end
%Converts Mean Particle Height into Matrix from cell
MeanPartHeight=cell2mat(hparmean);
%Calculates the Mean of the mode particle heights
ParticleMeanHeightMean=mean2(MeanPartHeight)
%Converts Mode Particle Height into matrix from cell
ModePartHeight=cell2mat(hparmode);
%Calculates the Mean of the mode particle heights
ParticleModeHeightMean=mean2(ModePartHeight)
%Extracts the max no. of data points in each cell.
maxSize = max(cellfun(@numel,h));
%Expands the cells into a matrix padding any values up to the
maximum with
%NaN
H=cell2mat(cellfun(@(x)cat(1,x,NaN(maxSize-
length(x),1)),h,'UniformOutput',false));
%Concatenates the columns into one column
Height=vertcat(H(:),1);
%All NaNs removed
Height(any(isnan(Height),2),:) = [];
%The total no. of height data from traced fibrils
TracedHeight=size(Height);
%Selects the first value of the matrix, removes the ,1.
TracedHeightPixels=TracedHeight(1,1)
%Finds any pixel with a value greater than 2nm, 1nm and 0.5nm
TotalHeightover2=find(img.z>2);
TotalHeightover1=find(img.z>1);
TotalHeightover0_5=find(img.z>0.5);
%Calculates the no. pixels over 2nm
NumofPixelsover2=length(TotalHeightover2)
NumofPixelsover1=length(TotalHeightover1)
NumofPixelsover0_5=length(TotalHeightover0_5)
%Converts the l cell array into a matrix Len
Len=cell2mat(l);
%Flips the horizontal orientation to vertical
Len=Len';
%Converts the no. of pixels to the length in nm (9.7704 dependent on
%resolution of AFM scanning 20um2 at 2048 x 2048 and 10uM2 at 1024x
1024 =
%9.7704
CL=Len*9.7704;
%The contour length values squared
CL2=CL.^2;
```

```matlab
%The sum of the squared contour lengths - single as data is double
SumCL2=sum(single(CL2));
%The sum of the contour lengths - single as data is double
SumCL=sum(single(CL));
%Generation of the length of the longest fibril in nm
LongestFib=max(CL)
%loop which finds particles which are of length less than or equal
to the
%value p from 1 up to the longest fibril length
for p=1:LongestFib
    CDF1{p}=find(CL<=p);
end
%Extracts the no. of particles in the cells generated from the loop
CDF2=cellfun('prodofsize', CDF1);
%Calculation of the Cumulative Distribution function - each value of
CDF2
%divided my the total no. of fibrils
CDFy=CDF2/n;
%Generation of Cumulative Distribution Function x axis in nm - upto
the
%maximum fibril length
CDFx=1:p;
%Calculation of Contour Length weighted mean
ContourLengthWeightMean=SumCL2/SumCL
ContourLengthWeightMeanSEM=(SumCL2/SumCL)/sqrt(s)
%Calculation of Mean Contour Length
ContourLengthMean=mean2(CL)
%Calculation of standard error of mean for the Contour Lengths
ContourLengthSEM=std(CL)/sqrt(s)
%Calculation of the mean Height from all height values
HeightMean=mean2(Height)
%Calculation of the Standard Deviation for Heights (all height
values)
HeightSD=std(Height)
```