Mechanical response of Amyloid fibrils probed by molecular dynamics simulation.

Hlengisizwe Ndlovu

School of Physics and Astronomy

University of Leeds

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

February 2013

Publication Statement:

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Jointly authored publications:

The work described in Chapter 3 is based on the publication in *Biophysical Journal* (2012) entitled *Effect of Sequence Variation on the Mechanical Response of Amyloid Fibrils Probed by Steered Molecular Dynamics Simulation* by H Ndlovu, AE Ashcroft, SE Radford & SA Harris. The contribution of the candidate was to wholly originate, design, execute and analyse the simulations described in this paper under the direct supervision and guidance of AE Ashcroft, SE Radford & SA Harris. The candidate co-wrote the bulk of the text contained in this article with assistance from SA Harris whilst AE Ashcroft & SE Radford provided amendments of an editorial nature.

The work presented in Chapter 4 has been accepted for publication in the *Beilstein Journal of Nanotechnology* (2013) after successful peer review. It will appear in print under the title: *Molecular dynamics simulations of mechanical failure in polymorphic arrangements of amyloid fibrils containing structural defects* by H Ndlovu, AE Ashcroft, SE Radford & SA Harris. The contributions of the candidate and co-authors are identical to those stated above. The candidate co-wrote the bulk of the text contained in this article with assistance from SA Harris.

Acknowledgements:

Thanks: I would like to thank everyone associated with the White-Rose Doctoral Training Center for an interesting interdisciplinary experience at the life/physical sciences interface. I acknowledge the the EPRSC for provision of funding not only the research but also my day to day existence. Additionally, I thank members of the Polymers & Complex Fluids group (Soft Matter Physics) for the seminars, Monday lunchtime talks, gradclub and random cake sessions over the years. In terms of the journey of scientific discovery during the course of my research, I acknowledge my supervisory team, Prof Alison Ashcroft, Prof Sheena Radford and Dr Sarah Harris. Thank you for the advice, latitude and space given to allow me to pursue a good chunk of this work 'semi-autonomously'. I am sure this will stand me in good stead whatever future career path I embark on. To my Mum, Dad, Brother and Sisters thank you for the support, unconditional love and for providing the environment that made me who I am. Finally to my wife Lucy and son Nathan, thank you for keeping me grounded in reality and above all letting me pursue my childhood ambition of being a scientist. A very big thank you the Lord God almighty with whom all things are evidently possible.

The effort put into this thesis is dedicated to the loving memory of my mum.
22 September 1956 - 19 March 2011.

Abstract

The presence of self-assembled protein aggregates known as Amyloid fibrils are associated with an increasing number of human conditions such as type-II Diabetes, Parkinson's, Huntington's and Alzheimer's disease. A link has been made between the fragility of these normally robust structures and an enhancement of their toxic effects. This highlights a need for a firm understanding of the factors that govern their mechanical properties if effective therapeutic strategies are to be developed. The main aims of this thesis are to probe, at a molecular level, the key interactions that contribute to the mechanical stability generically exhibited by amyloid fibril systems and then explore ways in which these may be modulated. A variety of atomistic detailed fibrils are computationally modelled and then subjected to molecular dynamics steering forces from a variety of directions that characterise their fragmentation through disruption of the stabilising interactions. This work focuses on models of proto-fibrils associated with type-II Diabetes and mature fibres linked with Alzheimer's disease. Three separate investigations were undertaken to examine firstly the role played by the peptide sequence in mechanical resistance, then ways of modulating mechanical failure in polymorphic fibril arrangements and finally the mechanical effects of heterologous cross-seed interfaces. The presence of structural defects in the otherwise highly ordered protein aggregates were consistently found to be able to dominate the overall mechanical characteristics of the fibrils. The results of the simulations give further insights into the fracture mechanisms as well as have wider implications for both therapeutic and potential nano-technological applications.

Contents

Abbreviations

Aβ	Amyloid beta
AFM	Atomic Force Microscope
C$_\alpha$	Carbon alpha
cv-SMD	Constant velocity Steered Molecular Dynamics
E.coli	Escherichia coli
EM	Electron Microscopy
FAD	Familial Alzheimer's Disease
hIAPP	human Islet Amyloid Polypeptide
IAPP	Islet Amyloid Polypeptide
JE	Jarzynski equality
k_BT	Thermal energy (Boltzmann constant & temperature)
MD	Molecular Dynamics
MM-PBSA	Molecular Mechanics Poisson Boltzmann Surface Area
NMR	Nuclear Magnetic Resonance
ns	nano second
PDB	Protein Data Bank
PMF	Potential of Mean Force
rIAPP	rat Islet Amyloid Polypeptide
RMSD	Root mean square deviation
S_c	Shape complementarity
SMD	Steered Molecular Dynamics
ssNMR	Solid state nuclear magnetic resonance
WT	Wild-type

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Protein Basics

Proteins are a class of biomolecules that are ubiquitous in all living organisms. They are more abundant and structurally varied relative to the other classes of biomolecules such as nucleic acids, sugars and lipids. They utilise their unique tertiary structures to carry out a host of activities within and outside of the cellular environment. The diversity of functions proteins perform include acting as enzymatic catalysts, ion channels and gates, structural scaffolds, molecular motors & switches, signal receptors/transmitters as well as gene regulators. Each of these distinct functions is directly related to the proteins unique 3-D structure, which is derived from its amino-acid building blocks. The amino-acids form the main chemical composition of proteins and are attached as side-chain residues on the backbone of a polypeptide. There are 20 known naturally occurring amino-acids and these have properties that range from being (un)polar, (un)charged, hydrophobic, hydrophilic to acidic or basic. The order in which the amino-acid sequence is arranged (primary structure) determines the properties of the polypeptide as it influences the native fold adopted. As the shape of a protein is closely related to its function, different combinations of the primary sequence result in a uniquely folded protein. This, in part, accounts for the multitude of roles these biomolecules are able to undertake.

1.1.1 Folding and Misfolding

How proteins transform from an unfolded or linear polypeptide chain to the native functional state is a subject that has been widely studied though not yet fully understood. Current models to describe folding events treat the process as stochastic, resulting from the combined effects of a number of interactions (1; 2; 3). These are thought to originate from the intrinsic properties of the amino-acid sequence which give rise to van der Waal, hydrophobic, coulombic and hydrogen bond interactions. Folding models utilise the concept of a funnel shaped rugged energy landscape on which a polypeptide chain would explore the available conformational space before reaching a unique native fold at the global minimum (3; 4). The energy landscape is distinct and specific to each polypeptide chain as it is encoded into the amino-acid sequence.

During the folding process, a peptide exploring its conformational space within the energy landscape may become kinetically trapped in a local minimum away from the global minimum (fig 1.1). This would leave it in a partially or mis-folded state that is effectively unable

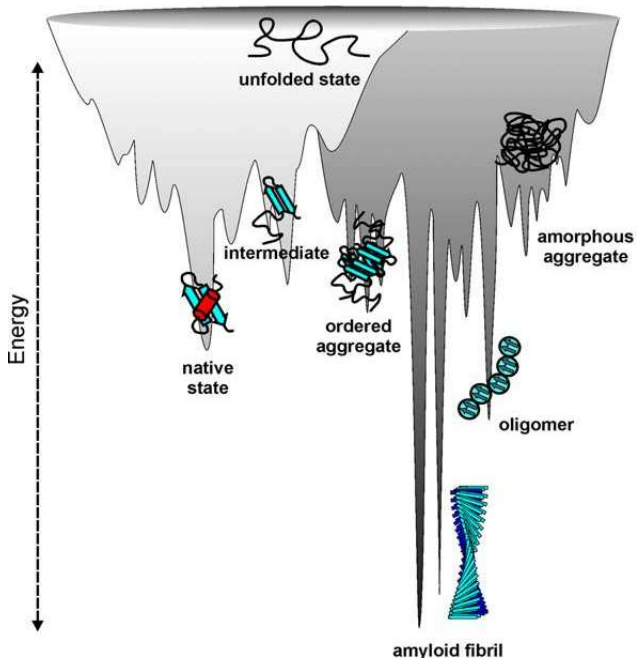

Figure 1.1: A sketch of the conformational energy landscape available to an unfolded protein. The region on the left (light grey) depicts a pathway that eventually leads to the correct native fold. The region on the right (dark-grey) relates to the landscape in which the mis-folded proteins aggregate to form amyloid fibrils. (Image reproduced from (3)).

to function in its shape related role. Given the biological importance of the range of protein functions, the predominance of mis-folded proteins may have severe consequences for the host organism. As such, a number of mechanisms are in place to identify and aid the mis-folded proteins reach the desired fold. This is carried out by a class of molecular chapreons such as GroEL and Heat Shock Proteins (HSP) which not only ensure the correct fold is achieved but also that the folding process occurs at a rapid rate (5). In cases where the mis-folded peptide is unable to correctly fold despite repeated attempts by the chapreons, it is eventually marked for proteolytic degradation.

However, under certain environmental conditions the meta-stable partially folded proteins may increase in population. These then self-assemble via nucleation dependent polymerisation mechanisms into ordered aggregates known as amyloid fibrils (1; 3; 6). The formation of the amyloid aggregates is usually undesirable as their occurrence *in vivo* is almost invariably linked a number of fatal human diseases. The interactions involved in attaining

the native protein fold are ironically thought to be the same as those that drive both the mis-folding and aggregation events (2; 4).

The work outlined in this thesis is primarily concerned with understanding the factors that influence mechanical stability in amyloid fibrils, through the use of computational simulation methods. The relative importance of the mechanical characteristics of fibrils is exhibited in the fibril propagation kinetics as well as fibril toxicity. The remaining sections of this chapter thus serve as both an introduction to and a discussion of the nature of amyloid fibrils. Strong emphasis throughout the discussion is placed on how the mechanical properties relate to various aspects associated with the amyloid state. The chapter then concludes with an explicit statement of the overall aims of the thesis.

1.2 Amyloid Fibrils

Introduction

Amyloid fibrils are formed when peptides or mis-folded proteins self-assemble into filamentous aggregates. Although the presence of the amyloid aggregates is associated with a number of neurodegenerative and pathological human diseases (7; 8) it is still not clear if the fibrils are the cause or merely symptomatic by-products of the disease state (9). As such, a lot of research has gone into understanding not only the mechanisms of amyloid formation but also their physical and structural properties. The goal of such research has focused on the development of therapeutic measures to alleviate effects of amyloid related disease as well as to gain fundamental insights into protein (mis)folding.

1.2.1 Kinetics of amyloid formation

Amyloidogenic proteins are able to form fibrils both *in vivo* and *in vitro*. Observations of *in vitro* amyloid formation suggest that the fibrillisation events proceeds via a nucleation dependant pathway in which fibril growth is achieved as protein monomers are added to a transiently populated aggregation nucleus (2; 3; 4). The formation of the nucleus is thermodynamically unfavourable as there is a characteristic lag phase prior to any fibril growth. Once the initial nucleus has formed however, fibril elongation becomes very favourable and growth proceeds rapidly until the available monomers have been consumed. The lag

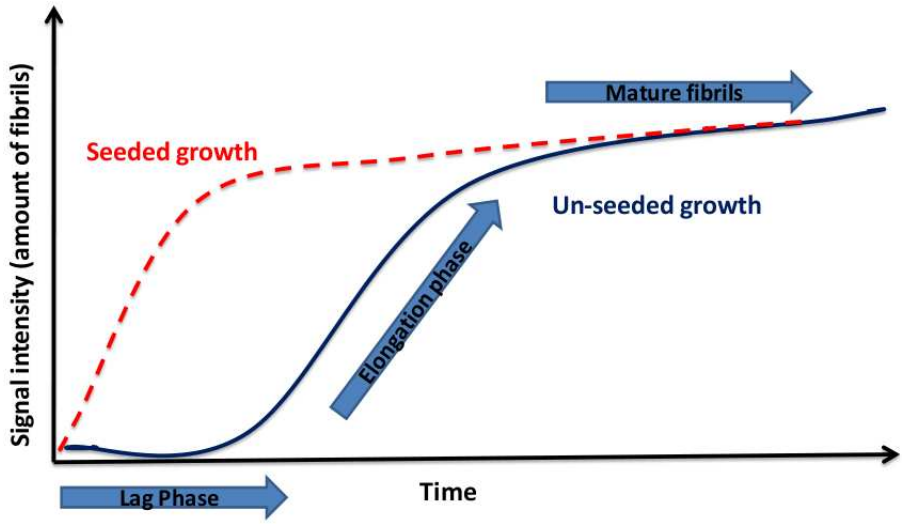

Figure 1.2: Schematic representation of typical growth kinetics for amyloid fibrils. The blue curve shows un-seeded growth in which there is a characteristic lag phase followed by a rapid elongation phase. The red curve shows elimination of the lag phase when seeding material is added.

phase can be significantly shortened or eliminated by 'seeding' growth with preformed fibril/peptide fragments that ensure fibrillisation kinetics directly go the elongation phase. The characteristic kinetic profiles exhibited during normal and seeded fibril growth are contrasted in figure 1.2. Since the seed fragments are nucleation sites on which polymerisation is initiated, there is a resultant increased rate of fibril formation (2; 3).

At present, there is neither a firm understanding of the physical mechanisms that drive the formation of the precursor transient oligomers that precede amyloid fibrillisation nor is there any structural information about the oligomers available. In this regard, computer models and simulation have proved useful in the characterisation of nucleation barriers of amyloid formation. Coarse grained molecular dynamics on the self-assembly of a system made up of eighty 12-residue peptides have, for instance, suggested a two step condensation ordering mechanism in which the disordered oligomeric aggregates are formed first which then reorganise into ordered fibrils (6). The form of the free energy barriers calculated from such simulations provide further insights and rationalisation for the universal features of protein aggregation e.g non-specific hydrophobic interactions and hydrogen bonding.

The fibrillisation kinetics can be further understood if the breakage or fragmentation of growing individual fibrils is taken into account (10; 11). *In vitro* experimental studies into

the effects of mechanical agitation during the polymerisation phase highlight a drastic reduction in the time taken to form fibrils, with the lag phase visibly absent (10). This is consistent with nucleated growth as the fibril fragments provide additional sites, on either end, on which monomers can attach thus increasing the amount of fibril material. Fragmentation events *in vivo* may occur either spontaneously due to thermal fluctuations, applied mechanical stresses or through the action of molecular chaperones such as HSP104 (12). Consequently, a firm understanding of the mechanical character of the fibrils may prove invaluable in designing therapeutic strategies against amyloid related diseases. Additionally, such insights would be a requirement in the development of effective fibril based nano-materials. These ideas are discussed further in section §1.3.

1.2.2 Properties and Structural organisation

The classification of a proteinaceous aggregate as amyloid is dependent on it satisfying a number of criteria. For historical reasons and disease association, amyloid has been defined from a pathological perspective. In this viewpoint, amyloid is an aggregate of un-branched protein fibers that are found *in vivo* as either extra or intra-cellular deposits whose presence is linked to disease (13). There is also an additional requirement that these aggregates bind to Congo-red dye which then displays green birefringence when viewed with crossed polarisers (1; 13). However, given that a host of proteins and peptides are able to form fibrils *in vitro* that are structurally similar to those found *in vivo* which may have no obvious disease association, a structural based definition of amyloid has been more useful. In this structural viewpoint, amyloid is best defined as any self-assembled proteinaceous aggregate of un-branched fibres that displays a cross-β fibre diffraction pattern and also preferentially binds to dyes such as Congo-red and Thioflavin-T.

Attempts at structural determination of amyloid fibrils by traditional biophysical techniques such as X-ray crystallography and NMR have proved unsuccessful as the fibrils not only lack long range order but are also insoluble. This has meant structural information has been limited to models inferred from a combination of AFM/EM images, x-ray fiber diffraction patterns, and solid state NMR distance restraints. An interesting outcome from these techniques has been the observation that regardless of the peptide sequence under study, all amyloid fibrils posses similar structural features.

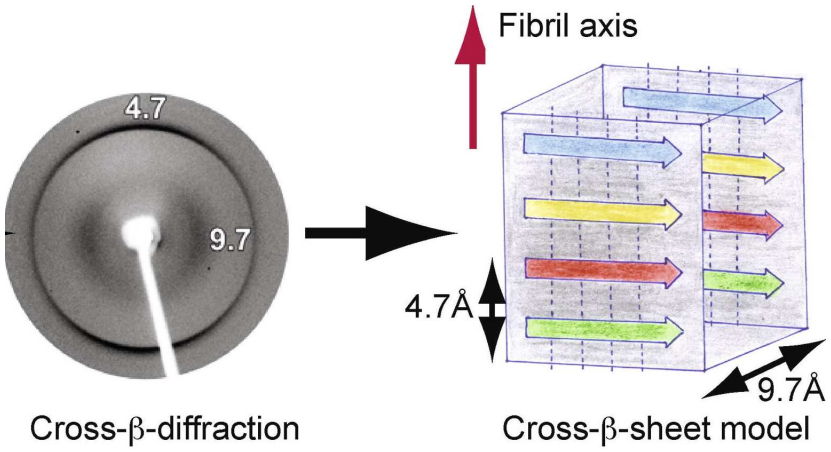

Figure 1.3: X-ray fibre diffraction pattern for amyloid from which the cross-β structural organisation is inferred. The characteristic reflections at 4.7 Å and 9.7 Å relate to inter-strand separation and inter-sheet side chain spacing respectively. (Reproduced from (15)).

Images from AFM and EM probing for instance collectively show that the mature fibrils have an un-branched morphology (14). Meanwhile, X-ray fibre diffraction patterns reveal that fibrils are composed of stacked β-sheets that are parallel to the fibril long axis, with the peptide strands perpendicular to this axis (1; 3; 4). This is referred to as the cross-β motif and it is a defining feature of all amyloid. Figure 1.3 shows a typical cross-β fibre diffraction pattern from which the structural organisation within the fibril is inferred. The meridional reflection signal at $\approx$ 4.7 Å gives the spacing between individual peptides in a single β-sheet, each of which adopts the extended beta-strand motif. An equatorial reflection at $\approx$ 10 Å in the pattern represents the separation between a pair of interfaced stacked β-sheets. The individual peptides within a single β-sheet are held together by inter-peptide backbone hydrogen bonding whilst the interactions between the pair of sheets are mainly driven by hydrophobic forces.

Polymorphism in Amyloid

Even though all amyloid fibrils share the the same cross-β motif as a building block despite differences in the amino-acid sequence, there is nonetheless a wide variety of morphologies adopted by the mature fibrils once they form. These include sheets, ribbons, fibers and tubular structures (fig 1.4). When environmental conditions such as pH, temperature and salt concentration are varied, diversity in fibril morphology can also manifest even when

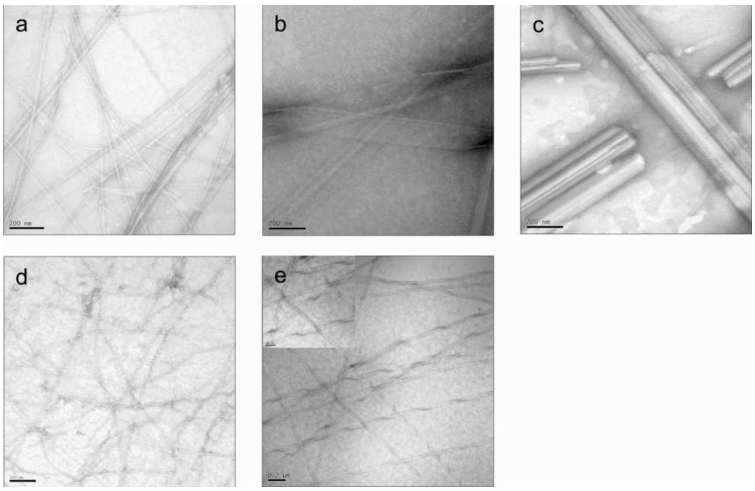

Figure 1.4: Transmission Electron Microscopy images of different fibril morphologies formed under various environmental conditions. The amyloidogenic GNNQQNY peptide can be found in a variety of fibrillar forms that include (a) straight fibres, (b) ribbons or (c) micro-crystals. Other possible morphologies include twisted fibrils as shown in panels (d) and (e) for Aβ_{40} and KFFEAAAKFFE peptides respectively (Adapted from (14)).

the same peptide sequence is used. This characteristic is known as polymorphism and is an intrinsic property of amyloidogenic peptides. The polymorphs may exhibit different chemical properties in addition to their distinct structural organisation. A number of stable polymorphic fibril forms have also been found to exist simultaneously under identical solution conditions (1; 3; 16; 17). It is posited that the ability to form polymorphs maybe the mechanism by which different strains of infectious prions are produced (18).

Polymorphism is possible due to the number of structural organisation options available to the fibrils. Firstly, changes in the numbers of proto-filaments that twist round each other to make a fiber can give rise to different types of morphologies. Secondly, the relative arrangements of the proto-filaments within a fibril such as registry shifts between a pair of β-sheets also contribute to the multitude of polymorphic forms available. Finally, the arrangements of the peptides within the β-sheet (parallel or anti-parallel) and the orientations of stacked pairs of β-sheets relative to each other increase the diversity of fibril morphologies available to a particular sequence. The latter structural organisational option is associated with so-called 'steric zipper' classes which are described in following section.

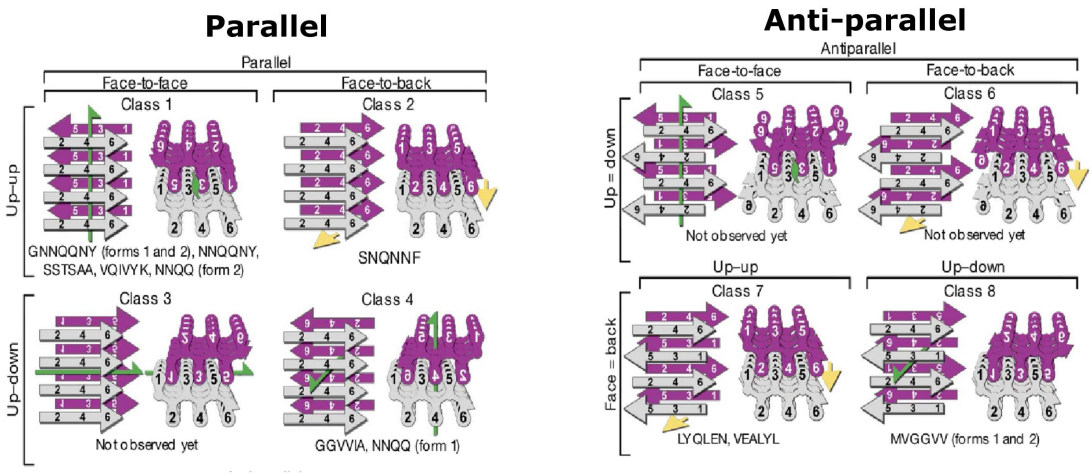

Figure 1.5: The eight possible steric zipper classes relating to polymorphic arrangements of peptides within fibrils. (Image adapted from (20)).

Steric Zipper classes

So-called steric zippers where first observed in X-ray crystallography experiments on micro-crystal arrays formed from short segments of full length amyloidogenic proteins (19; 20). In these experiments, both fibrils and micro-crystals of the short peptides (e.g GNNQQNY) were shown to have a similar morphology to fibrils obtained from the parent proteins. The micro-crystals revealed that the side-chains on interacting stacked pairs of sheets fit tightly together in a very complementary manner like teeth in a zipper (19). This tight fitting interface is devoid of water molecules and is thought to be a fundamental structural feature of the fibrils. The experiments and subsequent symmetry considerations suggest that there are eight possible configurations that a pair of β-sheets may adopt (20). The eight steric zipper classes are defined in terms of; the orientations of the strands within a sheet (parallel vs anti-parallel), the orientation of strands between the sheets (up vs down) as well as the relative orientations of the pair of sheets with each other (face to face vs face to back). Figure 1.5 shows the peptide arrangements of each of the eight steric zipper classes.

In each class of steric zippers, the major interactions responsible for the stability of the fibrils is the dense network of hydrogen bonds between peptide ladders within each β-sheet. It would be expected, if one considers the Pauling-Corey model of a β-sheet, that the anti-parallel arrangements of peptide strands would be more favourable than the parallel

counterpart.This assertion has been confirmed by a systematic study of all eight steric zipper classes modelled on fibrils formed by various (non-G/Q rich) peptide sequences. The symmetry class was found to directly affect fibril stability, with anti-parallel arrangements being more thermodynamically stable than the parallel ones (21). Interestingly however, both parallel and anti-parallel steric zipper polymorphic forms of an amyloidogenic segment of the human amylin protein (hIAPP) have been experimentally observed simultaneously (16; 22). Both these polymorphic forms are the subject of the inquiry described in chapter 4.

1.3 Amyloid Perspectives: The Bad, the Good & the Useful

Perspectives of the amyloid state in relation to native protein folds have greatly shifted since a study revealed that the ability to form amyloid structures is a generic property of all peptides, regardless of amino-acid sequence (23). In this generic hypothesis of amyloid formation, it is surmised that the ability to assemble into ordered cross-β structures is not an unusual feature possessed by a small number of proteins/peptides with special sequence or structural properties but rather an inherent characteristic of all polypeptide chains (5; 8; 23; 24). This means that the amyloid state can be viewed as an alternate conformational state available to the protein in addition to the the native fold. It is thought that the energy minimum for individual chains in mature amyloid fibrils may be much deeper and sharper than that of the native monomeric proteins (fig 1.1). A consequence of this paradigm shift is that focus has moved on from just the undesirable effects of amyloid to understanding how nature utilises the properties of this protein state. The section below gives three different accounts of how amyloid may be viewed as either undesirable, biologically functional or as a material in nano-technology devices.

1.3.1 The Bad: Disease implicated Amyloid

Amyloid fibrils are at present associated with an more than 30 neurodegenerative or pathological human diseases. The range encompasses conditions like Alzheimer's, Huntington's, Transmissible Spongiform Encephalopathy, Parkinson's, Type II diabetes, Dialysis related amyloidosis (DRA) and Senile systemic amyloidosis (7; 8). These are all linked to the mis-folding and self aggregation of normally functional proteins which, for the previously listed diseases, are Amyloid-Beta, Huntingtin, Prion protein (PrP), α-synuclein, Amylin, β-2-microglobulin and Transthyretin respectively. Amyloid material has been found in both

extra and intra cellular deposits in organs such the brain of patients with Alzheimer's diseases, joints of DRA patients and in the pancreas of patients with type II diabetes (7). A recently published review on human diseases offers a detailed and current understating of the role played by amyloid fibrils in neurodegenerative or pathological conditions (13).

Although the link between the presence of amyloid deposits and the diseased state has been made, there isn't at present any consensus on the identity of the toxic species. While some studies strongly suggest that the oligomeric structures arising during the early stages of fibril formation are responsible for toxicity (9; 25; 26), other studies indicate that the mature fibrils are also potent toxic agents as well (27; 28; 29; 30). Furthermore, the mechanism by which the fibrils or the on-pathway oligomers induce cytotoxicity are not yet understood either. Possible mechanisms by which the toxic species may act is to bind to the cell membranes and either permeabilize them (31) or form channels that can affect for example, the homeostasis of ions like Ca^{2+} (27).

In recent years, a link has been made between the decrease in length of fibrils, caused by fragmentation, and an increase in their toxic effects (28; 32). Fractured fibrils of reduced length seem more readily able to disrupt model liposome membranes and a selection of cell lines than their longer counterparts (33) thus implying the shorter fibril fragments are more cytotoxic (for β-2-microglobulin, α-synuclein & PrP) . This feature however, doesn't seem to be generic character in all fibrils. A study of two conformationally distinct polymorphs of the Syrian hamster prion protein for instance, reveals that fibril fragments were indeed more toxic for one polymorphic state but that fragments of the other state were substantially less toxic than the intact fibrils (28). These results highlight that in order for effective therapeutic measures against amyloid related diseases to be developed, serious consideration has to also be given to their mechanical characteristics.

1.3.2 The Good: Functional Amyloid

The self-assembly, mechanical stability and high resistance to degradation of amyloid fibrils may combine to frustrate attempts at biological clearance or therapeutic treatments in the disease state. These very characteristics however, make them ideal structures for exploitation by nature in a functional capacity. A number of organisms have been found that do indeed take advantage of the amyloid motif to form so called 'functional' amyloid. The

formation of these functional fibrils is always a highly regulated process geared towards minimising potential risk of toxicity and harm to the organism. Examples of known functional amyloid found in diverse organisms can be found in the relevant literature (34), here only a summary is presented.

Bacteria such as *E. coli* utilise amyloid Curli protein fibres as the main constituent of extracellular bio-films (35). The bio-films are essential to the survival of bacterial communities as they provide a protective coating that is highly resistant to the action of antimicrobial agents or thermal degradation. Fibril production is highly regulated in that the aggregating monomers (CgsA subunit) in the Curli protein only self-assemble when bound, extracellularly, to a protein (CgsB subunit) anchored in the bacterial membrane (35). This avoids exposing the bacteria to harmful effects of fibril formation as initiation, production and elongation can only occur outside the organism. These Curli fibrils also aid in the surface adhesion and binding activities of the bacteria (35).

Amyloid and amyloid-like building blocks are also found in the web and silk spun by spiders (36). Spider silk is renowned for its mechanical properties which are directly derived from the presence and structural organisation of amyloid-like fibrils (fig 1.6). The mechanisms by which spiders are able to produce and store soluble proteins before conversion or secretion of insoluble silks are carried out in highly specialised glands (36). Such mechanisms, though not fully understood, may be optimised to protect spiders from disastrous effects of amyloid. Other organisms that exploit the mechanical characteristics provided by the amyloid state include fish, silkmoths and various insects. They use amyloid fibrils formed from chorion proteins as a protective material in egg envelopes (37; 38). The fibril material in the egg envelops protect the embryos/larvae of these organisms from undesirable external chemical and environmental conditions such as heat, degradation by proteases and mechanical stresses.

The occurrence of amyloid in humans has traditionally been linked in one way or the other with disease. In recent years however, discovery of functional amyloid vital for melanin biosynthesis in humans has been made (40). Amyloid fibrils formed from a fragment of the Pmel17 protein are used as templates for the orientation and polymerisation of melanin precursor monomers. The synthesis takes place in specialised organelles of the cell called

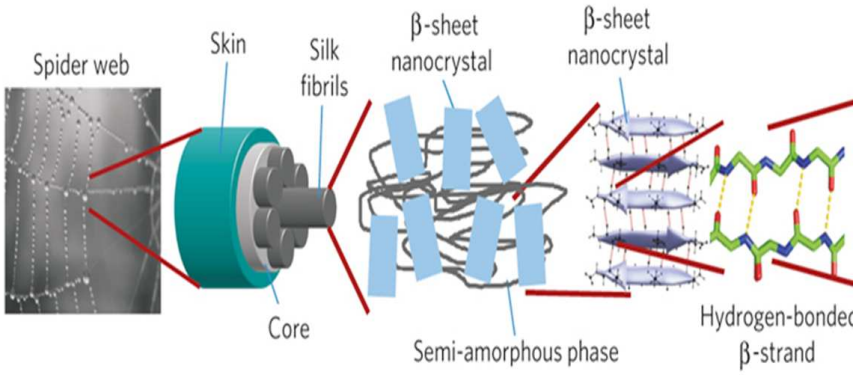

Figure 1.6: Hierarchical structural organisation in spider silk. The presence of amyloid-like assemblies (schematically shown as rectangles) are responsible for the high tensile strength. (Adapted from (39))

melanosomes. Interestingly, other fibrils from proteins such as Aβ & α-synuclein have been shown (*in vitro*) to be able to act as templates for melanin polymerisation (40) thus implying that Pmel17 fibril structure alone is required for its function. However, comparisons of the rates of fibrillisation show that the aggregation of Pmel17 is at least 4 orders of magnitude faster than rates found in Aβ or α-synuclein amyloid formation (40). This extremely rapid Pmel17 fibrillisation rate is thought have been selected through evolution for melanin biosynthesis as it avoids the formation or accumulation of intermediate toxic oligomeric states. The extremely rapid fibrillisation rate and use of specialised organelles is an example of the extent to which functional amyloid production is highly regulated.

1.3.3 The Useful: Amyloid Derived Nano-materials

The intrinsic properties of amyloid structures also make them very attractive candidates for the building blocks of novel nano-materials (41). Their *in vitro* self-assembly makes them particularly suited for the bottom-up construction of nano-devices, especially given that this ability is available to any peptide sequence. In addition, peptides are highly functionisable and fibrils adopt a range of morphologies that can be modulated by minor changes in pH, salt concentration and temperature. These characteristics coupled with mechanical robustness and resistance to degradation make a compelling case for utilisation of amyloid in nano-technological applications. An abundant wealth of information about these assemblies already exists in the literature as they have been extensively studied due to their pathogenic links. To date, a number of potential nano-technological applications have been

proposed with some successful proof of principle demonstrations already available (41; 42).

The most successful demonstration of the utility of amyloid based nano-material design has been the construction of electrically conducting nano-wires (43). Peptides formed from the amyloidogenic Sup35p protein were modified to include surface-accessible cysteine residues. The peptides were then functionalised thorough covalent links between the cysteine residues and colloidal gold particles. The resulting self-assembled fibrils were able to conduct current with low electrical resistance and displayed ohmic behaviour. Interestingly, when the voltage was ramped up above a critical value, the current flow stopped completely as the fibrils became vapourised, thus acting as a nano-scale fuse (43).

Amyloid based materials have also been proposed as potential drug delivery devices. Two schemes that have been demonstrated include functionalised fibrils that allow the controlled slow release of peptide monomeric drugs (44) and an amyloid matrix material in which drug molecules are enclosed until they reach a target site (45). Other implementations have included using amyloid as a material in construction of structural & tissue engineering scaffolds (46; 47), hydro-gels (48) and thin films (49).

All these schemes (drug delivery, nano-wires etc) rely in some part, on the ability to control the mechanical strength of the fibrils in distinct ways to function effectively. The nano-wires for example, have to be as long as possible but not liable to fragmentation whilst the drug delivery devices only have to be mechanically robust until they reach the target site whereupon they have to be brittle enough to release their cargo. A thorough understanding of ways to modulate the mechanical character of the chosen fibril system is therefore needed in the development of each of these applications. The next section of this chapter is concerned with the current methods by which the mechanical properties of fibrils are assessed.

1.4 Mechanical Characterisation of Amyloid

Amyloid fibrils are mechanically stiff structures with remarkable elastic properties (50). Insulin fibrils for instance, have a strength of ≈ 0.6 GPa which is comparable to the value for steel and a stiffness (Young's modulus) of ≈ 3.3 GPa which is comparable to that of silk

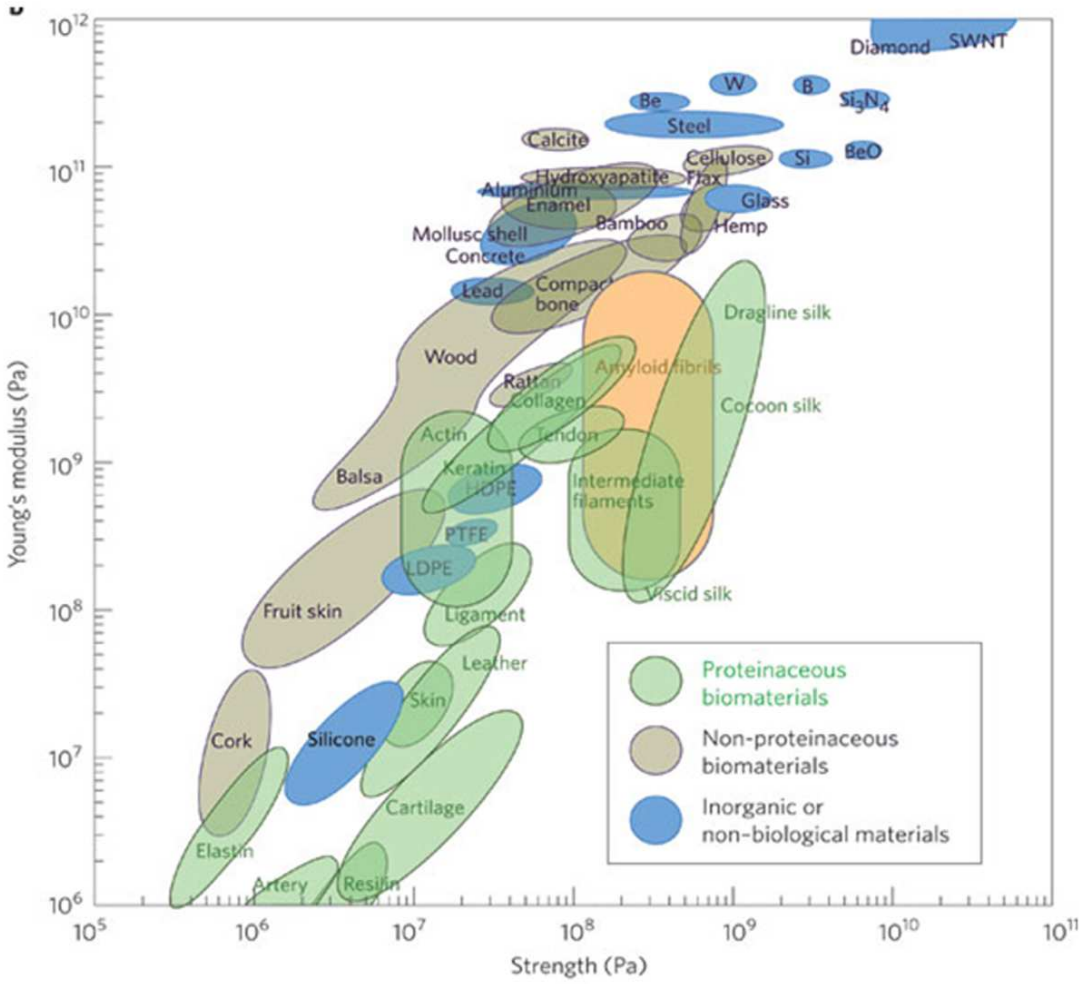

Figure 1.7: Comparisons of Young's modulus (Stiffness) as function of Strength for various materials. Amyloid fibrils are represented in orange, other proteinaceous biomaterials in green, non-proteinaceous biomaterials in brown and inorganic materials in blue. (Adapted from (50)).

(51). They are amongst the most mechanically stable biological structures made solely of a proteinaceous composition (fig 1.7). The major contribution to their mechanical character stems from individually weak backbone hydrogen-bond interactions that form a dense network which then acts cooperatively during mechanical resistance. Measurements of amyloid elastic properties have been made from a combination of biophysical and computational methods.

1.4.1 Biophysical mechanical measurements

Despite the obvious biomedical importance of the mechanical properties of amyloid fibrils, only few biophysical techniques are currently able to characterise their responses to force experimentally at the molecular level. This is due in part to practical limitations on how the fibrils can be prepared and tethered while the deformation is applied. Atomic force microscopy (AFM) is a means by which fibrils are imaged and primarily force probed. AFM has been used to mechanically probe an adhesive secreted by a marine invertebrate (Entobdella soleae) (52; 53) that contains mechanically functional amyloid.

The AFM technique has also been used to measure the elastic modulus of a number of amyloid and amyloid-like fibril species through force spectroscopy and nano-indentation (51; 54; 55; 56; 57) experiments that record the local response of a selected region rather than that of the fibril as a whole. Analysis of AFM topographic data taken from images of a number of fibril species, to obtain their bending rigidity, reveals that fibrils have a high elastic modulus whose molecular basis stems from backbone intermolecular hydrogen bonding (58). Other methods include the use of synchrotron x-ray diffraction at high pressures to study fibrils under mechanical compression (59) which reproduce elastic moduli comparable those in AFM experiments.

1.4.2 Computational mechanical measurements

Computer simulations, in which the geometry of the system being studied and the manner in which force is applied can be precisely controlled, are particularly valuable in complementing experimental studies. Steered molecular dynamics simulations have been used to probe the force-induced unbinding of a single peptide from Aβ fibrils and demonstrated an anisotropic response relative to the pulling direction (60). Another set of simulations were

Figure 1.8: Human and rat versions of amylin have very similar amino-acid sequences. The places where the sequences differ (marked in red) are believed to confer amyloidogenic properties in the human version not present in the rat analogue.

able to measure the Young's modulus for Aβ fibrils through a combination of compressive and tensile loading conditions (61; 62; 63). Simulations replicating AFM experiments can also characterise the elastic properties of proteins such as titin (64; 65). They reproduce force-extension profiles similar to those found in AFM measurements. However, such simulations tend to report peak forces many orders of magnitude higher than in the experiments due to the time steps and pulling rates used. They do nonetheless, provide dynamical information not accessible to the experiments about the relevant interactions and structural reorganisations in the fibrils due to the applied force. In the work described in later chapters of this thesis, the steered molecular dynamics simulation method is the main probing mode applied to characterise mechanical aspects of different amyloid systems. A brief introduction and background to the proteins from which each studied amyloid system arises is therefore provided in the next section.

1.5 Meet the Protein models

In the work presented in this thesis, two distinct amyloid systems are modelled in atomistic detail and then mechanically probed through the use of computational methods. The model systems are from the amyloidogenic proteins, amylin and amyloid-beta, which are associated with type II diabetes and Alzheimer's disease respectively.

Amylin: Islet amyloid polypeptide

Amyloid deposits are found in the pancreas of $\approx$ 95% of patients with type II diabetes (66). The main constituent of the fibrillar deposits is the normally soluble and physiologically functional amylin protein. Amylin, also known as the islet amyloid polypeptide (IAPP), is a 37 amino-acid residue protein believed to play a role in the regulation of glucose as it

is co-secreted with insulin (31). The amino-acid sequence of amylin is highly conserved across several species with only slight variations from organism to organism (66). An interesting observation is that despite similarity of sequence between human and rodent analogues (fig 1.8), rats do not develop type II diabetes nor do they show signs of fibril deposits in their pancreas (31). This led to speculation that the regions where these two sequences differ could be responsible for the amyloidogenic nature of human version. This is confirmed by studies on amylin fragments utilising peptides where rat and human amylin differ (region 20 to 29 in the sequnce). The 10-residue human version, hIAPP$_{20-29}$, readily forms fibrils whilst the the rat analogue, rIAPP$_{20-29}$, is unable to form stable fibrils (67). The hIAPP$_{20-29}$ fragment of amylin (**SNNFGAILSS**) is a convenient archetypal model used in both experimental and computational studies of amyloid related phenomena due to its relative small size and amyloidogenicity.

Amyloid Beta: Aβ protein

One of the hallmarks of the neurodegenerative age related Alzheimer's disease is the presence of amyloid deposits in the brain of affected patients (7). The deposits consist mainly of aggregated Amyloid-Beta (Aβ) proteins whose physiological function is as yet undetermined. Aβ is derived from a much longer transmembrane protein, β-amyloid Precursor Protein (APP), which is also of unknown physiological function. Proteolytic cleavage of APP leads to a much shorter 39 to 43 residue Aβ protein, whose exact length depends on the location of β- and γ-secretase cleavage sites (fig 1.9). This results in a heterogeneous mix of different length Aβ peptides, of which the 40 and 42 residue peptides (Aβ_{40} & Aβ_{42}) are the dominant species (fig 1.10).

Aβ_{42} has a much higher aggregation ability and neurotoxicity relative to Aβ_{40} (68). Although both species have been extensively studied, only fibrils formed by Aβ_{40} have so far been structurally determined. A combination of ssNMR and other biophysical techniques have led to two distinct models for the fibrils (69; 70). Each model is made up of U-shaped Aβ_{40} monomers which stack to form either a 2- or 3-fold symmetric structure. The 2-fold symmetric model is thought to represent a proto-filament while 3-fold symmetric structure is that of a mature fibril (70). A schematic representation of the 3-fold symmetric model as determined from ssNMR constraints is shown in figure 1.11.

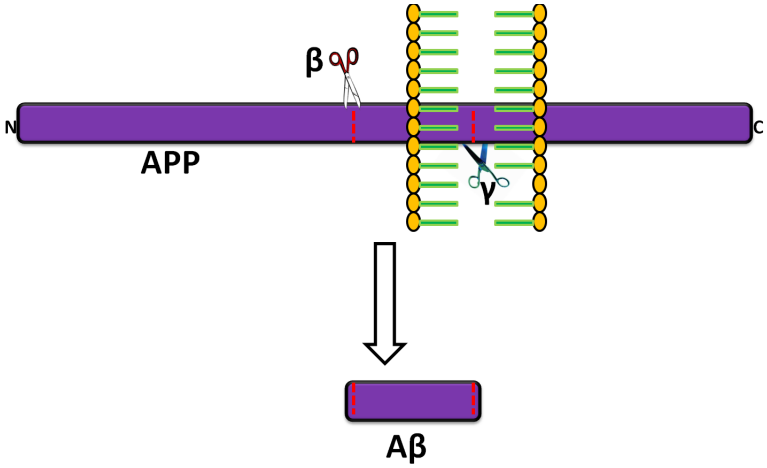

Figure 1.9: A 365 residue long transmembrane β-amyloid Precursor Protein (purple) is proteolytically cleaved by action of β- and γ-secretase (along dotted red line) to form the shorter Aβ peptide.

1 10 20 30 40

$A\beta_{(40)}$: DAEFRHDSGY EVHHQKLVFF AEDVGSNKGA IIGLMVGGVV

$A\beta_{(42)}$: DAEFRHDSGY EVHHQKLVFF AEDVGSNKGA IIGLMVGGVV IA

Figure 1.10: The Aβ_{40} and Aβ_{42} protein sequences both associated with Alzheimer's disease differ only in two additional isoleucine and alanine residues at the C- terminus of Aβ_{42} .

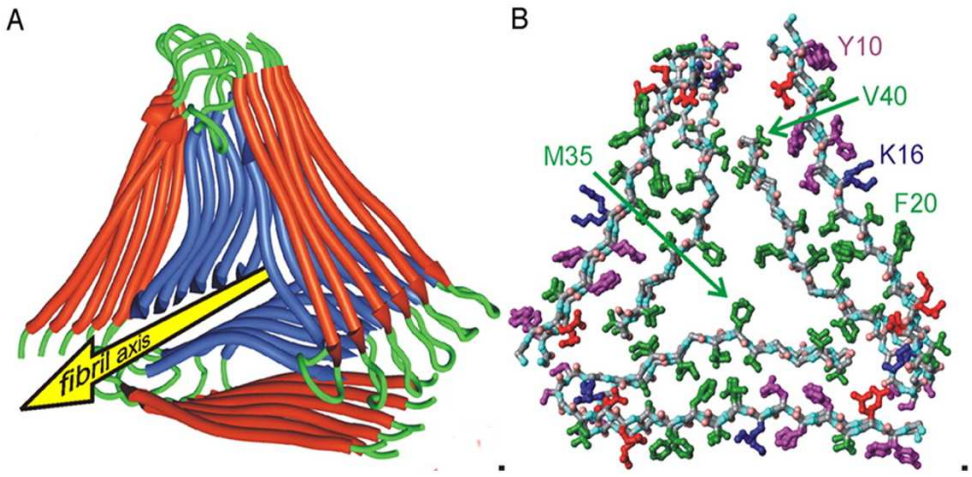

Figure 1.11: A) The 3-fold symmetric Aβ_{40} fibril model obtained from ssNMR data. B) Three U-shaped Aβ_{40} monomers of a single layer of the fibril. (reproduced from (70)). The 3-fold fibril model is exclusively used in the relevant chapters of this work.

In early on-set Alzheimer's disease, the characteristic symptoms become evident before the age of 65, with a rapid decline in the affected individual. Early on-set Alzheimer's disease has been linked to mutations in the gene that expresses APP (71). These mutations then lead to single amino-acid substitutions in Aβ that tend to cluster around residue 21 to 23. Each Aβ variant peptide displays enhanced aggregation and neurotoxicity and is associated with the Familial Alzheimer's Diseases (FAD) i.e early on-set Alzheimer's disease (72). Examples of FAD related Aβ variants include Flemish (A21G), Arctic (E22G), Dutch (E22Q), Italian (E22K) and Iowa (D23N) mutations.

1.6 Aims of This Work

The main aims of the work presented in this thesis are to understand, at a molecular level, how the intrinsic interactions within amyloid fibril structures influence their mechanical responses to externally applied forces. The motivation behind this is to gain further insights into the process of fibril fragmentation which, as already discussed here, has direct implications for fibril proliferation, toxicity and potential utility. The computational steered molecular dynamics (SMD) simulation technique is primarily used to probe and disrupt the stabilising interactions in a variety of fibril structures, which are modelled in atomistic detail. The theoretical basis for the simulation methods and details of their implementation on the fibril models are outlined in chapter 2. Three distinct investigations are undertaken in pursuit of the above stated aims. The details and outcome of each are presented separately in chapters 3, 4 and 5.

The investigation outlined in chapter 3 is concerned with the role played by the amino-acid side-chains in determining the mechanical character of amyloid fibrils. This study uses a number of sequence modified fibril models of the SNNFGAILSS fragment of the amylin protein implicated in type II diabetes. The focus is on how the structural order imposed by the amino-acid residue substitutions on the β-sheet assemblies impacts on both the thermodynamic and mechanical stability of fibrils.

The study in chapter 4 examines how the different steric packing arrangements found in fibril polymorphs can be manipulated as a means to modulate their mechanical properties. Three polymorphic structures of amylin SNNFGAILSS derived fibrils, with various

modifications, are compared and contrasted during the different force probing modes.

The final investigation in chapter 5 is partly inspired by the observations made in the preceding studies (chapters 3 & 4). It focuses on the role of structural defects that may arise due to hetrologous cross-seeding events in fibril formation. This work uses the larger Aβ fibril models implicated in Alzheimer's disease, with cross-seeding modelled between the wild-type Aβ and selected Familial Alzheimer's Diseases (FAD) variants.
The major findings of all three investigations are summarised in chapter 6 along with a discussion of possible avenues of future research opened up by this work.

Chapter 2

Theory and Methodology

Synopsis

This chapter presents the theoretical foundations, strengths and limitations of the simulation techniques employed to achieve the stated aims of this work. An outline of the methods developed to construct, refine and mechanically probe the various fibril models are discussed. Finally, the analysis methods and tools used through the work presented in chapters 3, 4 and 5 are given.

2.1 Molecular Dynamics

Overview

Molecular dynamics (MD) simulation is a computational technique originally developed to study condensed phase matter in 1957 using hard sphere models. It wasn't until 1977 that the first successful MD simulation of a bio-macromolecule was reported (73) on the folded globular protein bovine pancreatic trypsin inhibitor (BPTI). Since then a number of theoretical and technological advances have seen MD grow from a complementary tool for established experimental biophysical techniques to arguably a field in its own right. This view is particularly highlighted in a recent review of biomolecular simulation focusing on the progress made in the past 35 years (74) in areas such as structure prediction, protein folding & interpretation of experimental phenomena from a host of biomolecules. Despite its wide successes, the theoretical principles underpinning MD are relatively simple. Given an initial structure (atomic positions) of a molecule, molecular dynamics seeks to evolve the system until new atomic positions can be determined by simply considering their energetic interactions. By doing this iteratively over a given period, a trajectory of the time evolution of the system is built up from which a wealth of information on the molecule can be obtained which would not have been accessible from just the static model.

Below is a brief description of the theoretical basis and means of implementation of the MD technique. The section heavily references three excellent resources, books by Leach (75), Schlick (76) and Mccammon (77), which detail the theoretical formulation of the MD method as applied in biomolecular simulation.

2.1.1 Theoretical Formalism

The central physical principle that governs MD simulation is Newtons second law of motion eqn (2.1) which relates the force, **F**, experienced by mass, m, in motion to its rate of change of momentum. For a system with a known starting configuration i.e atomic positions, **r**, it is possible to numerically integrate the equation of motion iteratively such that new snapshots of the next atomic positions are generated at each time step.

$$\mathbf{F} = m\frac{d^2\mathbf{r}}{dt^2} = m\mathbf{a} \tag{2.1}$$

Typically, positions of atoms in bio-molecules are obtained from experimental techniques such as X-ray crystallography or ssNMR. In the absence of experimental structures it is possible to rationally design and construct computational models of the bio-molecules of interest.

In order to make use of eqn (2.1) to evolve a molecular model, the forces on the individual constituent atoms can be computed from considering their energetic interactions with their neighbours. A classical empirical function is used to account for the possible inter-atomic interactions in terms of their relative positions. This potential energy function, $U(\mathbf{r})_{\text{total}}$, is also known as the forcefield and is made up of a summation of energy terms. The forces on each of the atoms due to interactions with their neighbours are thus obtained from the gradient of the forcefield, eqn (2.2).

$$\mathbf{F} = -\nabla U(\mathbf{r})_{\text{total}} \tag{2.2}$$

The molecular mechanics forcefield:

Since the potential $U(\mathbf{r})_{\text{total}}$ is an empirically derived function, it is not surprising that a diversity of biomolecular forcefields exist. These may have different terms in their functional forms, or maybe similar but posses different sets of parameters. Examples of popular forcefields used to MD simulation on macromolecules are AMBER, CHARMM, GROMOS, MM2/3/4 and OPLS. A common theme embodied in the forcefields is that potential energy function consist of a summation of bonded and non-bonded interaction energy terms. A

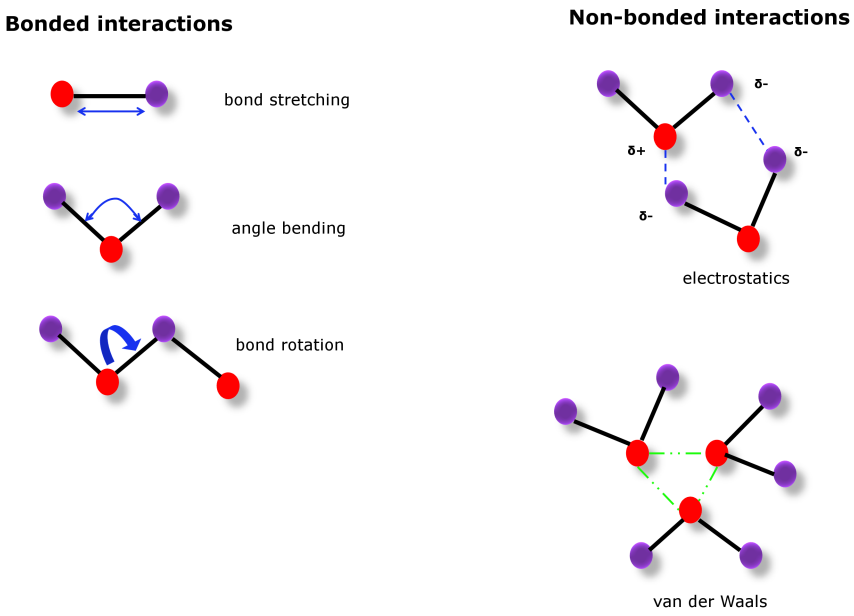

Figure 2.1: Schematic representation of the interaction energy terms that are found in a typical molecular dynamics forcefield. The bonded terms relate to bond stretching, angle bending and bond rotation. Non-bonded terms are electrostatic and van der Waal interactions.

generic form of the energy function is given in eqn (2.3).

$$U(\mathbf{r})_{\text{total}} = \sum_{\text{bonds}} K_r(r - r_{eq})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{eq})^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}} \right] \tag{2.3}$$

Figure 2.1 shows the origins of each term in eqn (2.3) and these are discussed in turn. The first three terms in eqn (2.3) relate to the bonded interactions between sets of covalently linked atoms in the molecule. The first term (summation over bonds) describes the stretching of a bond between two atoms about an reference distance r_{eq} and is represented by a harmonic term with the strength of the bond defined by a spring constant K_r. Similarly, the second term (summation over angles) relates to bond bending between three atoms about an equilibrium bond angle θ_{eq} whose strength is determined by a bending spring constant K_θ.

The use of harmonic potentials to describe bond stretching and bending is only valid for small fluctations about the equilibrium positions. For larger fluctuations however, the harmonic expressions of these interactions does not capture bond breakage or formation events that may occur in typical chemical reactions. Behaviour at large deformations like bond disassociation during bond stretching for instance are more accurately described by the Morse potential, eqn (2.4).

$$U(\mathbf{r})_{\text{bond}} = D_e(1 - e^{-a(r-r_{eq})})^2 \tag{2.4}$$

For two bonded atoms separated by a distance r, the bonded potential, $U(\mathbf{r})_{\text{bond}}$ between them will depend on the depth the well of the potential (disassociation energy), D_e, a reference bond length r_{eq} and a parameter a that describes how quickly the potential changes with increasing separation. The Morse potential is however rarely ever implemented in current forcefields as it is not only much more computationally intensive than the harmonic approximation but also needs to be parametrised for three terms for each bond.

The final bonded energy term in the forcefield, eqn (2.3), is a summation over the dihedral angles. This represents the torsional energy between a group of four atoms separated by three covalent bonds. In this periodic function, the parameters V_n, n and γ are the torsional force constant, periodicity and phase angle respectively. The summation however only accounts for atoms consecutively linked together in the sequence 1-2-3-4 i.e proper dihedral. As such, an additional term to account for improper out-of-plane torsional angles may also be included in some forcefields.

Non-bonded pairwise interactions between atoms i and j separated an inter-atomic distance R_{ij} that account for both the van der Waal and electrostatic energy in the forcefield (eqn (2.3)) are decribed as:

$$U(\mathbf{r})_{\text{non-bonded}} = \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}} \right] \tag{2.5}$$

The evaluation of this term is a major cause in computational bottlenecks when large numbers of atoms are involved. As such, a number of methods and approximations are used to make the calculations tractable. Van der waal interactions are represented by the 6-12 Lennard-Jones potential with the repulsive and attractive coefficients A_{ij} and B_{ij} determined by the type of interacting atoms. The 6-12 Lennard-Jones potential is designed to

reflect that atom pairs beyond an optimal distance move close to each other due to mutual attraction but then then strongly repel when their separation is less than this distance. The attractive R_{ij}^{-6} relation is predicted by the Drude dispersive interaction model whereas the repulsive R_{ij}^{-12} component is used for computational convenience. As van der Waal interactions tend to be short ranged, distance cutoffs or switching functions are normally employed during MD simulations to reduce computational costs.

Meanwhile, the electrostatic interactions are a computed from the Coulombic expression where ϵ is the effective dielectric constant for a medium between a pair of atoms of charge q_i and q_j. As this is a R_{ij}^{-1} relation, the interactions are long ranged and decay very slowly. This means that the imposition of distance cutoffs to improve computational efficiency is not possible. Fortunately methods for evaluating these interactions for a system of made of N atoms have been developed that speed-up the calculation from a prohibitive $\mathcal{O}N^2$ to a reasonable $\mathcal{O}NlogN$ problem. One such example is the particle-mesh Ewald method (PME) which is used for systems that employ periodic boundary conditions. The main effect of the PME method is split the electrostatic evaluations into summations of components in real and reciprocal space which both rapidly converge. The speed-up in the reciprocal space summation in particular is aided by the use of the fast Fourier transform which requires each atomic point charge to be distributed over a lattice grid. In most forcefields, hydrogen bond interactions are implicitly contained in the non-bonded potential eqn (2.5), however in some cases additional explicit terms are also used to describe them.

The forcefield equation (2.3) has a number of parameters and constants for the various interactions of different types of atom species. These parameters are derived from a combination of quantum calculations on very small molecules and experimental methods such as vibrational spectroscopy (infrared radiation and Raman scattering). Once an appropriate forcefield has been selected for a particular starting structure of a macromolecule, the next step is to understand how this is transformed into a trajectory by the MD simulation process.

Integrating the equation of motion:

Given that a continuous forcefield function is used to describe the inter-atomic interactions, it is not possible analytically solve for the forces exerted on the atoms as their interactions

are coupled together. As such, Newtons equation of motion has to be solved through numerical integration over a very short time-step δt. It is the integration of the equation of motion that gives a trajectory which describes the how the positions ($\mathbf{r}$), velocities ($\mathbf{v}$) and accelerations ($\mathbf{a}$) of atoms vary as a function of time. Central to this is the assumption that positions, velocities and accelerations can be approximated by a Taylor series expansion. A number of integration algorithms have thus been developed to solve for atomic positions during MD simulation. The most commonly used integrator is the Verlet algorithm eqn (2.6)

$$\mathbf{r}(t+\delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\delta t) + \mathbf{a}(t)\delta t^2 \tag{2.6}$$

which determines new atomic positions at time $t+\delta t$ from considering the current atomic positions $\mathbf{r}(t)$, their current accelerations $\mathbf{a}(t)$, their position at a previous time interval $\mathbf{r}(t-\delta t)$. The accelerations $\mathbf{a}(t)$ are obtained from the gradient of the potential through eqn (2.7).

$$\mathbf{a}(t) = -\frac{1}{m}\nabla U(\mathbf{r})_{\text{total}} \tag{2.7}$$

The first stage before an actual MD simulation can be carried out however is to specify the starting atomic positions, the forces on the atoms due to their interactions and their initial velocities at time t. The initial atomic positions are expressed in structure files (e.g PDB) as cartesian coordinates, initial atomic velocities are then randomly assigned using the Maxwell-Boltzmann distribution set at the desired simulation temperature whilst forces are determined from the gradient of the forcefield. The next stage is to compute the atomic positions at a later time step $t+\delta t$. When these have been determined, the forces at the new positions are once again calculated from the potential function thus allowing for the atomic positions and velocities at time $t+2\delta t$ to be found. This procedure is repeated over a multiple steps through which a trajectory of the system is slowly built up in time.

For an accurate and stable representation of all atomic motions in a molecule, the MD integration time-step δt has to be sufficiently small to account for the highest vibrational frequencies of the bonded atoms. In biomolecules such as proteins and DNA, bond stretching involving links to hydrogen atoms tend to give rise to the largest vibrational frequencies. In order to capture these motions, a simulation time-step δt of 1 femto-sec has to be used. This time-step is however computationally inefficient for covering phase space for systems with large numbers of atoms. Fortunately though, the high frequency vibrational modes of hydrogen atoms do not contribute significantly to the bulk motions of the molecule.

As such, these can be fixed throughout the entirety of the simulation by the application of constraints. Popular algorithms for constrained dynamics include SHAKE and RATTLE methods which are incorporated into the integration algorithm and evaluated for the specified atomic bonds. These allow a time-steps of 2 femto-sec to be used for biomolecular systems resulting in simulation timescales of tens of nanoseconds to be achieved by atomistic MD.

Despite the documented successes in helping understand a host biomolecular interactions and properties, atomistic MD simulations have a major drawback in that most biologically relevant phenomena occur in time-scales that range from micro-seconds to milliseconds and longer. This means that most of the interesting biological changes such as protein folding or ligand unbinding are not accessible to to technique as it is limited to nanosecond scale. One way of addressing this is to move from a full atomistic description of the system to a coarse grained representation. Coarse grained MD has allowed for simulation times in the μs range to be reached but this is at the cost of the loss in detailed information during the dynamics. In recent years, advanced MD simulation methods have been developed in order to either improve on phase space sampling or to accelerate physical processes. These methods include Replica Exchange MD, umbrella sampling, Steered MD and Targeted MD.

Of particular interest is Steered molecular dynamics (SMD) method which applies external forces to the system such that events that cannot be normally accessed in MD simulation time-scales (e.g conformational changes) can be accelerated and observed (64). Crucially, the atomistic details of the system are retained in this method. The SMD technique (in addition to atomistic MD) has been extensively used in the work described in chapters 3, 4 and 5 of this thesis. This has been primarily to induce amyloid fibril fragmentation and fracture through the application of external forces. As such, a detailed description of the SMD technique is presented in the following section.

2.2 Steered Molecular Dynamics

Steered molecular dynamics is an advanced MD method that has been widely & successfully implemented in studies of the elastic properties of proteins, protein unfolding, ligand binding and movement of ions through channels in membrane pores. It is a technique that has been primarily useful in the characterisation of the mechanical properties of bio-molecules.

Trajectories from a SMD simulation can be used to produce force-extension profiles as well as details of the work done by the steering force. The Steered molecular dynamics method is inspired by biophysical force probe techniques such as atomic force microscopy (AFM) and optical tweezers. SMD not only complements these types of experiments but also gives atomistic details on the probed molecule that is not yet accessible to the established physical methods. In a typical SMD simulation, an eternal force is applied to an atom (or group of atoms) in the molecule whilst one end of the molecule is held fixed (78). This allows for the mechanical response of the molecule along a chosen pulling direction (reaction coordinate) to be monitored during the simulation. The external steering force is applied in addition to the forcefield that describes the energetic interactions between atoms in the simulation.

2.2.1 Types of steering forces

There are a number of different ways in which the applied SMD force can be used to steer the biomolecule under investigation. Currently, one can apply either a constant force to the system during the dynamics or alternatively a variable time-dependent force maybe used. Although these two force schemes are fundamentally different in their execution, they do however provide complementary information in their force-time profiles about the mechanical character of the molecule. In a constant force SMD simulation, a fixed linear force is applied directly to an atom (or group of atoms) whilst the extension/displacement is recorded. Constant force SMD simulation trajectories are particularly amenable to analysis by the theory of mean first passage time which characterises the unfolding barriers. This approach allows for short lived intermediate protein states during unfolding to be identified. Additionally, rotational motion in regions of a biomolecule can be induced through application of a constant torque.

In the variable force scheme, an atom (or group of atoms) is linked by restraints to a point in space through a harmonic potential. The point in space is then moved at a constant velocity in a predefined direction thus forcing the linked atom and the rest of the molecule to respond. The constant velocity i.e variable force SMD scheme is the primary steering force mode used in the simulations carried out in chapters 3, 4 and 5 of this thesis. An in-depth description of this steering mode is therefore presented in the following sections.

2.2.2 Constant Velocity SMD mode

The constant velocity SMD (cv-SMD) mode is a computational analogue for AFM experiments in which the cantilever is attached to one end of a molecule and retracted at a constant rate while the force response is measured. A virtual spring is used in cv-SMD instead of the AFM cantilever and this is coupled to a 'dummy' atom that is pulled at a constant velocity along a set direction. The dummy atom is actually a restraint point in space linked to the pulled atom in the molecule via an external harmonic potential and this is shifted in the set direction. The functional form of the cv-SMD steering potential is expressed in eqn (2.8) and this is applied in addition to the inter-atomic forcefield.

$$U_{smd} = \frac{1}{2}k\left[vt - (\vec{r} - \vec{r}_0)\cdot\vec{n}\right]^2 \tag{2.8}$$

The terms in the potential, U_{smd}, are: stiffness of the virtual spring k, pulling velocity v, initial position of pulled atom $\vec{r}_0$, current position of pulled atom $\vec{r}$, unit vector pointing in the pulling direction $\vec{n}$ and current simulation time-step t. The force at each time step is then evaluated from negative gradient of the steering potential, $-\nabla U_{smd}$. A schematic representation of the implementation of the pulling mode is shown in fig 2.2. The values of parameters in U_{smd} heavily influence the results obtained in SMD simulations. As such, a detailed discussion relating to the choice of parameters in eqn (2.8) used for the simulations in later chapters is given in § 2.4.2.

It should be noted that although cv-SMD simulations reproduce similar types of force profiles to AFM experiments and offer much more atomistic details, there are however some computational limitations associated with this method. Firstly, the timescales in which AFM experiments are conducted in are at least six orders of magnitude ($\approx 10^6$) slower than what can be achieved in current simulation timescales. This means that extremely high pulling rates relative to AFM have to be used in cv-SMD to bring about the desired conformational changes. Hence the recorded peak forces are always much higher than the experimental values. Secondly, a direct consequence of using high pulling rates in cv-SMD is that irreversible work is done which may complicate the analysis of the trajectory if equilibrium properties are to be extracted. Fortunately though, it is now possible to extract 'exact' equilibrium information from a non-equilibrium process through the use of analytical methods such as those described in the next section. Despite the shortcomings

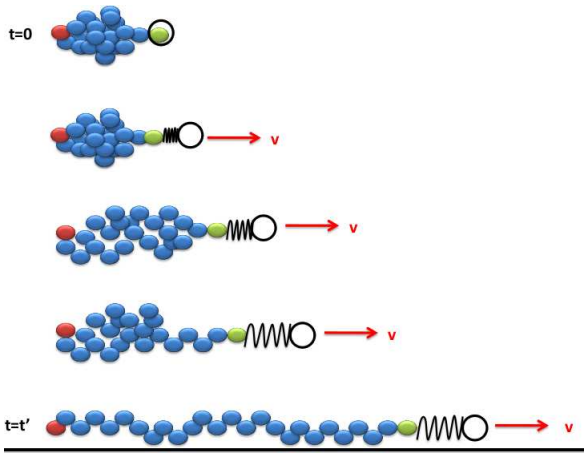

Figure 2.2: Schematic representation of constant velocity steered molecular dynamics. One end of the molecule is held fixed (red) and the other end pulled (green). The pulled end is attached to a dummy atom (black circle) by a harmonic spring which is pulled at a constant velocity.

and exaggerated force peaks, cv-SMD still offers a powerful qualitative way in which to mechanically compare molecules.

2.2.3 Non-equilibrium SMD analysis

A typical cv-SMD trajectory can give a wealth of qualitative information at the atomistic scale of key interactions that are destabilised or form during the steering process. This is supplemented by force-time profiles which may indicate the location of relevant intermediate states of the system. However, collecting quantitative thermodynamic information from a cv-SMD trajectory is not possible as the simulations are inevitably done out of equilibrium unless the pulling speed is sufficiently slow (limit of low force) enough to make the process reversible. This means that useful equilibrium thermodynamic descriptions such as free energy changes between the native and deformed states cannot be used for these processes. The second law of thermodynamics states that the average work, W, done on a system when moving it from one equilibrium state to another cannot be smaller than the change in free energy, ΔF, between the two states, with equality only being realised if the process is quasi-static.

$$\Delta F \leq \langle W \rangle \tag{2.9}$$

In 1997, Jarzynski derived an expression that allows for the calculation of free energy difference from a non-equilibrium process (79). The Jarzynski equality (JE) relates the

equilibrium free energy difference between two states to the non-equilibrium work required to move from one of those states to the other (79).

$$e^{-\beta\Delta F} = \langle e^{-\beta W} \rangle^{1} \tag{2.10}$$

This expression has been since subjected to a lot of theoretical and experimental scrutiny and so far seems valid. Crucially for constant velocity SMD, the JE holds regardless of the speed of the process and is thus a potent tool in calculating the free energy change from the cv-SMD trajectories. Just from an ensemble of trajectories and a suitable reaction coordinate, one can utilise the work done during the non-equilibrium process to compute the free energy landscape as a function of that chosen reaction coordinate. This is known as the potential of mean force (PMF) and its form is a good indicator of the stability of a system as you do work on it along a reaction coordinate path. The ability to produce the PMF allows one to determine the overall thermodynamics (stability) and kinetics (relaxation time) of the system.

There is a limitation however with the usefulness of the JE as the exponential term to be averaged is dominated by small work values that are rarely ever sampled during the simulation (80). This means that if a statistically insufficient number of trajectories is analysed and is lacking in these rare events, there is very poor convergence between the computed and the actual PMF. As a result, approximations have to be made in order to make the JE practically applicable. One such approximation is that of cumulant expansion of the natural logarithm of the averaged exponential term.

$$ln\langle e^{-\beta W} \rangle = \left[-\beta\langle W \rangle + \frac{\beta^2}{2}(\langle W^2 \rangle - \langle W \rangle^2) + ... \right] \tag{2.11}$$

This expression is truncated to a second order expansion, and provided the distribution of the work values is obtained in the simulations is Gaussian, it is sufficiently accurate to approximate the real PMF (81). It has been shown that the Gaussian distribution of the work values can be guaranteed by employing a *stiff* spring to do the pulling in cv-SMD (82) thus allowing for direct calculation of the PMF.

The number of cv-SMD simulation repeats carried out for the work described in chapters 3, 4 and 5 of this thesis are not statistically sufficient for accurate analysis by the Jarzynski

[1] Where β is $1/k_B T$.

equality method. As the prime aim of the work in those chapters is to characterise the relative mechanical resistance between the different fibril models, the peak force profiles and calculated stiffness constants suffice in this regard. It should be noted however that the JE was implicitly considered when deciding appropriate SMD parameters to use such as the stiffness of virtual SMD spring. The details of this are discussed later in § 2.4.2 .

2.3 Construction of fibril models

The biomolecules that are modelled, mechanically probed and discussed in later chapters of this thesis are amyloid fibrils formed from two distinct peptide sequences. Namely, these are proto-fibrils formed from the amyloidogenic fragment of amylin (hIAPP$_{20-29}$) and fibrils formed from the amyloid beta (Aβ_{1-40}) peptide, both of which are structurally determined from solid state NMR (ssNMR) experiments. The ssNMR coordinates are used as templates for construction of fibril models to which modifications such as fibril elongation and amino-acid composition are made as appropriate to each simulation investigation in chapters 3, 4 & 5. Below is a list of the general steps used to transform the ssNMR coordinates to viable computational models. The details unique to each of the two model structures are then described in turn.

General steps to build models:

1. Obtain experimentally determined atomic coordinates for the system in PDB format.
2. Use initial atomic coordinates as a template to elongate the fibril taking care to maintain the experimentally derived inter-peptide separation gap. This is done via the Nucleic Acid Builder (NAB) software package. If necessary, make appropriate changes to peptide sequence.
3. Check elongated model for any steric clashes, if present correct through energy minimisation methods.
4. Add missing hydrogen atoms to models since ssNMR is not able to resolve them. This is automatically done by use of the LEAP module of the AmberTools package.

5. If titratable residues are present in the amino-acid sequence, assign protonation states based on pH conditions the ssNMR models are grown. The empirically based fast pK_a predictor method (ProPka) is ideal for this.

6. Check the net charge of the model, if not neutral add appropriate number of counterions untill overall charge is zero by use of the LEAP module of the AmberTools package.

7. Adequately solvate the model in explicit water molecules and energy minimise the whole system.

2.3.1 Building hIAPP$_{20-29}$ protofibril models

The initial coordinates from the ssNMR data of fibril models from the **SNNFGAILSS** sequence consisted of a pair of β-sheets, each comprised of two peptides i.e a 2x2 unit cell. A combination of ssNMR and X-ray fibre diffraction techniques led to the identification two polymorphic steric zipper arrangements adopted by the fibrils, namely parallel and anti-parallel conformations (16). These were supplied as PDB coordinate files by David Middleton. Each 2x2 unit cell served as a template from which different sizes of fibrils could be built according to the steps outlined in the preceding section. The NAB software package was used to extend the length of the models by translating and joining copies of the 2x2 templates along the fibril axis, maintaining the 4.7 Å inter-peptide strand distance. Two fibril lengths were chosen for models studied in this work, namely a pair of β-sheets each made of 8 peptides (8x2 model) and a pair of β-sheets each made of 16 peptides (16x2 model). Additional modifications to the elongated fibril models were made in-line with the different investigations carried out in chapter 3 and chapter 4. These modifications relate to the amino-acid sequence & terminal capping groups respectively and are discussed in detail in the relevant chapters.

We were able to validate this method of fibril construction in a separate joint experimental & modeling investigation (83), independent of the work presented in this thesis, whereby cross-sectional area calculations from computational models of a variety of SNNFGAILSS derived fibrils accurately predicted cross-sectional area values consistent with those calculated directly from electro-spray ionisation mass spectrometry (ESI-IMS-MS) experiments (83).

2.3.2 Building Aβ_{1-40} and Aβ_{9-40} fibril models

For the Aβ_{1-40} fibrils, atomic coordinates derived from a combination of ssNMR and electron microscopy experiments were used (70). The experimental model is three-fold symmetric in that each layer in the fibre is made up of three monomers of u-shaped loops of Aβ_{1-40} monomers. Four low energy structures were provided by Robert Tycko as PDB files, each of which had six layers of the three-fold symmetric arrangement of monomers (six-mers). However, the first eight residues in the Aβ_{1-40} monomers were absent from the structure files as these show structural disorder thus cannot be resolved in the ssNMR experiments. As such, the bulk of the simulations in chapter 5 relate to fibril models built from the 32 residue Aβ_{9-40} monomers. For completeness however, a fibril model composed of the full length Aβ_{1-40} monomers was rationally designed by adding the eight residues in a random coiled configuration to the PDB files.

Elongation of the six-mer models through template translation was not possible since the turn regions in the monomer loops gave rise to severe steric clashes and misalignment when the interpeptide gap is maintained. To overcome this, a single layer of monomers was directly extracted from one of the low energy six-mer models. This then served a template for translations via the NAB software program and a six-mer model structure was constructed in this manner (fig 2.3). This and the low energy ssNMR structures were then each subjected to 45 ns of MD in explicit solvent after which comparisons were made. Visual inspection & analysis of the RMSD, enthalpy, secondary structure content and hydrogen bonding at the end of the dynamics showed that the designed six-mer model had similar structural characteristics to the four low energy ssNMR structures. As such, this procedure was used to build elongated fibrils of 12 layers of monomers (twelve-mers). Additional modifications to amino-acid sequence based on familial Alzheimer's disease mutations were then made to twelve-mers for various the cross-seeding investigations described in chapter 5.

Assignment of protonation states

The Aβ_{1-40} sequence, **DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV** is composed of a number of titratable side-chains (ASP, ARG GLU, LYS, HIS & TYR). The protonation states of these residues are strongly dependent on both the pH at which the fibrils are grown as well as the local environment of the residue within the protein.

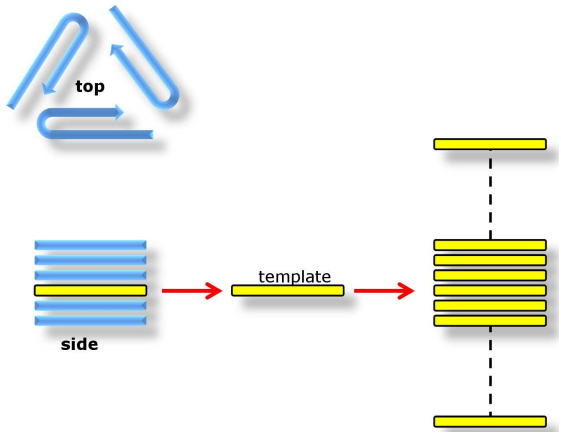

Figure 2.3: Steps in construction of elongated Aβ fibrils from low energy ssNMR six-mer models (blue, viewed from top and side). A layer of three monomers (yellow) is directly extracted and then used as a template for fibril elongation.

Additionally, the histidine residues are complicated in that at some pH values, they can be found in any one of three distinct protonation states (fig 2.4). As the ssNMR technique is unable to resolve hydrogen atoms, the coordinate PDB files were supplied without any hydrogen atom coordinates. Two methods were explored for the assignment of protonation states for the titratable residues. One uses Poisson-Boltzmann electrostatic based pk_a calculations, H^{++} via MEAD (84), whilst the other uses empirically based fast pK_a predictor algorithms, ProPka (85). Although the Poisson-Boltzmann pk_a based calculations are theoretically rigorous, they were however found to be computationally intensive for systems with large numbers of atoms such as the Aβ-12-mer models. Additionally, as these calculations depend on the atomic radii and their partial charges, using different forcefields (Amber99sb & Charmm-22-cmap) leads to very distinct results in assigned protonation states. Meanwhile assignments done using ProPka complete the prediction within a few seconds (85). Crucially, these only consider the hydrogen bonding geometry of residues relative to each other as well as the pH in assignment of protonation states.

Thus for computational expediency as well forcefield independence, the empirical predictions from ProPka were exclusively used for all models with titratable side-chains. ProPka calculations for the Aβ-12-mer models were evaluated using a pH of 7.5 to match buffer growth conditions in which the fibrils were grown. After protonation assignments, the mod-

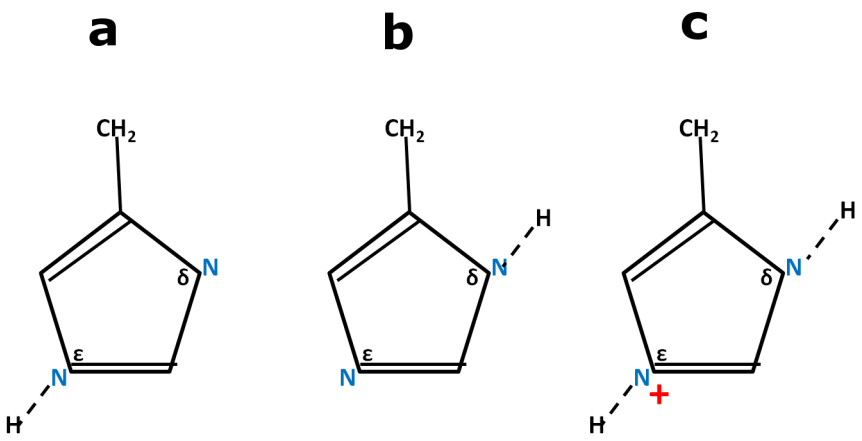

Figure 2.4: The three possible protonation states of histidine. a) and b) show singly protonated histidine with a hydrogen atom at the ϵ-nitrogen and δ-nitrogen sites respectively. c) Doubly protonated histidine with hydrogen attached at both δ-nitrogen and ϵ-nitrogen sites, this carries a net positive charge.

els possessed a net charge thus an appropriate number of ions were added to each system to make it charge neutral prior to solvation, minimisation and MD simulation.

2.4 Characterising mechanical failure in amyloid models

In order to characterise the mechanical properties of the various fibril models, steered molecular dynamics simulations were used to apply sufficient mechanical stresses to induce failure in the systems. In this instance, mechanical failure relates to the forced fragmentation of the fibril into two or more aggregates that are separated by intervening solvent molecules. In the normal implementation of the SMD technique for protein unfolding, one end of the molecule is typically pulled whilst the other end is held fixed as in figure 2.2. This approach has already proved successful in a recent ambitious survey of over 17000 folded protein structures (86). At present, the sample size is significantly much more than has been mechanically characterised through experimental biophysical techniques. Each protein is ranked according to the maximum force, F_{max}, required to induce rupture events when being pulled end-to-end (as in fig 2.2). This has allowed for the identification of several types of force clamps and offered new insights into the relationship between protein structure, mechanical stability and function (86; 87; 88). The computed F_{max} values and related properties for the $\approx$ 17000 proteins surveyed are stored in the Bio-molecule Stretch-

ing Database (BSDB) (89).

It should be noted however, that despite the large sample size in the BSDB survey, it currently only covers small protein structures made of up to $\approx$ 250 residues. Crucially, the selection criteria and the manner of the stretching protocol employed also excludes amyloid fibril structures from the analysis. This in part is because the standard protein unfolding end-to-end SMD pulling mode is inadequate for study of amyloid fibrils since their structure is a discontinuous collection of peptides that are held together by non-bonded interactions such as inter-peptide hydrogen bonding and inter-β-sheet interactions. As such, there was a need to create bespoke pulling geometries that could directly probe the various stabilising interactions of the fibrils as a whole from different directions. To this end, a total of four types of fragmentation schemes were designed to probe a host of fibril models in this work. The SMD fragmentation protocols were only applied at the conclusion of the standard production MD simulations once convergence of relevant RMSD values had been achieved. The following section describes the design and implementation of each of the four pulling geometries with emphasis on the types of stabilising interactions they disrupt.

2.4.1 Design of Fragmentation protocols

Each deformation protocol interrogates the physical response in a specific direction by either affecting the hydrophobic core interactions between a pair of β-sheets or the inter-peptide hydrogen bond network within each β-sheet. The four schemes are named Peel, Stretch, Slide & Shear and are schematically shown in fig 2.5.

In all cases, the implementation requires a selected group of backbone carbon-α atoms in the relevant peptides to be held fixed for the duration of the simulation whilst the centre of mass of another group of backbone carbon-α atoms is pulled with a variable force in a specified direction. All other atom types in the peptides are free to move during the simulations. The pulling directions **n1**, **n2** and **n3** in figure 2.5 are computed from a unit vector that lies to normal the plane of chosen atom groups and is located at their centre of mass. The four pulling geometries all utilise the constant velocity pulling SMD approach.

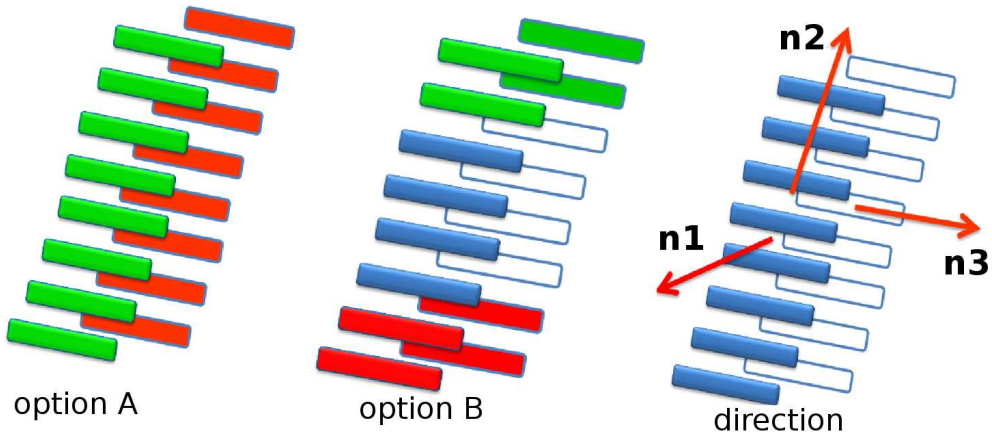

Figure 2.5: Illustration of the deformation schemes applied to fibril models. Each protocol is defined by pulled/fixed atom choice (either **A** or **B**) and then a pulling direction (**n1**, **n2** or **n3**). (Green) Peptides whose backbone carbon-α atoms are pulled. (Red) Peptides with carbon-α atoms held fixed. All other atoms in these peptides are free to move. **Peel deformation:** option A and n1. **Stretch deformation:** option B and n2. **Slide deformation:** option A and n2. **Shear deformation:** option B and n3.

Peel Fragmentation

The Peel deformation involves pulling the pair of β-sheets apart in order to separate them. This scheme primarily addresses the stabilising hydrophobic interactions between the pair of β-sheets. The peel SMD protocol requires all backbone carbon-α atoms in one of the sheets to be fixed for the duration of the simulation whilst the center of mass of all backbone carbon-α atoms in the other sheet is pulled with a variable force. The pulling direction is perpendicular to the peptide axis. This is represented in figure 2.5 as pulled/fixed option: **A**, pulling direction: **n1**.

Stretch Fragmentation

The Stretch deformation pulls on the fibril along its length until it reaches its tensile limit. The pulling direction is parallel to the fibril long axis. This probes the hydrogen bonding network between the peptides as the steering forces are applied along the same direction. The stretch SMD protocol requires the backbone carbon-α atoms of peptides on one end of the fibril ('bottom') to be held fixed whilst the center of mass of the backbone carbon-α atoms at the other end ('top') of the fibril are pulled linearly. This selection takes into

account the random fraying of strands at the ends of the fibril during the MD phase. The peptides in between these two selections and all other atom types are free to move. This is represented in figure 2.5 as pulled/fixed option: **B**, pulling direction: **n2**.

Slide Fragmentation

The slide deformation scheme drags one β-sheet across another with the pulling direction parallel to the fibril long axis. The hydrophobic core interactions as well as the steric packing arrangements between the β-sheets are of interest in this scheme. The slide SMD protocol uses the same pulled/fixed atom choices as in peel, however, the pulling direction is along the long axis of the fibril instead. This is represented in figure 2.5 as pulled/fixed option: **A**, pulling direction: **n2**.

Shear Fragmentation

The shear deformation applies steering forces perpendicular to the fibril axis thus addressing the hydrogen bonding network between the peptides from another direction. This differs from the stretch scheme in that the pulling direction is along the peptide axis (i.e perpendicular to the hydrogen bond network) although the same atom selections are used in both cases. This is represented in fig 2.5 as Fixed/pulled option: **B**, pulling direction: **n3**.

2.4.2 Choice of SMD parameters for fragmentation protocols

Before applying the above mentioned fragmentation protocols, careful consideration was given to the selection of key quantities that govern the recorded mechanical responses. Clearly the choice of pulled/fixed atom selections as well as pulling direction affect the trajectory taken by the molecule and ultimately influence the force time profile produced by the SMD simulation. Another crucial contribution is the choice of parameters used in the steering potential in eqn (2.8) namely the magnitude of the pulling velocity v, and stiffness of the virtual spring k . In addition to these factors, the response of the molecule may depend on the MD forcefield that is chosen to account for inter-atomic interactions during the simulation. Given the importance of each of these variables, the following section is dedicated to explaining the reasoning and experimental justification behind the various choices made for the quantities employed in this thesis.

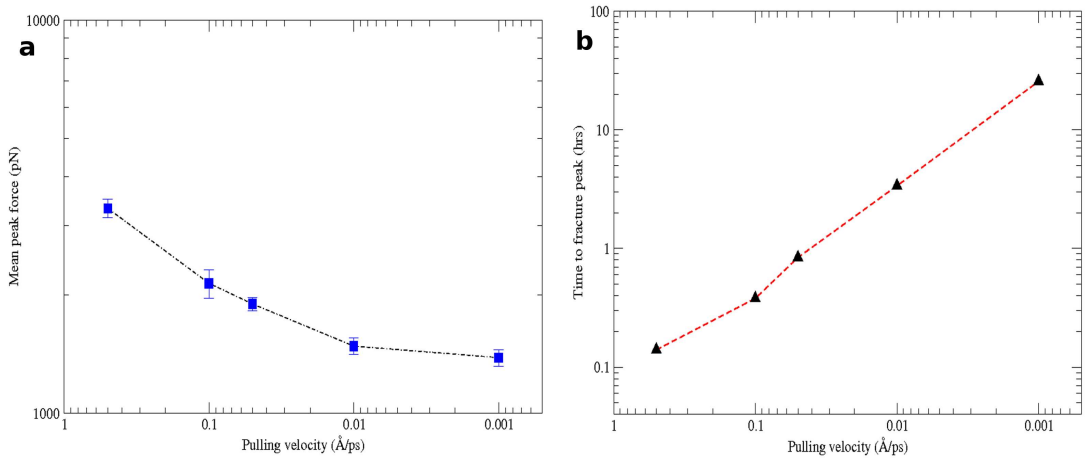

Figure 2.6: Trial simulations of peel SMD using different pulling rates for the 8x2-hIAPP$_{20-29}$-F4L variant fibril model. (a) Mean peak forces recorded, error bars are standard error in mean from four repeats in each case. (b) Benchmark performance of the computational cost required to reach each peak for the different velocities. All simulations run using 16 processors on Arc1-Leeds computational architecture.

1: Choice of the Pulling Velocity v

Ideally, a constant velocity SMD simulation should be carried out as slowly as possible as this not only gives the system enough time to respond to the mechanical perturbation but also keeps it close to equilibrium. However, the use of very slow pulling velocities may prove both computationally intensive and expensive especially for large systems such as those modeled in this work. Meanwhile, the use of very fast pulling speeds tend to lead to highly exaggerated peak forces being recorded. A compromise between the two extremes is to seek an optimal pulling velocity that is sufficiently slow to adequately probe the system but not too slow as to be computationally restrictive. As such, trial simulations were run in order to select an ideal pulling velocity for use in full scale SMD. In these trial runs, pulling velocities of 0.5, 0.1, 0.05, 0.01 and 0.001Å/ps were used for the peel deformation with each simulation repeated four times. The relative differences in mean-peak forces recorded between successive pulling speeds were then contrasted against computational cost in terms of the time required to reach the relevant peak force (fig 2.6). As expected higher peak forces are recorded when high pulling rates are used. However, for the two slowest speeds (0.01 & 0.001Å/ps) used, similar mean peak forces are observed. Given that the slower pulling velocities tend to more accurately capture the mechanical response of

the probed molecule relative to faster ones, the computational cost in running both these simulations was then contrasted. The benchmark data shows that by using a pulling velocity of 0.01Å/ps, fibril fracture could be achieved $\approx 8\times$ faster than at 0.001Å/ps even though similar mean peak forces are recorded in both cases. Thus in relative terms, 0.01Å/ps seems to strike an optimal compromise between computational cost and reasonable accuracy. This velocity was deemed suitable for the bulk of the work described in later chapters and used exclusively unless otherwise stated. The order of magnitude for the pulling velocity is also typical of that reported in the literature concerning SMD simulations ($\approx$ 0.01 - 0.1 Å/ps) of various protein molecules (64; 65; 90).

2: Choice of the Spring Constant k

The choice of the stiffness constant, k, for the virtual spring used in the SMD steering potential (eqn 2.8) has significant impact on both the trajectory and relative fluctuations in key observables. In deciding on an appropriate stiffness, one might be influenced by actual values used in analogous AFM pulling experiments. These tend to employ soft cantilevers with nominal k-constants such as those used in amyloid nano-indentation experiments which lie in the 0.05 - 0.60 N/m range. As discussed previously, cv-SMD uses pulling rates that are many orders of magnitude faster than found in the experimental realm. This makes the minimisation of the thermal fluctuations in the position of the pulled atom(s) essential if detailed information is to be extracted from such simulations. The displacement amplitude δx of the steered atom(s) along the reaction coordinate due to thermal energy fluctuations at a given temperature T can be expressed in terms of the Boltzmann constant k_B as:

$$\delta x \approx \sqrt{\frac{k_B T}{k}} \tag{2.12}$$

In order to minimise spatial thermal fluctuations, eqn (2.12) implies that a sufficiently large stiffness constant has to be employed. The use of a large spring constant is also an essential requirement for accurate PMF reconstruction based on the cumulant expansion approximation of the Jarzynski equality eqn (2.11). However, the use of a very stiff spring also leads to large fluctuations in the external force δF which correlates to significant fluctuations in the computed PMF profiles.

$$\delta F \approx \sqrt{k_B T k} \tag{2.13}$$

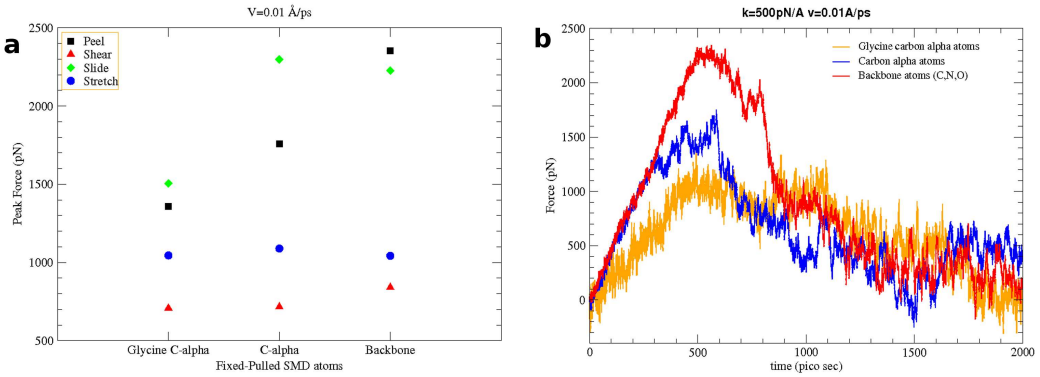

Figure 2.7: Results from trial simulations utilising three different types of pulled/fixed atom selections in 8x2-hIAPP$_{20-29}$ anti-parallel polymorph. (a) Peak forces obtained from the four fragmentation protocols (b) The force profiles selected from the peel simulations (black squares on left plot).

Extensive computational and theoretical studies to establish just how stiff the spring can become before there are significant fluctuations in the computed PMF have been carried out (81; 90). These studies surmise that a value of 500pN/Å is sufficiently large to ensure that thermal fluctuations in both the position of the pulled atom(s) and computed PMF are adequately minimised. Although PMF profiles are not extracted in analysis of the work described in this thesis, the practical & theoretical reasons for choosing a stiff spring in the SMD are still valid. As such, a stiff spring of 500pN/Å was selected for the SMD simulations described throughout this thesis. This ensures that the analysis of SMD data generated in this work can be extended in the future if computational resources permit for statistically significant SMD repeats for the large systems presented here.

3: Fixed & pulled atom selection

Selecting different atoms in the same region of a molecule to be pulled or fixed can have an impact on the recorded mechanical behaviour. Amyloid fibrils are mechanically robust structures that are composed of collections of individual interacting peptides. This makes the choice of appropriate atom groups that are harmonically steered particularly challenging since they have to be of a significant number to bring about the desired response in these mechanically robust structures. The selections also have be sequence independent if comparisons between fibrils with amino-acid modifications are to be made. This leaves just the atoms located at the peptide backbone as potential candidates from which the selections

could be made. A number of trial SMD simulations were carried out using three possible types of backbone atom selections for the different pulling schemes outlined in §2.4.1. These were (i) the carbon-α atoms in the central residues of the peptides, (ii) carbon-α atoms from all residues and (iii) backbone atoms from all residues (carbon, carbon-α, nitrogen & oxygen). The results of the trial simulations are depicted in fig 2.7. This shows that selection (i) gave rise to force profiles with a lot of noise. Although selection (iii) gave the least noisy force profile, there wasn't a great spread between forces that address similar interactions from different directions (i.e peel vs slide & stretch vs shear). Selection (ii) which employs the carbon-α atoms from all residues in relevant peptides gave good discrimination between peel, slide, stretch and shear simulations. Ultimately, this was chosen as the standard in all SMD investigations as it also allowed for introduction of amino-acid substitutions without changes in pulled/fixed atoms.

4: Forcefield validation

When interpreting the force response from SMD trajectories, its important to distinguish if the observed trends are physically valid or if they are just purely artifacts of the chosen forcefield. As such, the SMD simulations were validated by use of two different biomolecular forcefields, namely Charmm22-cmap and Amberff99SB. The trends in mechanical properties observed from the forcefields were contrasted against each other and these are described in the relevant chapters.

2.5 Simulation details and set-up

Once the various fibril models had been constructed and the fragmentation schemes designed, the next phase was to run dynamic simulations. Molecular dynamics were carried out to energetically refine the models as well as collect thermodynamical information. At the conclusion of the MD simulations, the model fibrils were mechanically probed from different directions by SMD simulation. Below is a detailed description of both the MD and SMD protocols common to all models investigated in chapters 3, 4 and 5.

2.5.1 Molecular Dynamics Simulation Protocols

The MD simulations and the associated analysis were performed by the use of the AMBER9, AMBER10 & AMBER11 (91) suite of programs, utilising the amber99SB all-atom

forcefield. The fibril models were explicitly solvated in a periodic box of TIP3P water molecules. Depending on the net charge of the amyloid molecule, appropriate numbers of Na^+ counterions were added via the LEAP module in AmberTools to make the systems charge neutral. The solvated models were then subjected to careful minimisation, heating and equilibration steps prior to the production MD runs:

Minimisation, Heating and Equilibration steps:

The constructed structures (elongated or sequence modified) were not energetically optimised, it was necessary to carry out energy minimisation and equilibration of each system before any long MD runs could be made. The minimisation was carried out in two stages, first with the protein solute held fixed thus minimising the solvent. In the second stage both the solvent and solute were allowed to move resulting in the entire system being minimised. Both minimisation stages were carried out at constant volume. After this each system underwent equilibration at constant pressure using a multistage protocol with positional restraints placed on the protein. The systems were heated from an initial temperature of 100K to a final constant temperature of 300K. For each step the SHAKE algorithm was used to constrain all bonds to hydrogen atoms. The purpose of this was to remove largest oscillation frequency from the system thus allowing a 2 femto-second timestep in the integration of Newton's equations of motion for the ensuring MD. Each of these steps lasted 10 ps with the restraints getting progressively smaller in each stage. The last step in the equilibration process was with the all restraints removed and MD run for 100 ps, allowing the system time to relax. A summary of the steps in this process is given in table 2.1.

Production Dynamics set-up:

After the equilibration phase, long MD simulations were conducted. All MD was run at constant temperature (300K, Berendsen thermostat) and constant pressure (1 atm). Long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) method with a 9 Å cut-off. The duration of the simulations ranged between 20-40 ns with the length determined by convergence of the root mean square deviation (RMSD) value of backbone atoms. Additional forcefield validation MD was run using the charmm22-cmap forcefield via NAMD2.7b1. For the NAMD validation simulations, the minimisation and equilibration the steps used for AMBER MD were repeated with the RATTLE algorithm

Steps prior to MD phase:		
Stage	**Details**	**Duration**
Minimisation	Solute frozen, solvent & ions free to move	5000 cycles
Minimisation	Solute, solvent & ions all free to move	5000 cycles
Heating	Solute restrained k =100, initial temp=100K	10 ps
Heating	Solute restrained k =100, target temp=300K	10 ps
Equilibration	Solute restrained k =50, temp=300K	10 ps
Equilibration	Solute restrained k =25, temp=300K	10 ps
Equilibration	Solute restrained k =10, temp=300K	10 ps
Equilibration	Solute restrained k =5.0, temp=300K	10 ps
Equilibration	Solute restrained k =2.5, temp=300K	10 ps
Equilibration	Solute restrained k =1.0, temp=300K	10 ps
Equilibration	restrains removed, temp=300K	100 ps

Table 2.1: Multistage minimisation, heating and equilibration protocol employed prior to full scale production MD. The Units of the positional restrains, k, are in kcal/molÅ^2.

employed to constrain bonds to hydrogen atoms. Constant pressure was maintained with a Langevin piston of 100 femto-sec decay period and 50 femto-sec damping time. The PME cut-off was 12 Å for long range electrostatics whilst a switching function (with twin range of 10 Å and 12 Å) was used for van der Waals interactions.

2.5.2 Steered Molecular Dynamics Simulation Protocols

The fibril conformations obtained at the conclusion of the long MD simulations were then used as the starting structures for the SMD simulations to characterise mechanical stability. These were run on the NAMD2.7b1 (78) package, utilising the charmm22-cmap forcefield. The models were re-solvated in a larger periodic water box that would allow full mechanical deformations of fibrils without self interactions. The duration of each simulation allowed for displacement of the pulled peptides to be achieved as described in §2.4.1 . The simulations were carried out at constant temperature (300K) and constant pressure (1 atm). The constant velocity pulling approach was employed with a pulling rate of 0.01Å/ps used throughout unless otherwise stated. The stiffness constant of the virtual spring was set at 500pN/Å. Each SMD simulation was repeated four times, with randomised starting velocities according to the Maxwell-Boltzmann distribution used in each repeat to ensure the trajectories sampled different areas of phase space. Additional validation simulations to check the dependence of mechanical responses on the forcefield were carried out using the amber99SB all-atom forcefield.

2.5.3 Computational Resources

High performance computing (HPC) resources were used to carry out the simulations described in this work. Computational resources were initially supplied by the UK National Grid Service (NGS) and then by the University of Leeds through *Everest & Arc1* nodes. Everest is an AMD Opteron based cluster which has since been decommissioned and replaced by Arc1 whose architecture is Intel Nehalem based. The parallel multiprocessor platform of these resources allowed for the AMBER and NAMD codes to conduct large scale simulations at computationally reasonable rates.

2.6 Analysis Methods and Calculations

A variety of methods were used to analyse both the MD and SMD trajectories. The analysis allowed for determination of the factors influencing structural, thermodynamic and mechanical properties of the model fibrils throughout the different computational investigations in chapters 3, 4 and 5. For MD trajectories, the analysis was carried out only after convergence of relevant RMSD values had been achieved.

Secondary Structure Content

Secondary structure content such as β-strands, turns and random coils were assigned by use of the DSSP method (92) via the PTRAJ module of the AMBER-Tools package. The content of random coil and beta motifs in the peptides was monitored every 1ps for the final ≈ 10 ns of the converged MD simulations. The loss or gain of secondary structure motifs by the peptides in the fibrils are a good indicator of relative stability as a simulation progresses.

Hydrogen Bonding

Hydrogen bonding occurs when a hydrogen atom that is covalently bonded to an electro-negative atom (donor) interacts with another electro-negative atom (acceptor) through non-bonded short range interactions. The HBONDS utility of VMD (93) was used to analyse the occupancies of backbone and side-chain interstrand hydrogen bonds. The presence of a hydrogen bond was defined by distance and geometry considerations. These were satisfied if the distance between the donor and acceptor atoms was less than 3.5Å

with a 135° cut-off for the angle between the three atoms. The number of hydrogen bonds was computed every 1ps from the final ≈ 10 ns of the converged simulations.

Shape Complementarity, S_C

Shape complementarity, S_C, is a statistic designed to quantify the geometrical packing of protein interfaces (94). This compares the directions of the unit vectors from nearest points of opposing surfaces. The quality of fit between the two surfaces is obtained from the average dot product of the pairs of vectors. For two perfectly fitting surfaces this value is 1 whilst for surface that have no matching the value becomes 0. The S_C of the interface between two pairs of β-sheets was measured using the CCP4 toolkit. The default settings for the parameters used in the measurements were implemented without alteration. Peptides at the ends of the fibrils were excluded from the calculations due to their random fraying. Hydrogen atoms were stripped to allow the progression of the calculation. Measurements were taken in 10 ps intervals from the last ≈ 10 ns of the converged simulations.

Electrostatic Potential

The electrostatic potential energy around a system of point charges can be thought of as the amount of work required to bring those charges from a distance at infinity to their final positions. For a system made up of N charged species, the electrostatic potential energy, $U_{(ep)}$, can be expressed as in eqn 2.14 whereby V is the potential of the ith particle with electric charge q_i at location $\mathbf{r}$.

$$U_{(ep)} = \sum_{i}^{N} q_i V(\mathbf{r}_i) \tag{2.14}$$

The electrostatic potential of the inter-sheet interface of the fibrils was calculated and mapped out by use of the Adaptive Poisson-Boltzmann Solver (APBS) program (95). The final configuration of each fibril model at the conclusion of the MD simulations, stripped of their solvent molecules was used for these calculations. Non-bonded interactions between the β-sheets during the SMD deformations were also monitored by the NAMDenergy utility in VMD (93).

Enthalpy & Binding Energy Calculations

The Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) methodology as implemented in AMBER11 was used to calculate the enthalpies of the fibril models. These post-processing calculations were carried out after the explicit solvent molecules from the MD trajectories had been stripped off. The MM-PBSA method calculates the enthalpy, ΔH, from summing the energetic contributions from the molecular mechanics forcefield term, E_{MM}, and the continuum solvation energy term, E_{PBSA}, (eqn 2.15).

$$\langle \Delta H \rangle = \langle E_{MM} \rangle + \langle E_{PBSA} \rangle \tag{2.15}$$

The molecular mechanics term is composed of the averages of energy interactions that are directly mesured from the bonded and non-bonded components of the forcefield (eqn 2.3) from which the original simulation with explicit solvent was run. Namely, these are the sum of the averages of bond, angle, dihedral, van der Waals and electrostatic energies (eqn 2.16).

$$\langle E_{MM} \rangle = \langle E_{bond} \rangle + \langle E_{angle} \rangle + \langle E_{dihedral} \rangle + \langle E_{vdw} \rangle + \langle E_{elec} \rangle \tag{2.16}$$

Meanwhile, the molecular solvation free energy term is made up of two terms; the electrostatic contributions from molecular solvation, E_{PB}, and the solvent accessible area, A, which is scaled to the surface tension, γ of the solvent as shown in eqn 2.17 where the Poisson-Boltzmann equation is then used to evaluate E_{PB}.

$$\langle E_{PBSA} \rangle = \langle E_{PB} \rangle + \gamma \langle A \rangle \tag{2.17}$$

The MM-PBSA method was also used to compute the binding energy between pairs of β-sheets. All MM-PBSA calculations were complemented by equivalent MM-GBSA (molecular mechanics/generalized Born surface area) calculations. Snapshots were taken in 10 ps intervals from the last ≈ 10 ns of the converged simulations.

Chapter 3

Role of peptide sequence in mechanical resistence of hIAPP$_{20-29}$ protofibrils

This chapter is based on the paper Ndlovu et al. published in the Biophysical Journal (96).

3.1 Overview

Synopsis

In this chapter, the effects of sequence substitutions on the mechanical properties of hIAPP$_{20-29}$ derived amyloid proto-fibrils are explored. Seven variant fibril models having systematic sequence changes that progressively evolve from the amyloidogenic human amylin fragment to its non-amyloidogenic rat counterpart are constructed and then probed from four independent directions. The contribution of sequence specific intermolecular interactions as well as the coupling between sequence and the nature of pulling on the mechanical properties are studied. The main focus of this chapter is how the structural order imposed by the amino-acid residue substitutions on the β-sheet assemblies impacts on both the thermodynamic and mechanical stability of the proto-fibril models. The results presented here have wider implications for understanding the importance of sequence-dependent mechanical properties of amyloid systems such as their growth kinetics (seeding), the role of fragmentation events in cytotoxicity and in the design of materials for nanotechnology applications.

Introduction

The specific role played by the amino-acid sequence of a peptide in amyloid formation is still not fully understood. There is evidence to suggest that the side-chain interactions that determine the different characteristic folds of native proteins may also dominate the aggregation forces depending on the environmental conditions (3; 4). The aggregation propensity of a peptide may be increased by changes to the sequence through site mutations if the effect of the amino-acid substitution leads to to either an increase in hydrophobicity, decrease in the net charge, decrease in the alpha-helical propensity or an increase the beta-sheet propensity of the sequence (97). Curiously however, it has been suggested that the ability to assemble into ordered cross-beta structures is not an unusual feature possessed by a small number of peptides with special sequence or structural properties but is rather an inherent characteristic of all polypeptide chains (23). Under the right environmental conditions, such as temperature, pH or salt concentration, any polypeptide can be amyloidogenic. Despite the details of the sequence composition, amyloid fibrils

exhibit very similar structural characteristics, stability and mechanical properties. Given the diversity in proteins/peptides that readily form amyloid fibrils, it is intriguing to decipher the molecular origins of their robust mechanical character. Since the fibrils share common backbone compositions and conformations, this might suggest that the generic properties common to all aggregates, such as backbone hydrogen bonding, may play much more of major role in their mechanical stability than the side-chain interactions. Experimental work on short peptide micro-crystal arrays of the GNNQQNY sequence have shed light on the role of the amino side-chains suggesting they play an integral role in maintaining the structural stability of the pair of β-sheets that form the unit of a fibril (19). The work reveals that the side-chains on interacting pairs of sheets fit tightly together in a very complementary manner like teeth in a zipper (19). This dry and tight steric fit interface between the pairs of β-sheets is thought to be a fundamental feature of all amyloid-like fibrils (19; 20) and is mediated by the side-chain residues.

A precise understanding of the relationship between the peptide sequence and the mechanical properties of amyloid fibrils is important since mechanical stability is linked to frangibility which in turn has implications for cytotoxicity and secondary nucleation events (proliferation of seeds). In early onset type-II diabetes characterised by a single amino-acid change in the amylin protein (98) for instance, increased fibrillisation kinetics and enhanced cytotoxicity relative to wild-type amylin is observed (99). This is entirely consistent with increased brittleness in the fibril which is most likely due to the single site amino-acid change in the hIAPP$_{20-29}$ region (**SNNFGAILSS** $\rightarrow$ **GNNFGAILSS**). As the brittleness of amyloid fibrils may affect aggregation kinetics and toxicity, understanding the factors that attenuate their mechanical character could aid in the interpretation of seeding & toxicity experiments and could ultimately provide a basis for the design of therapeutic agents to combat amyloid disease. Knowledge of the sequence-dependent factors that drive mechanical stability in amyloid fibril assemblies would also be particularly useful in the bottom-up design and construction of novel nano-materials.

In the work described in this chapter, SMD simulations are used to apply sufficient mechanical forces to induce fragmentation in a number of fibril models from different directions in order characterise the role of the amino-acid sequence in determining their mechanical properties. The simulations probe the stabilising hydrogen bond and hydrophobic-core interactions in the fibril models. Fibrils formed from the SNNFGAILSS sequence are used and

substitutions that progressively evolve from the amyloidogenic human amylin fragment to its non-amyloidogenic rat counterpart are made (see table 3.1). The choice of the SNNFGAILSS sequence for this particular investigation was inspired by the fact that it is a computationally tractable model system widely used to study amyloid related phenomena for which a wealth of experimental and computational data is available. The role of the amino-acid sequence on the amyloidogenic properties of amylin derived peptides has been extensively explored through alanine and proline scanning on various segments of the amylin protein (100; 101; 102). These studies mainly highlight the importance of certain residues in the overall ability of the peptide to form amyloid fibrils. Substituting aromatic residues for instance does not affect the ability to form fibrils but significantly slows down the rate at which they form (101). Meanwhile, some substitutions with proline along the sequence lead to the total loss of amyloidogenicity (102) . Despite this catalogue of work however, very little has been done to understand how these sequence substitutions affect the mechanical stability of fibrils they eventually form. The choice of the Rat-IAPP sequence to guide the systematic residue changes in the work outlined in this chapter is due to its inability to form stable amyloid fibrils in vitro (67) and as such acts as an ideal control for the SMD simulations. Furthermore, the Rat-IAPP inspired substitutions do not lead to the addition or removal of residues that are a potential donors or acceptors of side-chain hydrogen bonds thus the effects of the side-chain interactions can be unambiguously evaluated without direct changes to the hydrogen bond networks. A total 112 SMD simulations were carried out to mechanically characterise the wild-type and sequence variant fibril models.

3.2 Simulation Details

3.2.1 Sequence variants Construction:

The initial fibril model was built from coordinate files derived from solid state NMR data for the parallel polymorph of SNNFGAILSS fibril. The explicit details for the elongation and refinement methods used to construct model fibrils from the ssNMR coordinate template are described in § 2.3.1. The constructed fibril consisted of a pair of β-sheets, each containing eight peptides (an 8x2 fibril model). The peptide sequence of the 8x2 model was then altered to produce six distinct fibril variants in addition to the original SNNFGAILSS model, as shown in table 3.1 using the MUTATOR module in the VMD program (93). Peptide backbone angles were kept the same as in the original 8x2 model and atoms common to

Seven sequence variants of the hIAPP$_{20-29}$ proto-fibrils	
Fibril Model	**Amino-acid sequnce**
Wild-type (hIAPP)	**SNNFGAILSS**
F4L	**SNNLGAILSS**
A6P	**SNNFGPILSS**
I7V	**SNNFGAVLSS**
F4L-A6P	**SNNLGPILSS**
F4L-A6P-I7V	**SNNLGPVLSS**
Rat-IAPP	**SNNLGPVLPP**

Table 3.1: Models built from substitutions of amino-acids from the original hIAPP$_{20-29}$ sequence. The specific changes are indicated in red.

both the substituted residue and the original amino acid were left unchanged from the ssNMR structure. Positions of the remaining atoms that make up the substituted residue were then designed based on considerations of the internal coordinates of that particular amino acid. The relevant substitution was applied to all 16 peptides in the pair of β-sheets. The naming convention adopted for these substitutions arise from numbering the amino-acid residues of the SNNFGAILSS sequence from 1 to 10 (table 3.1). For example, the F4L variant relates to a substitution of Phe in the fourth position with a Leu residue.

MD simulation details

After the models had been constructed, MD techniques were used to energetically refine and extract thermodynamic properties from the fibrils. The MD simulation protocols used in this study are outlined in great detail in § 2.5.1, only key details are summarised here. MD simulations were run in both the AMBER9 and NAMD2.7b1 software packages, with the all-atom Amber99SB and the Charmm22-cmap forcefields, respectively. The 8x2 ssNMR model and the associated fibril variants were explicitly solvated in a periodic water box of TIP3P molecules (fig 3.1). The fibrillar assemblies were then subjected to the careful minimisation and equilibration steps listed in table 2.1. Inter-strand backbone hydrogen bond distance restraints were temporarily imposed for 1 ns prior to MD production runs, which allowed the fibril models to adopt relaxed conformations with a net twist. Unrestrained production MD lasting 22 ns for each fibril model was then run at constant temperature (300K) and pressure (1 atm). Convergence was monitored using the root mean-square deviation and was sufficient for all models within the final 6 ns of 22 ns production MD (fig 3.2). No

significant structural changes were observed after 10 ns of forcefield validation simulations rerun in NAMD using Charmm22-cmap.

SMD mechanical probing

Conformations of the fibrils from the end of the unrestrained MD simulations served as initial structures for steered molecular dynamics (SMD) simulations to characterise mechanical stability. SMD was performed using the NAMD2.7b1 package and the Charmm22-cmap forcefield. The four SMD fragmentation protocols (peel, stretch, shear, and slide) as described in § 2.4.1 were applied to each of the seven variants. Each simulation was repeated four times, with randomised starting velocities to ensure the trajectories sampled different areas of phase space thus a total of 112 individual SMD simulations were run. Each SMD simulation lasted 4 ns with exception of peel-SMD which ran for 2 ns due to the smaller displacements required for this deformation mode. For the rat amylin sequence (rIAPP$_{20-29}$), which collapses during the equilibration phase into a disordered aggregate, some of the SMD protocols (such as sliding sheets past each other) become ambiguous and difficult to implement. In these cases, the deformation simulations were carried out for the rIAPP variant using carbon-α atom selections consistent with those used in the other models. Additional simulations to check the dependence of mechanical response on the choice of forcefield were carried out using the Amber99SB force field. The trends recorded in the simulations using the Charmm22-cmap and Amber99SB force fields were similar (fig A.1a in Appendix A). A series of peel simulations was also performed at different pulling velocities (0.5, 0.1, 0.05, 0.01, and 0.001 Å/ps), which demonstrated that the trends in peak force between the sequence variants are not unique to the pulling velocity chosen during this deformation (fig A.1b in Appendix A).

The SMD simulations produced noisy force-time profiles, especially in the case of the shear deformation. To characterise the magnitude of the noise in the signals, further steered MD was performed for an already fractured aggregate. This simulation measured the force response of a detached fibril fragment purely as a result of being pulled in the water box environment. The resulting noise signal (fig 3.3) indicates the baseline level for the force on a fragment in the absence of interactions with the rest of the aggregate. The data was smoothed using a moving median filter with averaging over a 100 ps timescale. The moving median filter noise reduction strategy was applied to all the force-profiles obtained from the

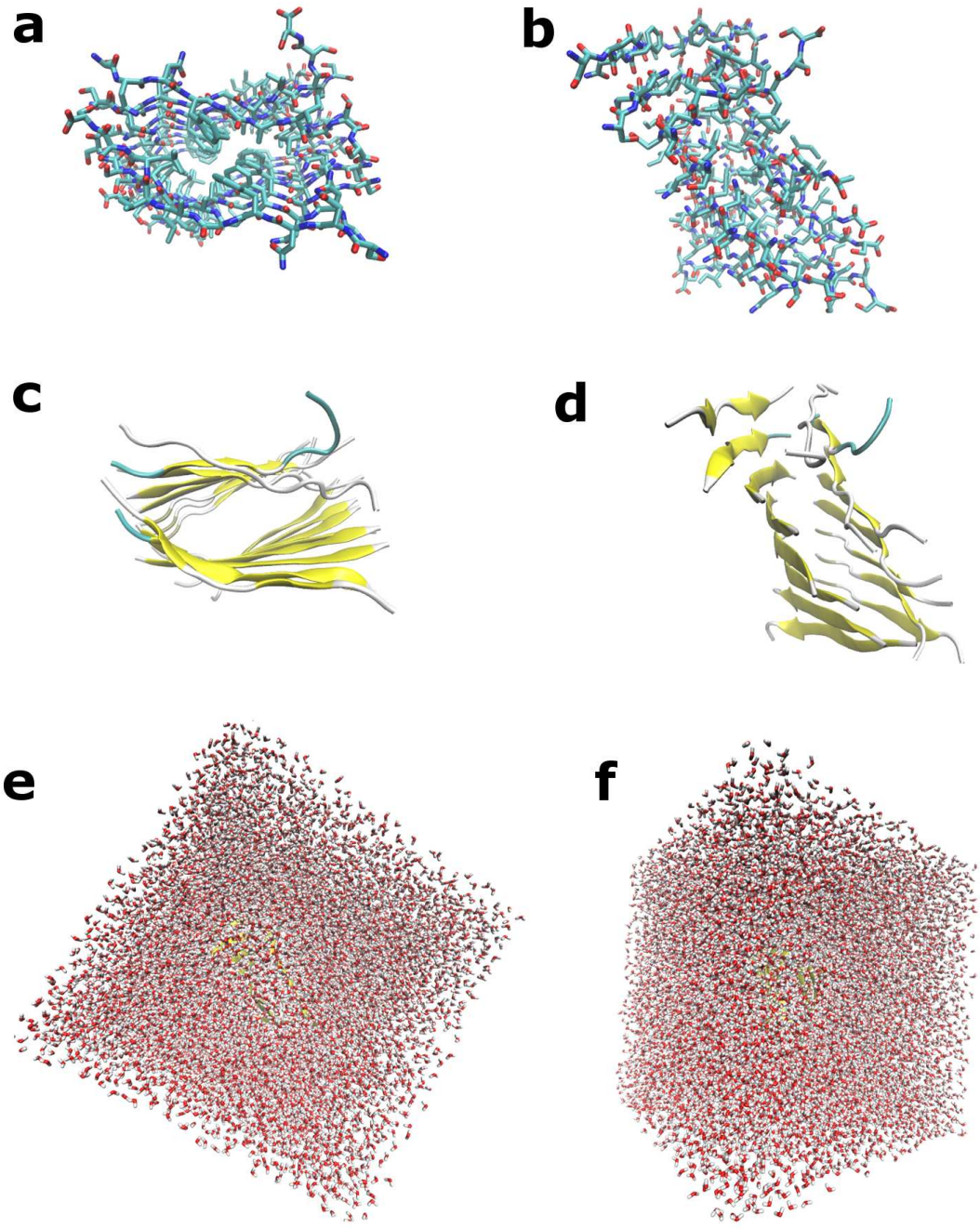

Figure 3.1: The 8x2 wild-type hIAPP$_{20-29}$ model prior to SMD simulation. Images on the left (a, c and e) show the top view of in 'stick', 'cartoon' and atomistic representations respectivley. Images on the right (b, d and f) show the corresponding side views. The fibril model is surrounded by a box of explicit TIP3P water molecules during MD (e and f) and the solvated model consists of ≈ 50000 atoms.

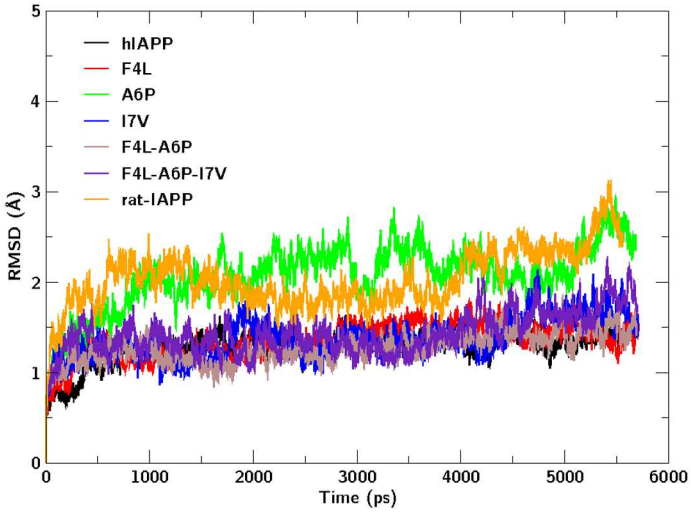

Figure 3.2: Root mean square deviation (RMSD) of the backbone carbon-α atoms for the final 6 ns of the 22 ns MD simulations. The edge strands on both ends of the fibrils were not considered in the RMSD calculation to eliminate the effects of end-strand fraying. Although in this time frame most of the models have converged, there are significant fluctuations in the rat-IAPP and A6P mutants due to their disordered nature.

four SMD modes (peel, slide stretch and shear) to enable cross comparability in the peak force results.

3.3 Results: Structural reorganisation in models

The wild-type sequence fibril and its six derivatives were all allowed to evolve freely under MD conditions from the initial ssNMR structure. The configurations of each fibrillar system at the conclusion of the production MD simulations are shown in figure 3.4. These final configurations reveal structural organisation, due to sequence substitutions, that range from subtle changes to drastic reductions in the degree of order in the aggregates relative to the wild-type model. These changes were contrasted against experimental data, where available, on the expected morphology of the resulting aggregates.

In order to quantify and characterise the degree of order within the fibrils, secondary structure elements present during the course of the dynamics were monitored by use of the DSSP method, figure 3.5a. Additionally, the mean occupancies of the inter-strand backbone and side-chain hydrogen bonds are shown in figure 3.5b . Both calculations for secondary struc-

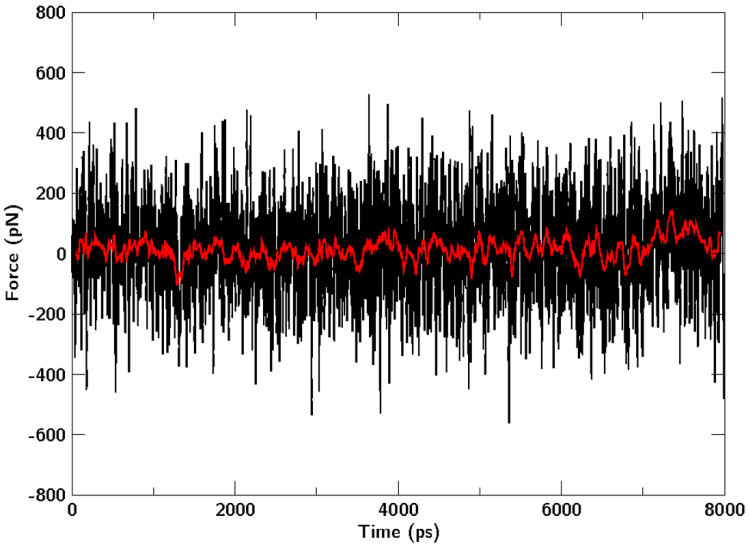

Figure 3.3: Background noise when a fragment is pulled through the solvent after initial fracture from the intact fibril.

ture elements and mean hydrogen bond occupancies were performed on snapshots sampled every 1 ps from the final 6 ns of the converged MD trajectories.

Effect of substitution on fibril conformation

The effect of altering the peptide sequence on the structural characteristics of each model relative to the wild-type are discussed in turn. Of the single point changes, the substitution of phenylalanine for leucine at position 4 (F4L) results in a model fibril that is most similar to that formed by the wild-type sequence after MD. The proportion of β-strands as well as the numbers of backbone and side-chain hydrogen bonds are identical (within error) in the two fibril models (fig 3.5). This observation is consistent with experimental data from electron microscopy (EM) and transmission electron microscopy images, which show that fibril aggregates form readily from the F4L variant, giving rise to fibrils that are morphologically identical to those formed from the wild-type peptide (67; 101).

For the isoleucine to valine (I7V) variant, EM images of fibril bundles obtained by Westermark et al. (67) indicate that they have a similar morphology to those formed from the wild-type peptide. However, during the MD simulations the I7V variant experiences a shift of one β-sheet relative to the other, with the N-terminal regions on the opposing sheets becoming more solvent-exposed than in the wild-type fibrils. Nevertheless, the aggregates

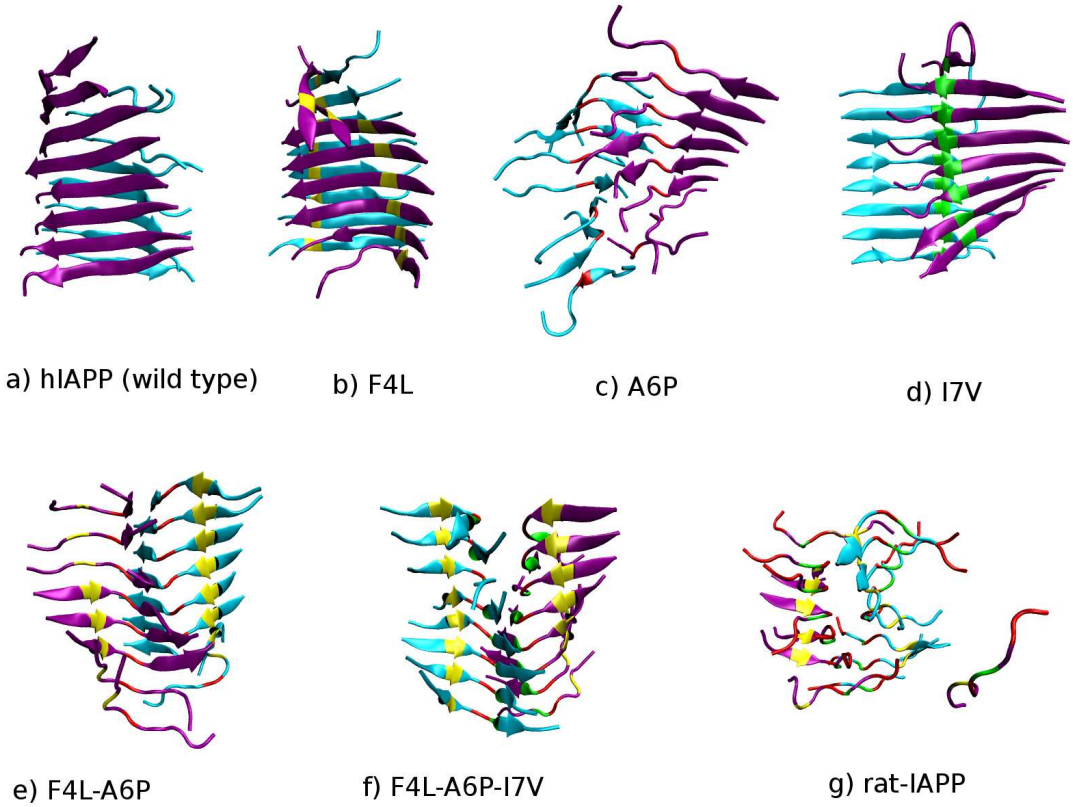

Figure 3.4: Configurations of the fibrils at the end of the 22 ns MD simulations performed in the absence of force. (a) Wild-type model with sequence SNNFGAILSS, (b) variant with F4 substituted with leucine (marked in yellow), (c) variant with A6 substituted with proline (marked in red), (d) variant with I7 substituted with valine (marked in green), (e) double substitution F4L/A6P, (f) triple substitution A6P/F4L/A6P, and (g) the non-amyloidogenic rat amylin fragment SNNLGPVLPP.

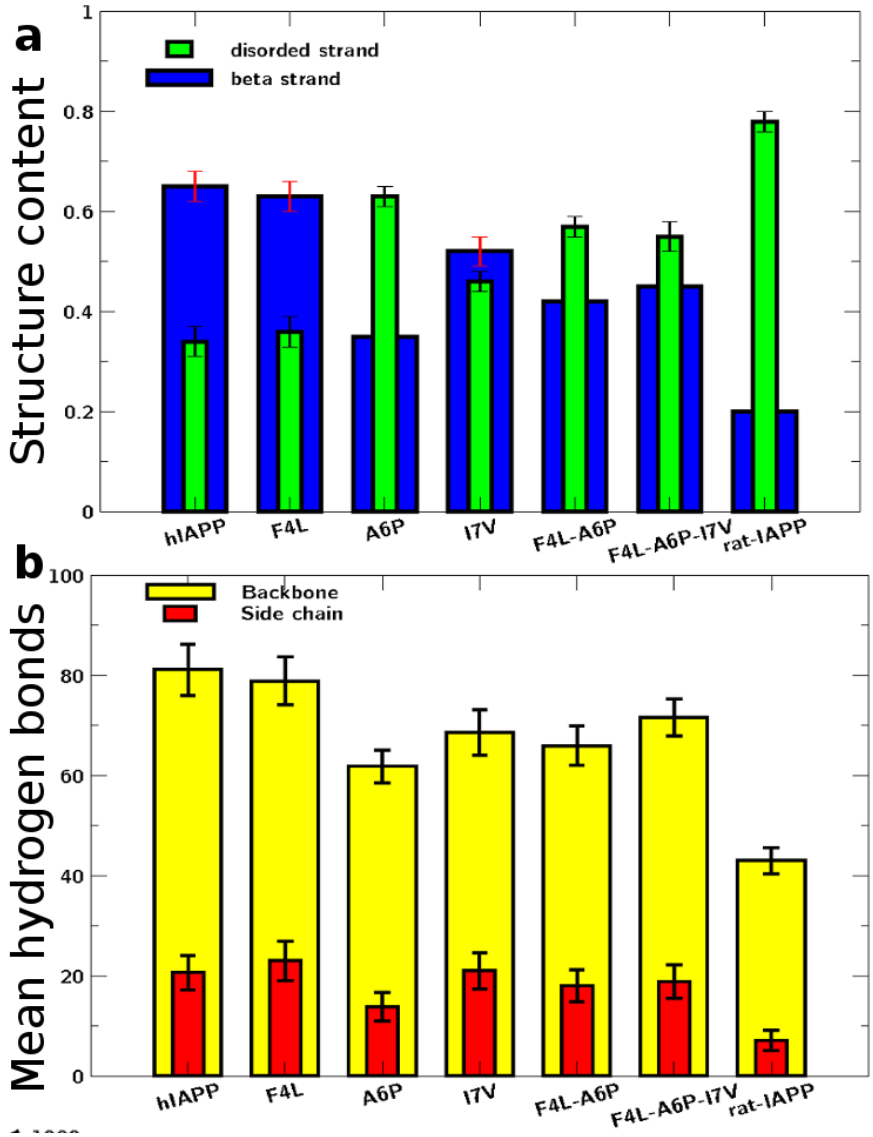

Figure 3.5: (a) Mean fraction of secondary structure elements present in each fibril model from 1 ps snapshots of the final 6 ns of MD simulation. (Blue) β-strand and (Green) random coil content. (b) Mean number of inter-strand backbone (yellow) and side-chain (red) hydrogen bonds recorded in the final 6 ns of MD

remain ordered, with only a 15% and 20% decrease in hydrogen bonding and β-sheet content respectively with no measurable change in the number of side-chain hydrogen bonds (fig 3.5).

In stark contrast to the above mentioned single point changes, the introduction of a proline residue in place of alanine (A6P) disrupts the fibril structure and results in an increase of random coiled regions at the expense of the β-strand motif, which is coupled to a reduction side-chain hydrogen bonding (fig 3.5). The resulting aggregates are also more visibly disordered relative to the wild-type (fig 3.4). The available images (from Westermark et al) from EM experiments show that fibrils formed from the A6P sequence are morphologically distinct from those formed by the wild-type peptide (67).

For the multiple point changes, the double (F4L-A6P) and triple (F4L-A6P-I7V) substitutions give rise to fibrils that share similar secondary structure characteristics to each other. The F4L-A6P-I7V variant has slightly more inter-strand backbone hydrogen bonds at the end of MD, with no detectable differences in side-chain hydrogen-bonding interactions between the two variants.

Finally, the fibril model mimicking the non-amyloidogenic rat sequence (Rat-IAPP) showed a dramatic transition from its initial ordered configuration into an amorphous globular-like aggregate of random coils. The presence of three proline residues resulted in a significant drop in both backbone and side-chain hydrogen bonding. These observations coupled with enthalpic calculations (see next section) are consistent with experimental studies that show that this sequence is unable to form stable amyloid fibrils (67). As such, the amorphous aggregates found in the Rat-IAPP model serve as the reference state for quantifying the mechanical resistance of ordered fibril models.

3.4 Results: Thermodynamic stability of models

The MM-PBSA methodology as implemented in Amber11 was used to calculate the relative enthalpy of each fibril model from snapshots sampled every 1 ps from the final 6 ns of the converged MD trajectories. The differences between the computed enthalpy in the variant fibril models are shown relative to the wild-type (fig 3.6). The individual energy terms contributing to the enthalpy for each model are listed in table 3.2

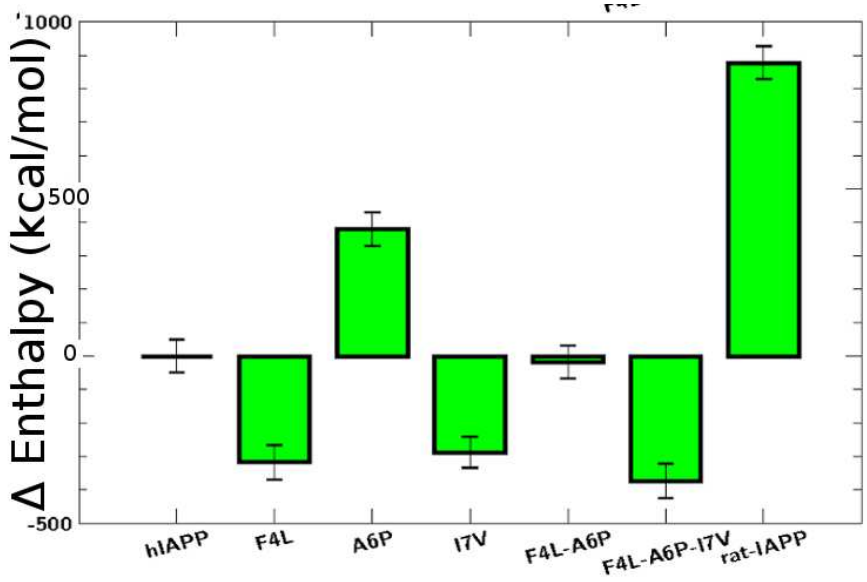

Figure 3.6: The configurational enthalpy of each variant model plotted relative to the wild-type fibril calculated from 1 ps snapshots of the final 6 ns of converged MD simulation.

Model	Int_{ene}	$Elec_{ene}$	vdW_{ene}	Sol_{ene}
Wild-type	2860.66 ± 32.16	-5389.51 ± 77.48	-524.18 ± 17.48	-2163.01 ± 70.85
F4L	2859.56 ± 33.41	-4513.50 ± 77.40	-501.51 ± 18.89	-3378.91 ± 72.89
A6P	3117.21 ± 33.19	-4212.08 ± 108.82	-495.09 ± 19.09	-3247.02 ± 105.45
I7V	2782.72 ± 31.10	-4185.31 ± 76.00	-454.00 ± 18.89	-3648.12 ± 70.91
F4L-A6P	3114.49 ± 31.63	-4384.55 ± 85.59	-574.31 ± 17.91	-3391.02 ± 77.40
F4L-A6P-I7V	3046.77 ± 31.57	-4807.71 ± 97.07	-538.79 ± 18.40	-3290.47 ± 97.14
Rat-IAPP	3482.45 ± 33.29	-4819.72 ± 69.74	-470.79 ± 19.09	-2531.60 ± 67.28

Table 3.2: Contributions to the enthalpy, calculated using MM-PBSA methodology. Units of kcal/mol, errors are the standard deviation in the mean. Int_{ene} Internal energy (sum of bond, angle and dihedral terms), $Elec_{ene}$ electrostatic energy, vdW_{ene} van der Waals contribution and Sol_{ene} solvation free energy.

A number of interesting features are evident from the relative enthalpy calculations. The wild-type sequence is not the most thermodynamically stable fibril structure. The F4L, I7V and F4L-A6P-I7V substitutions lead to the formation of more energetically stable fibril models whilst the double substitution does not affect the enthalpy. Apart from the solvation energy term in table 3.2, it is not immediately clear why the F4L, I7V and F4L-A6P-I7V variants appear more stable than the wild-type. Further analysis into the packing and contacts of the inter-sheet interface reveals that F4L, I7V and F4L-A6P-I7V have better

or equivalent shape complementarity (S_c) values, 0.72±0.03, 0.59±0.06 and 0.74±0.03 respectively in contrast to 0.61±0.04 for the wild-type.

The energetic effects of the sequence substitutions also appear to be additive i.e. the enthalpies of the double, triple and quintuple (rat-IAPP) substitutions have the sum properties of their constituent single substitutions. Finally, the massive reduction in enthalpy of the Rat-IAPP model relative to the wild-type is particularly reassuring as this is consistent with experimental studies which show it is unable to form stable amyloid fibrils (67).

3.5 Results: Mechanical response of fibril models

To characterise the different mechanical resistances of fibrils formed from the wild-type hIAPP$_{20-29}$ fragment and its six variants, structures taken at the end of the MD simulations were each subjected to probing by the four deformation SMD modes outlined in § 2.3.1 .

3.5.1 Mechanical responses to the SMD pulling modes

Figures 3.7, 3.8, 3.9 and 3.10 each show a selection of representative snapshots of the fibrils undergoing the peel, stretch, slide and shear SMD simulation respectively, with the relevant force-time profiles plotted above each snapshot. Figure A.2 in Appendix A shows all 16 force-time profiles, including repeats, measured for the wild-type sequence fibril in the four pulling directions. Although each repeat SMD simulation follows a slightly different path through conformational space, similar features are observed in force-time plots within each distinct set of pulling mode results. All force profiles recorded for the variant fibrils are included in Appendix A (figs A.3, A.4, A.5 and A.6). The magnitude of the mean peak force, calculated over the set of four repeat simulations for each pulling mode, was used as a measure of the mechanical resistance of each fibril to deformation along a particular direction. The peak force is defined as the maximum force exerted during a given SMD simulation after the data has been filtered to reduce thermal noise.

Mechanical performance across the sequence variants:

The recorded mechanical response for all seven fibril models across the different fragmentation pulling modes are summarised in figure 3.11. The mean peak forces, post noise

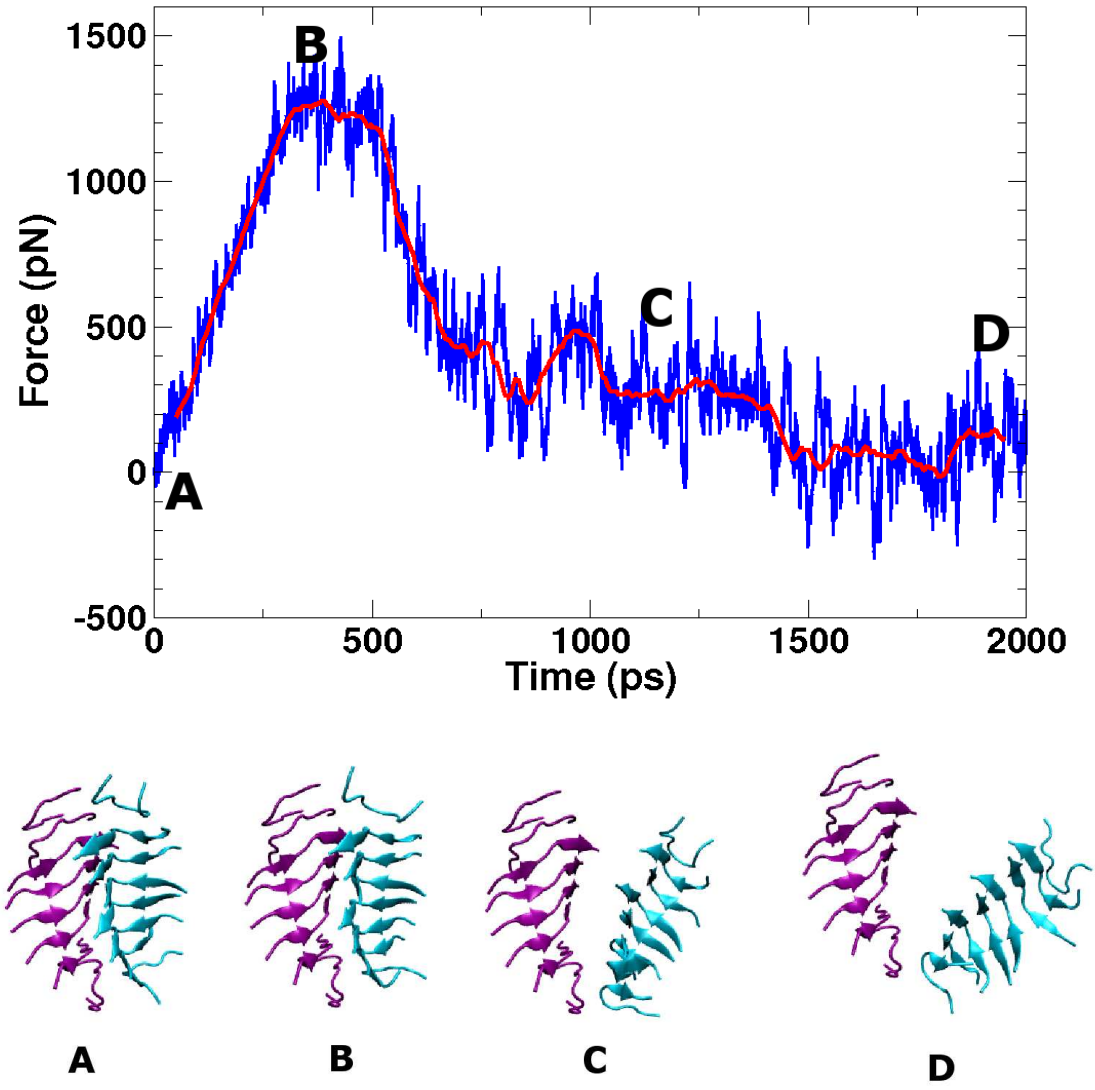

Figure 3.7: **PEEL SMD:** Force profile obtained during the Peel SMD pulling geometry for F4L-A6P fibril model trajectory. Points marked A, B, C and D on the plots correspond to the equivalently marked representative pictorial snapshots (below graph) of the fibrils evolving in time as the force is applied.

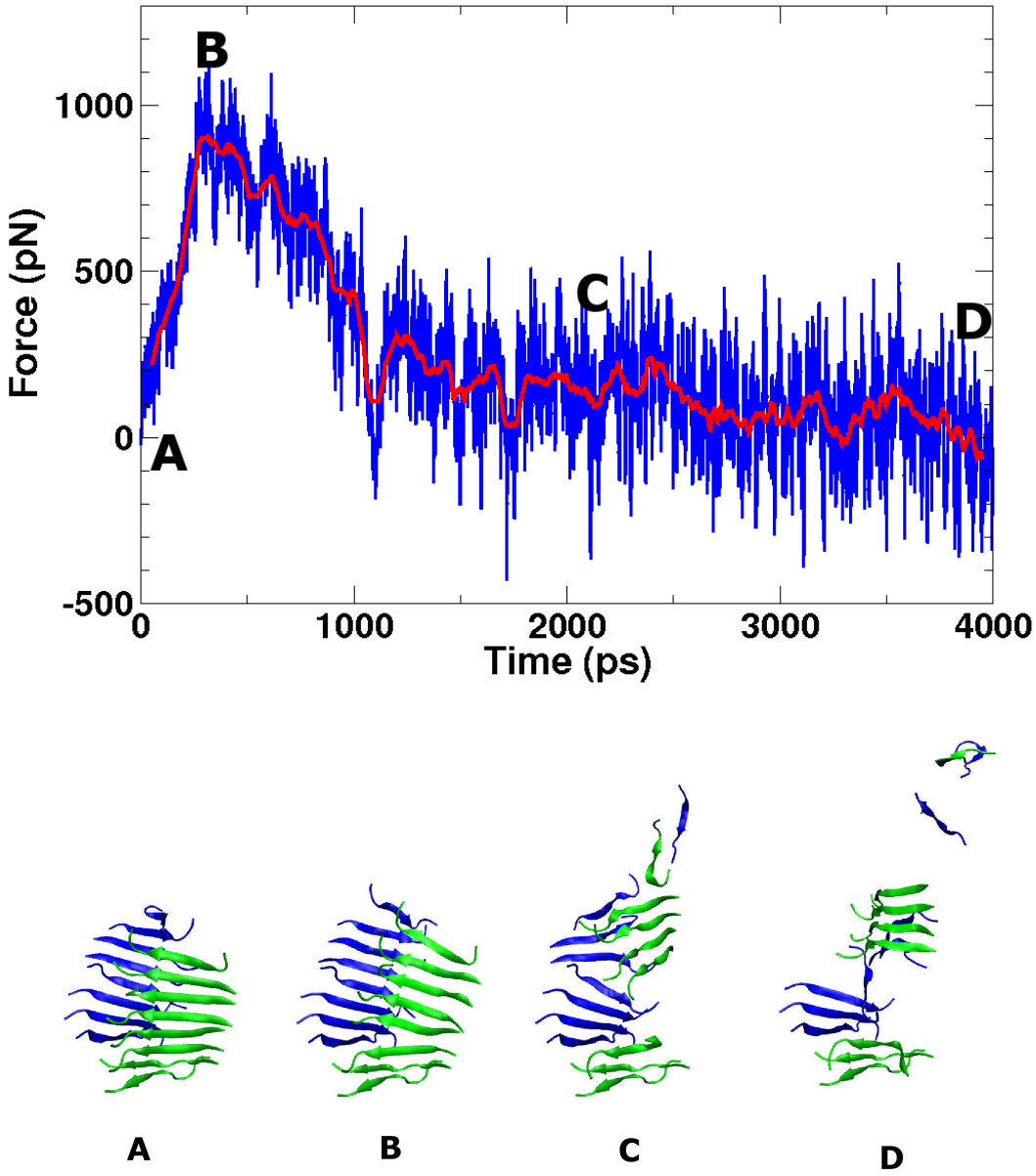

Figure 3.8: **STRETCH SMD:** Force profile obtained during the Stretch SMD pulling geometry for I7V fibril model trajectory. Points marked A, B, C and D on the plots correspond to the equivalently marked representative pictorial snapshots (below graph) of the fibrils evolving in time as the force is applied.

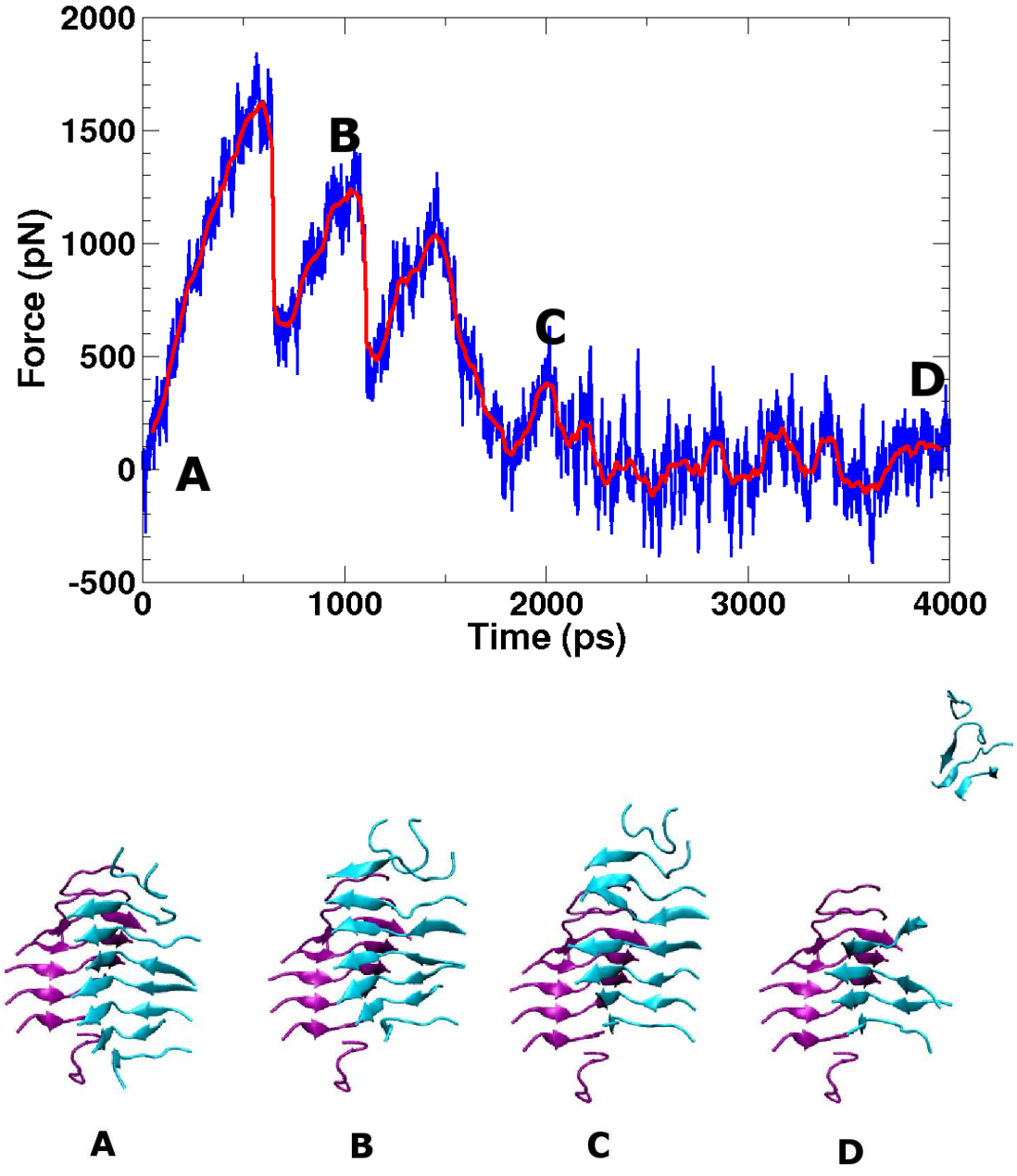

Figure 3.9: **SLIDE SMD:** Force profile obtained during the Slide SMD pulling geometry for F4L-A6P fibril model trajectory. Points marked A, B, C and D on the plots correspond to the equivalently marked representative pictorial snapshots (below graph) of the fibrils evolving in time as the force is applied.

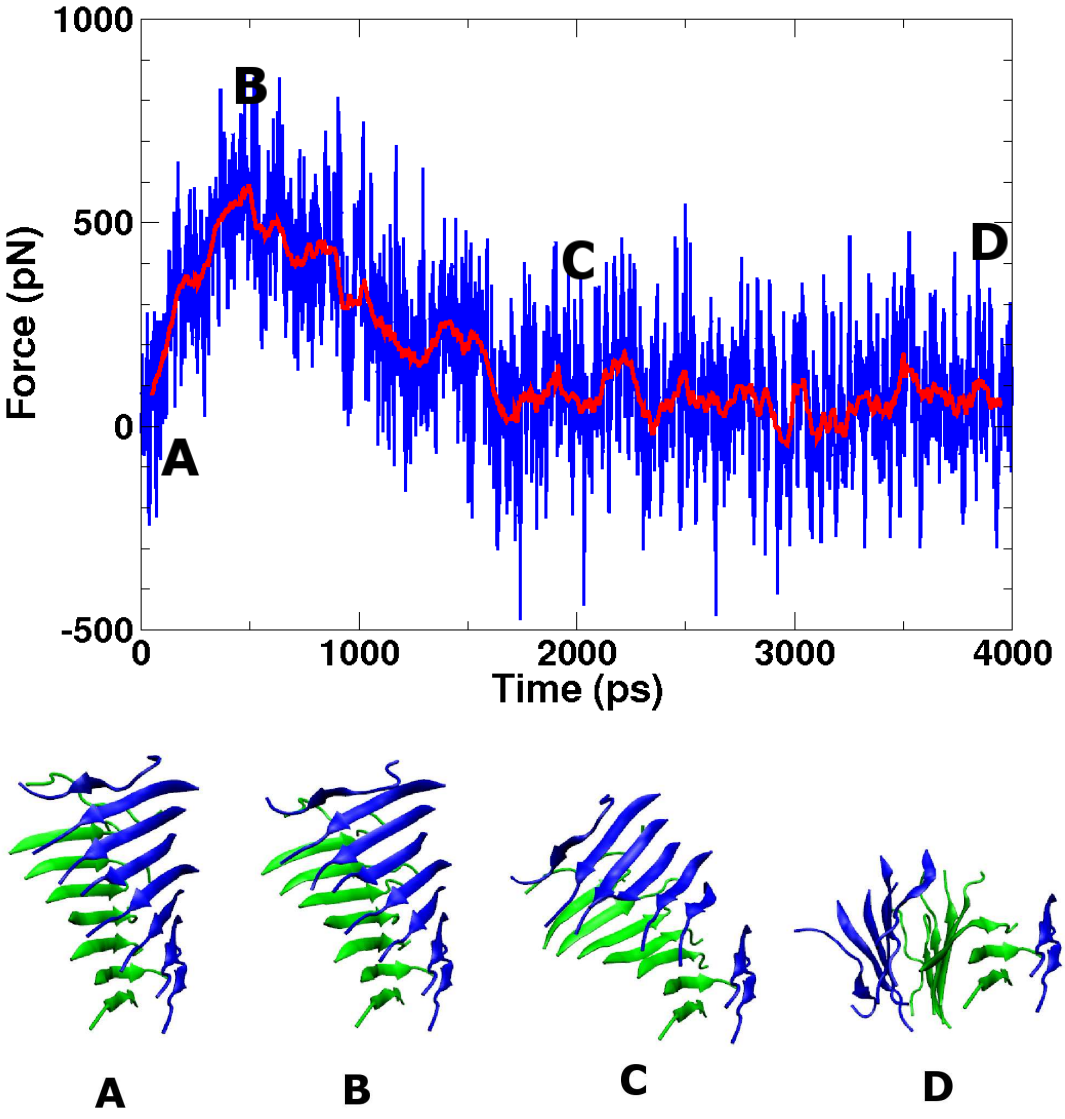

Figure 3.10: **SHEAR SMD:** Force profile obtained during the Shear SMD pulling geometry for Wild-type hIAPP$_{20-29}$ fibril model trajectory. Points marked A, B, C and D on the plots correspond to the equivalently marked representative pictorial snapshots (below graph) of the fibrils evolving in time as the force is applied.

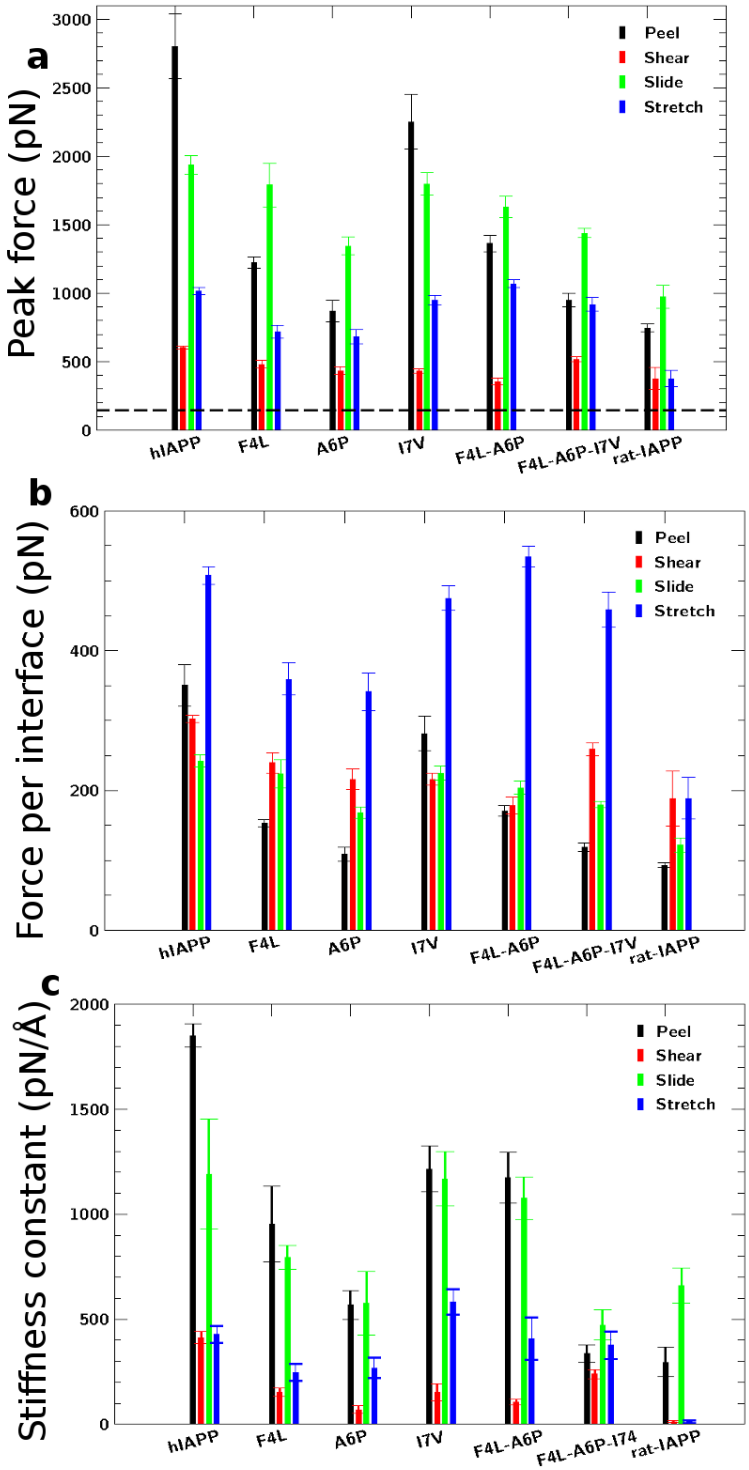

Figure 3.11: (a) Summary of the mean peak forces recorded for the seven fibril models during each type of fragmentation simulation, extracted post noise filtering. The dashed line is the maximum filtered noise signal level. Error bars are the standard error in the mean peak force from four repeat simulations. (b) Normalised peak force values in terms of number of interaction interfaces disrupted during the deformation (two for stretch and shear, eight for peel and slide). (c) Directional stiffness constants extracted from the gradient of the force versus displacement graphs in the linear region (before exertion of the peak force).

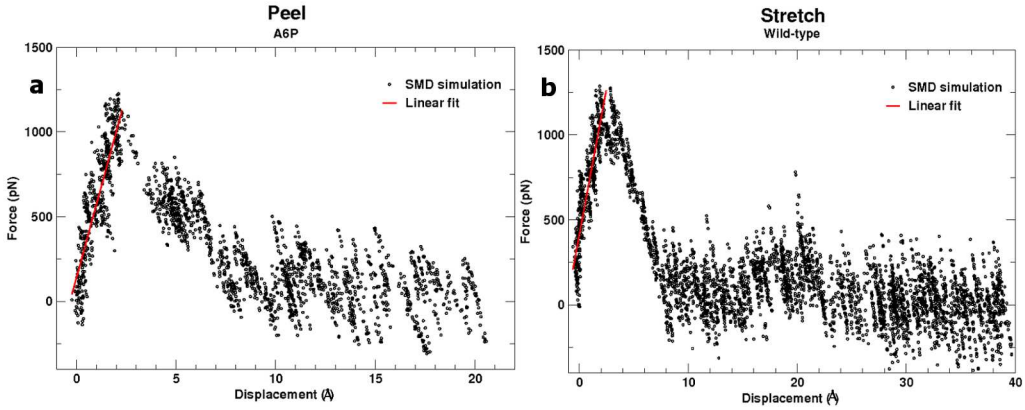

Figure 3.12: (a) and (b) show a representative selection of force vs displacement plots for peel and stretch SMD simulations respectively from which mechanical stiffness constants are computed. Black dots are simulation data points whilst the red line is a linear fit in the elastic regime whose gradient is used to characterise the stiffness (in pN/Å) of the fibril prior to failure. (Displacement in this case refers to the change in distance between the centres of mass of pulled and fixed atom pairs relevant to each SMD pulling mode.

filtering are shown in figure 3.11a. The mean peak force normalised to the number of interaction interfaces interrogated by each pulling mode are shown figure 3.11b. The directional stiffness constants for each model, measured for the four different SMD pulling directions, are shown in figure 3.11c. These directional stiffness constants were obtained from the gradient of the force versus displacement plots measured in the linear response regime (fig 3.12 is representative example) and averaged over the repeat simulations. Mean peak forces measured for the disordered rIAPP aggregate are the consistently the lowest for all deformations (fig 3.11a). The overall mechanical strength of each aggregate is governed by the weakest force required to induce structural failure in any direction. Using this criterion, the wild-type fibrils are the strongest, whereas the A6P and the rIAPP variants offer the least resistance to an applied force. Nevertheless, the disordered aggregates consistently have mechanical resistance above the levels of thermal noise due to the presence of van der Waals (table 3.2) and hydrogen-bonding interactions (fig 3.5b) between their peptide strands.

The peak forces required to fragment the model fibrils appear to depend on both the peptide sequence and the directionality of the deformation applied during the SMD. Stiffness constants measured in the linear response regime before the application of the peak force

are also sequence-dependent and anisotropic. Single molecule experiments and SMD simulations to probe the fracture forces necessary to disrupt the structure of folded proteins have previously shown that the mechanical resistance is dependent on the pulling geometry (103; 104; 105). The results presented here show that this is also the case in SMD simulations of peptide assemblies. This is consistent with the observation that there is no direct correlation between the enthalpic stability of a model fibril and its mechanical resistance *in silico*, as can be seen from comparing fig 3.6 and fig 3.11, because fibril fragmentation is a non-equilibrium process during the SMD calculations.

3.5.2 Role of inter-atomic interactions

Hydrogen bond interactions:

When the results of the SMD pulling modes are normalised in terms of the number of interfaces interrogated, all ordered fibrillar aggregates are most resistant to the stretch deformation (fig 3.11b). The stretch SMD mode directly interrogates the strength of the inter-strand hydrogen bonds along the long axis of the fibril that are primarily responsible for the general stability of amyloid fibrils. By comparing the peak forces per interaction interface to the secondary structure content it is clear that the resistance of fibrils to stretch is higher when the aggregate has greater β-sheet content and there are more inter-strand hydrogen bonds. The sequences of the models were chosen to minimise differences in side-chain hydrogen-bonding thus these are similar for all variants (with the exception of rat-IAPP & A6P). Consequently, the side-chain interactions play a relatively minor role in determining the sequence dependence of the resistance to the stretch pulling mode for this choice of variants. Fibrils are consistently less resistant to the shear deformation than to stretch as the shear SMD mode places stress on the hydrogen bond network perpendicular to the long axis of the fibril rather than along its length. This implies that the hydrogen bonding network provides a cooperative resistance to forces applied along the fibril.

Hydrophobic-core interactions:

The peel and slide deformations primarily address the hydrophobic core inter-sheet interactions within the fibrils. The peel deformation, in which the β-sheets are forced apart thus exposing the hydrophobic core to solvent molecules, was responsible for the two largest peak forces measured by the SMD simulations (2804.33 $\pm$ 237.15 pN for the wild-type

and 2253.88 ±201.26 pN for I7V). As might be expected, the relative response to peel and slide probing was particularly sequence-dependent since it depends on the molecular details of the packing within the hydrophobic core. Figure 3.13 and figure 3.14 both show a detailed analysis of the energetic changes experienced by the fibrils as they fracture during the course of peel and slide SMD simulations which mirror the distinct force profiles.

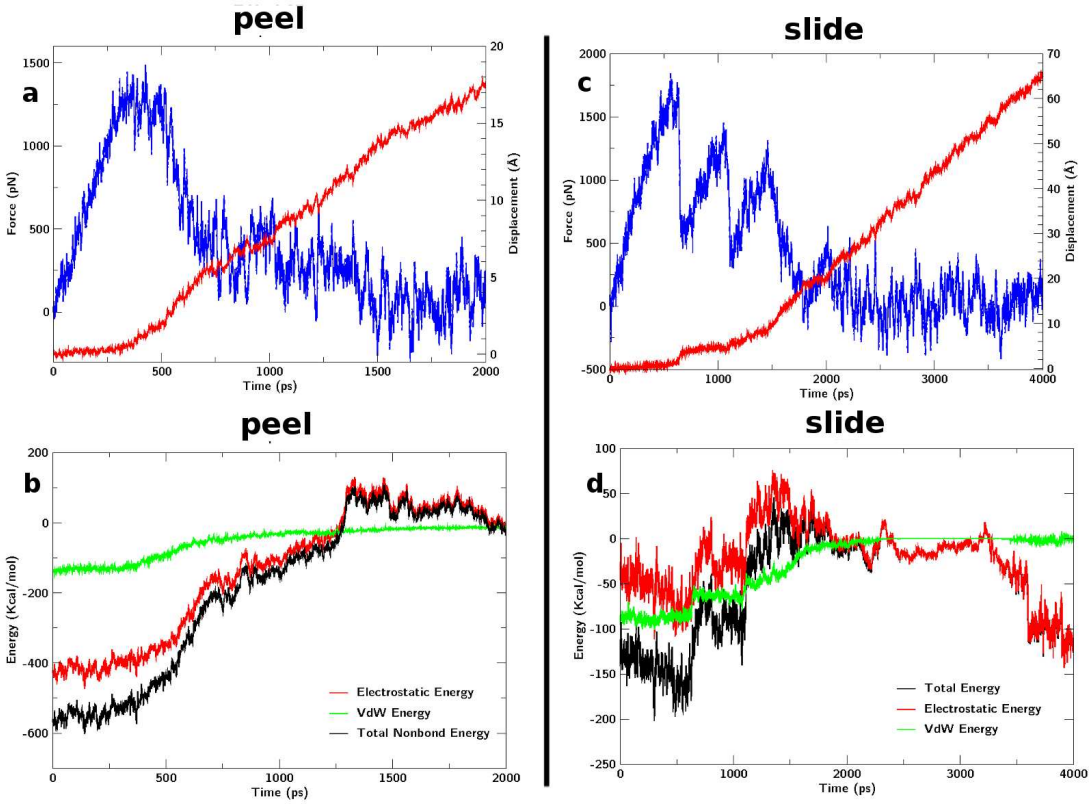

Figure 3.13: The energies associated with electrostatic and van der Waals interactions between detached fragments and the rest of the fibril as the distance between them increases calculated from two representative SMD simulations. (a) and (b) are for a peel simulation on the F4L-A6P variant and (c) and (d) refer to the slide SMD on F4L-A6P. The displacement in the plots (shown in red in panels a and c) indicates the change in separation between the centres of mass of the fragments and the remainder of the fibril.

3.5.3 Role of aromatic side-chains

Although the F4L variant has similar secondary structure content and backbone hydrogen bonds to the wild-type fibrils (fig 3.5), it records a significantly smaller relative mean peak force with a reduced mechanical stiffness constant (fig 3.11) in the peel simulations. The

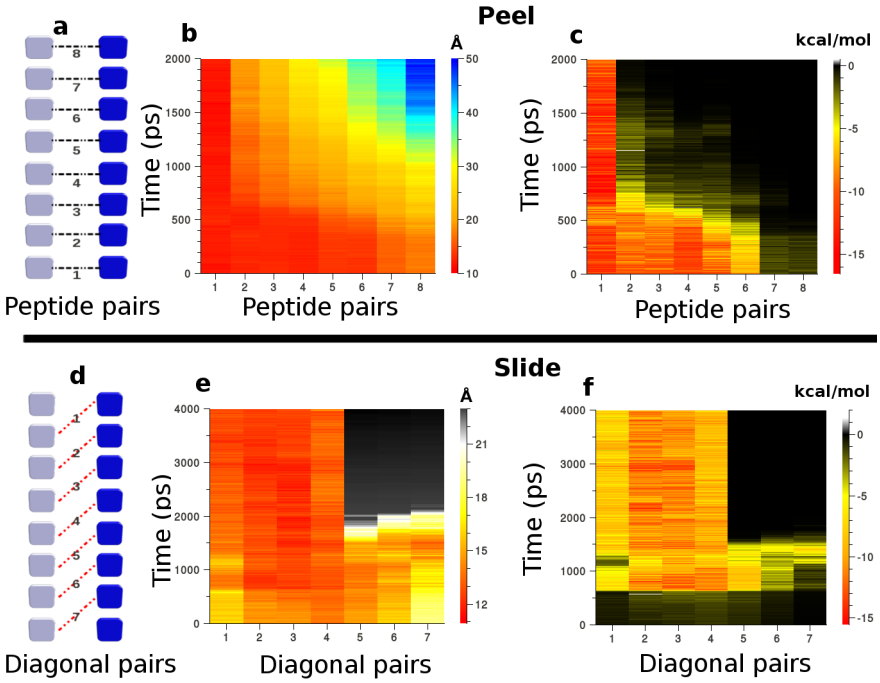

Figure 3.14: The separation distance and van der Waals interactions between individual pairs of peptides from opposing sheets during peel and slide SMD simulations for the F4L-A6P variant. **Peel:** (a) Schematic representation of peptide pair interactions monitored during peel SMD. (b) Variation of the distance between the centre of mass of the peptide pairs during peel SMD. (c) Inter-peptide vand er Waal interactions during peel SMD. **Slide:** (d) Schematic representation of the interactions monitored from peptides on opposing β-sheets during the slide SMD. (e) Variation of the distance between the centre of mass of the peptide pairs during slide SMD. Grey-to-black regions indicate large distances (from 24 to 70 Å). (f) Inter-peptide van der Waal interactions during slide SMD.

reasons for the differences in mechanical response of the wild-type sequence relative to F4L and all other models for the peel SMD mode can be traced to inter-sheet Phe-Phe interactions. A comparison of the peel trajectories of the wild-type and F4L fibrils (along with the other variants) reveals that the presence and orientation of Phe residues on opposite β-sheets plays a key role in determining the observed mechanical response. In the wild-type sequence, the Phe residues from opposite sheets orient themselves in such a way that they are directly interacting with each other (fig 3.15a and b). The interdigitation of these residues has an overall stabilising effect which confers the additional mechanical strength to the model fibrils as observed in the peel mean peak forces. During the course of the peel simulations, the pairs of Phe-Phe break one by one in a zipper like manner. In the slide simulations, the interdigitated Phe-Phe pairs aid in resisting the sliding motion of the pair

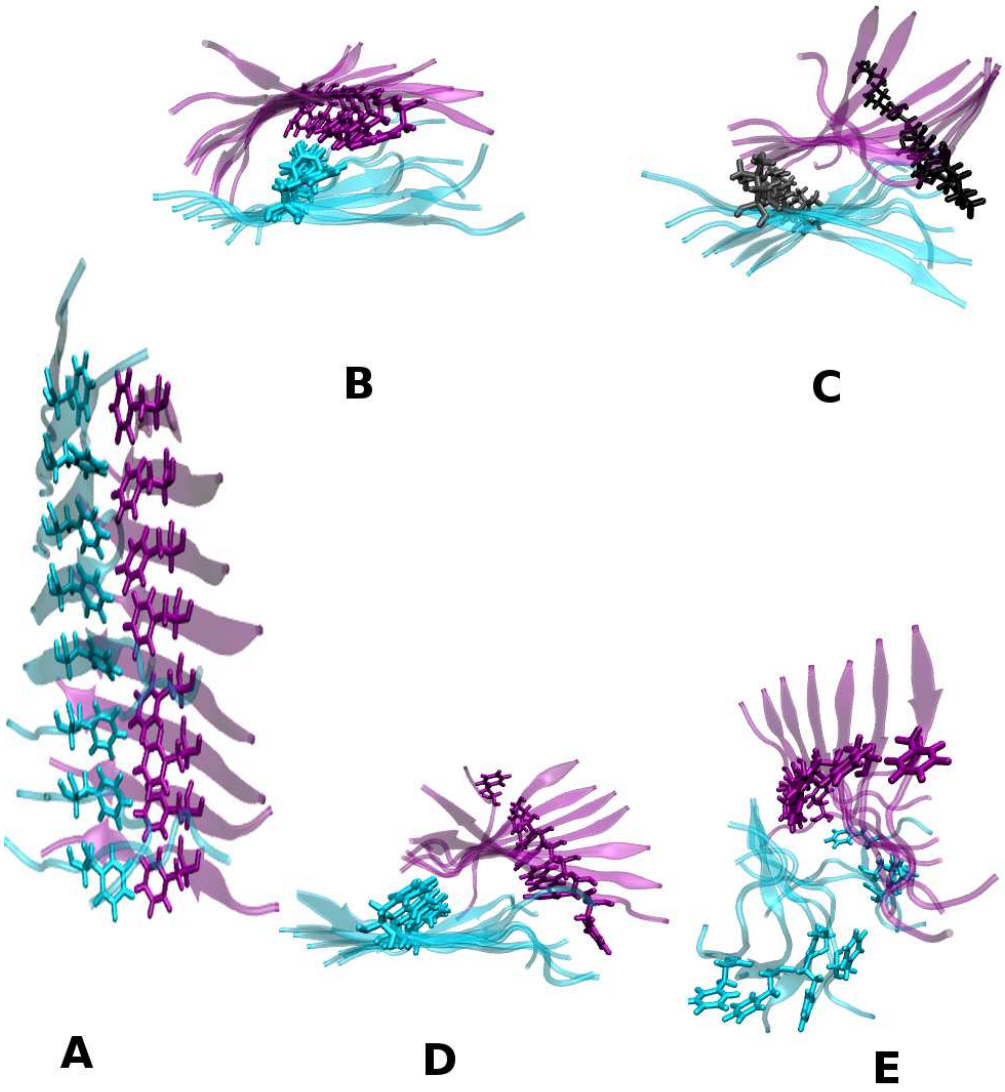

Figure 3.15: Orientation and alignment of Phe residues on opposing β-sheets. (a) Side and (b) top view for wildtype. Top view for (c) F4L (replacement Leu shown in black and gray), (d) I7V, and (e) A6P variants.

of β-sheets. In contrast, the reduced mechanical resistance to peel observed for all of the variants can be attributed to the absence of this arrangement of inter-sheet Phe residues, as shown in figures 3.15c, d, and e. For the F4L variant, the Phe substituted residues are significantly further apart (17.75 $\pm$ 0.34 Å between the center of masses of Leu pairs on opposite β-sheets relative to 9.70 $\pm$ 0.14 Å for corresponding Phe pairs in the wild-type). Similarly, in the I7V variant, which undergoes a shift of one β-sheet relative to the other, the Phe pairs on opposite sheets are much further apart (19.56 $\pm$ 0.63 Å) before SMD. For A6P, which has reduced β-sheet content relative to the wild-type, structural disorder in the model fibril gives rise to a random orientation of the Phe residues. The reduced mechanical resistance to peel observed for these variants is therefore attributed to the absence of close packed inter-sheet Phe residues in the model fibrils.

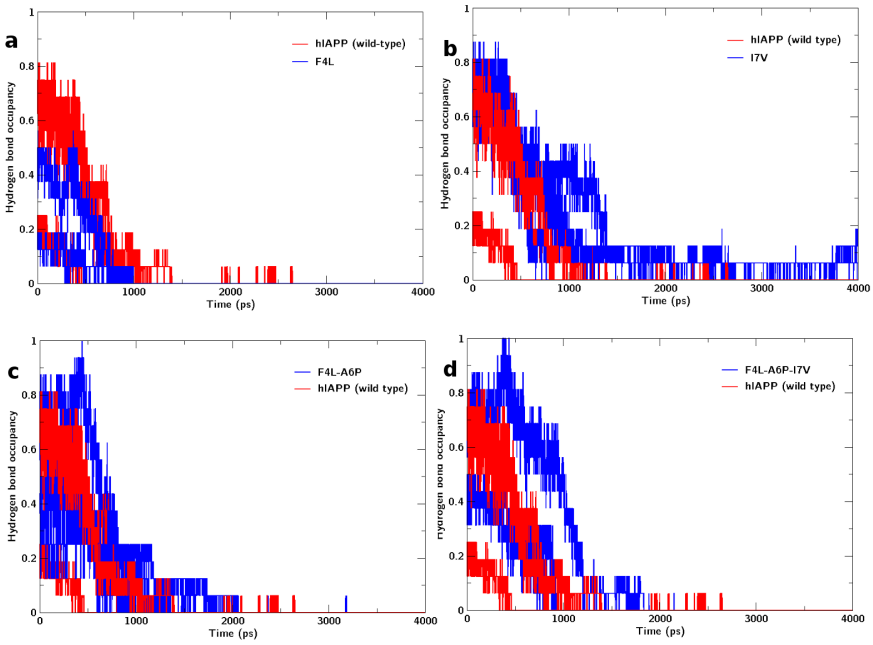

Figure 3.16: Backbone hydrogen bond occupancies between peptides at the fracture point of the fibrils during the stretch SMD. In each panel, the wild-type fibril (red) is compared with a selected variant (blue) for the four independent simulations. (a) F4L has an unexpectedly weaker response to force due fewer hydrogen bonds at the fracture site corresponding to occupancy of 0.2-0.4 whilst the wild-type peaks at 0.8. Plots b, c and d reveal the origin of similarity in peak force between the wild-type, I7V, F4L-A6P and F4L-A6P-I7V

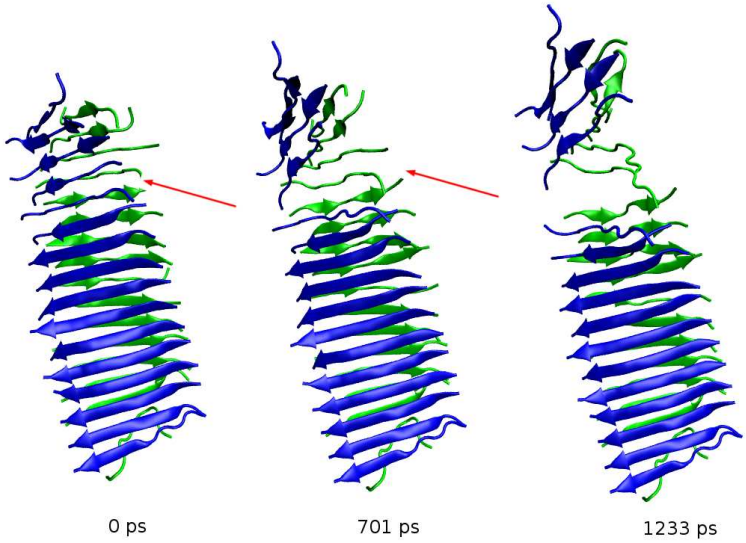

Figure 3.17: A defect site in the fibril structure, indicated by the red arrow, in which a randomly coiled peptide present within one β-sheet (shown in blue) acted as the primary fracture point after 700 ps of SMD using the stretch deformation mode. The second red arrow at 700 ps shows a subsequent defect in the remaining β-sheet (in green) which then lead to mechanical failure of the second β-sheet after 1223 ps of SMD, and ultimately the fragmentation of the fibril

3.5.4 Role of structural defects & weak points

All ordered fibrillar aggregates demonstrated the greatest resistance per interface to the stretch deformation. The wild-type, I7V, F4L-A6P, and F4L-A6P-I7V fibrils display mean peak forces under stretch of $\approx$ 1000 pN. However, the mean peak force for the F4L, A6P, and rIAPP variants is significantly lower (719.20 $\pm$ 46.39pN, 682.95 $\pm$ 53.65pN, and 377.79 $\pm$ 59.96pN, respectively) with correspondingly lower stiffness constants. The disruptive effect of the proline substitution in the A6P and rIAPP variants results in aggregates containing a higher proportion of disordered strands (63% and 78%, respectively, compared to only 34% for the wild-type), indicating that there are numerous positions where the chain of inter-β-sheet hydrogen-bonding interactions are interrupted (fig 3.5) i.e defects in the β-sheet. The reduced mechanical resistance of aggregates constructed from these sequences can be explained by the increased number of defects present in these structures. The defects act as fracture points that are the first to undergo mechanical failure under stress.

The importance of these defects is emphasized by the anomalously low mean stretch force

recorded for the F4L variant. Although aggregates of this sequence possess equivalent secondary structure content to the ordered wild-type fibrils, the SMD simulations performed start from a structure containing a defect in which a single peptide strand in a random coil conformation interrupts an otherwise ordered β-sheet (fig 3.16). This defect then acts as a weak spot in the fibril making it undergo mechanical failure at comparatively low forces.

To test the hypothesis that the defects can be the cause of structural failure in fibrils, a longer model (16x2) of the wild-type sequence was constructed containing a defect in the form of a random coiled β-strand as shown in figure 3.17. This was then subjected to the stretch SMD in an equivalent manner to the shorter (8x2) fibrils. As expected, the defect acted as the fracture point in every one of the four repeat simulations. These simulations indicate that the mechanical resistance of fibrils depends not only upon peptide sequence and the relative arrangements of the individual β-strands within the fibrils, but also upon the number of flaws in the quasi-crystalline peptide array. Consequently, growth conditions that decrease the number of defects, for example, if amyloid formation takes place very slowly under conditions that are close to equilibrium may result in fibrils that would be expected to be more resistant to mechanical stress than those grown so rapidly that numerous defects are annealed into their structure.

3.6 Conclusions

The results of this study reveal that the peptide sequence not only plays a role in the energetic stability of the hIAPP$_{20-29}$ derived proto-fibrils but can also be a major contributor, in some directions, to the the observed mechanical properties as well. The recorded mechanical strength of the fibrils is shown to depend upon the detailed choreography of the pathway that the fibrils are forced to take by the applied force, and not just on their equilibrium thermodynamic stability. The main findings of this chapter and their wider implications are summarised below.

Induced Structural effects:

The introduction of the substituted residues had varying effects on the on the overall structural properties of the fibrils relative to the original ssNMR conformation. At the conclusion of $\approx$ 22 ns of MD in the absence of mechanical forces, the wild-type, F4L and

I7V variants retained a high degree of β-sheet content and remained as ordered fibrillar arrays while the A6P, F4L-A6P, and F4L-A6P-I7V variants had increased levels of disorder relative to the wild-type. The Rat-IAPP variant became significantly disordered, comprised primary of randomly coiled peptide strands. These structural changes in the simulated models were consistent with available experimental observations on the morphology of fibrils formed from the different sequences as well the non-amyloidogenicity of the rat sequence.

Fibril response to mechanical probing:

The four SMD deformations applied to each fibril showed that the mechanical response is anisotropic. All aggregates (except the disordered rIAPP) showed the greatest mechanical resistance per interface to the stretch deformation, which directly tests the strength of the β-sheet hydrogen-bonding interactions parallel to the long axis of the fibril. Whereas stretch primarily interrogates the β-sheet hydrogen-bonding network common to all fibrils, the relative strengths of the other interfaces varied with sequence as these depend on the molecular details of packing within the hydrophobic core. Consequently, the contribution of sequence specific interactions, such as Phe-Phe interdigitation, can be significant in determining the mechanical response of amyloid fibrils in those directions. In force spectroscopy experiments that measure only the local response (such as AFM nano-indentation), care must be taken when reporting elastic moduli of amyloid fibrils as the results strongly depend both the direction and precise manner in which the force is applied.

Significance of structural defects in the fibrils:

Increased disorder coupled with a reduction in backbone inter-peptide hydrogen bonding lead to mechanically weak regions developing in some of the aggregates. These regions acted as structural defects within the fibril β-sheets, prone to mechanical failure, by consistently affecting their ability to withstand the applied forces in the stretch mode. The observation that defects are able to dominate mechanical properties is consistent with the extensive studies on inorganic materials such as metals and ceramics that highlight the role of dislocations, point and line defects in structural failure (106). The presence of defects implies that the mechanical properties of fibrils may also depend on the physical conditions under which they are formed. Slow nucleation and growth that takes place close to equilibrium conditions for instance would be expected to result in fibrils containing fewer defects

than those that have grown rapidly, and would therefore produce fibrils able to resist higher mechanical stress. It has been suggested that the preferential fracture of fibrils at defect locations can give rise to a self-healing mechanism that is responsible for the high levels of structural order present in fibrils (58), a hypothesis that is supported by the results of SMD simulations presented in this chapter. Moreover, a prion strain that has a molecular structure capable of accommodating numerous defects may well be inherently more promiscuous than one formed from fibrils with a unique and high regular quasi-crystalline structure. This would provide another molecular basis, in addition to seeding, for the observation that frangible prions have greater infectivity (107).

Implications for seeding events and cytotoxicity:

Understanding the role of the amino-acid sequence on the fibrils mechanical properties may also be of therapeutic importance. *In vivo* genetic mutations may lead to the formation of mechanically less stable fibrils that are prone to fragmentation. Brittle fibrils would not only potentially increase the likelihood of early onset of amyloid-related disease through increased seeding but may also increase toxicity because the shorter fragments have been shown to disrupt cell membranes and can have enhanced cytotoxicity (28; 32). Early onset type-II diabetes caused by a mis-sense mutation in the amylin gene within the Japanese population results in a change of the first serine residue to a glycine residue in the hIAPP$_{20-29}$ region i.e SNNFGAILSS becomes GNNFGAILSS (98). It has been reported that this amylin variant shows increased fibrillisation kinetics and enhanced cytotoxicity relative to wild-type amylin (99). Although this particular mutation was not explicitly explored in this study, nonetheless the SMD simulation results suggest that a plausible mechanism linking enhanced fibrillisation to seeding and to cytotoxicity could be changes in the mechanical properties resulting from the sequence substitution. SMD simulations such as those described here could provide useful insights to aid therapeutic design strategies for amyloid related diseases.

Implications for design of amyloid based nano-materials:

Knowledge of the factors governing the mechanical performance of amyloid or amyloid-like fibrils will be useful in the design and construction of nano-materials based on the self-assembly properties of amyloid. It may be desirable to engineer defects into self-assembling

systems to produce a material that is more malleable at the macroscopic level. This could be achieved, for example, by doping with a low concentration of synthetic peptides containing one or more amino-acid substitutions incapable of forming conventional hydrogen-bonding interactions within a β-sheet, or by mutating the dopant so it includes a disruptive proline residue, or by introducing a substitution designed to disrupt electrostatic interactions within the mature fibril. This strategy for nano-material design has already been used in computer simulations of tubular nanostructures formed from β-helical protein motifs that have shown that substitutions with synthetic amino acids (such as 1-aminocyclopentane-1-carboxylic acid) can alter their structural stability (108). Steered molecular dynamics simulations, such as those reported here, could well prove invaluable in the bottom-up design of novel nanomaterials.

Chapter 4

Modulating mechanical failure in polymorphic fibril arrangements

This chapter is based on a recently accepted paper due to appear for publication in the Beilstein Journal of Nanotechnology in 2013.

4.1 Overview

Synopsis

This chapter explores the influence of the steric packing of beta peptides within amyloid fibrils, due to polymorphism, on their mechanical properties. Three fibril polymorphs of the SNNFGAILSS peptide are mechanically probed with aim of determining ways in which their packing arrangements can be manipulated and exploited in nanotechnology applications as means of modulation of fibril strength. To this end, the effects of packing, fibril length and terminal capping are investigated with a view to potentially guide design strategies for nano-materials construction. Polymorphism and capping offer two attractive parameters that could be used to control the robustness of amyloid inspired nano-materials. A favoured polymorphic state for instance could be obtained from altering the growth environment whilst terminal capping is a simple step during standard peptide synthesis. Previous results from chapter 3 highlighted the importance of structural defects within the fibrils in determining their mechanical properties. Here, particular focus is given to the role played by such defects in the ability of the three polymorphs to resist an applied force. A hierarchy of factors governing the mechanical resilience of defect containing fibrils is observed.

Introduction

Amyloid fibrils, like many crystalline materials, exhibit polymorphism whereby a peptide sequence is able to form a variety of fibril structures with distinct features. The predominant polymorph obtained by the protein assembly depends on the environmental growth conditions such as pH, temperature, salt concentration and mechanical agitation (109). Since amyloid polymorphs have been observed with drastically different morphologies (110) and chemical properties (111), it is important to develop an understanding of how the polymorphic form influences the mechanical properties of fibrils. A wealth of information on the material properties of amyloid is already available from extensive pathological and biological studies that focus on the diseases aspect of amyloid. Over recent years however, intrinsic properties such as self-assembly, robustness and degradation resistance have

increasingly made amyloid fibrils attractive candidates for use in nano-technological applications (41; 42) that range from conducting nano-wires (43), drug delivery devices (44), structural scaffolds (46; 47) and functionalised hydrogels (48). A central theme in each of these distinct potential applications is an ability to control and modulate a desired property of the fibril aggregates.

The requirements for the mechanical robustness of ideal long conducting nano-wires for instance, is that they not be prone to fragmentation whereas a drug delivery device needs to be sufficiently robust to carry its cargo to the target site, but then be able to release it in response to an external signal. Amyloid fibrils are suited to property switching since a change in solvent pH or temperature can unambiguously favour one particular polymorphic state over another. If the mechanical characteristics of the fibrils change significantly between the two states, this could be then exploited as a simple strategy for modulation in some nano-technological applications. There is a need however, to understand how the arrangement of the individual β-sheets modulates the mechanical behaviour before fibrils with bespoke material properties can be designed.

In this chapter, three polymorphs of fibrils formed from the SNNFGAILSS fragments of the amylin protein (KCNTATCATQRLANFLVHSNNFGAILSSTNVGSNTY) as structurally determined by ssNMR are simulated in full atomistic detail. The SNNFGAILSS sequence is particularly interesting in that both parallel and anti-parallel polymorphs are simultaneously observed under identical growth conditions (16). Moreover, a separate ssNMR study only observed a single fibril type in the anti-parallel configuration, possibly due to the use of different terminal capping groups (22). Consequently, the differences in energetic and mechanical stability between polymorphs of SNNFGAILSS present a unique system to study the relevant interactions that play key roles in determining their observed properties.
The focus of this chapter is to investigate how the packing arrangements found in three different polymorph symmetry classes influences fibril mechanical behaviour. The fibrils are probed from different directions by use of four distinct pulling geometries. These SMD pulling geometries are designed to either disrupt the stabilising hydrophobic core or backbone hydrogen bond networks from a variety of directions. A secondary investigation to evaluate the efficacy simple ways to modulate the mechanical characteristics of the polymorphs is also carried out. In particular, assessments of how the mechanical properties are

attenuated by incorporating chemical capping groups to neutralise the N and C-termini of the peptides and whether the relative mechanical response is affected by changes to the length of the model fibrils. Attenuation in mechanical robustness with increasing length could, for instance, be important if viable long conducting nano-wires are to be created. The ability to control the length of fibrils by use of simple alterations in growth and storage conditions has been successfully demonstrated for bovine insulin fibrils (112). This could allow for the fibril length to become be a tunable parameter to influence the desired mechanical properties in nano-materials design strategies, if a length dependency exists. The comparative robustness might not only be different for the longer fibrils, but the stability ranking between the associated polymorphs might also be altered.

4.2 Simulation details

Construction of polymorph models

Two coordinate files of parallel and anti-parallel SNNFGAILSS polymorphs as determined by ssNMR experiments (16) were used to construct fibril models for simulation. A third additional fibril structure was rationally designed in a parallel polymorphic stacking arrangement. The three models are classified according to symmetry packing after the Eisenberg steric zipper nomenclature as Class1-P, Class2-P and Class6-AP where 'P' & 'AP' denote parallel and anti-parallel ladders of peptides in a β-sheet respectively. The peptide arrangements for each of these steric zippers classes are schematically shown in figure 1.5. The ssNMR coordinate files initially consisted of a pair of β-sheets, each of which was composed of two peptides of the SNNFGAILSS sequence. These coordinates were used as templates from which longer fibrils were constructed. The explicit details of the construction and refinement of these models are discussed in § 2.3.1. Two fibrils sizes were constructed for each polymorph namely a pair of β-sheets each containing eight peptides (8x2 model) and a pair of β-sheets each containing 16 peptides (16x2 model). Two versions of the different sized polymorphs were constructed with the peptide terminal ends in either the zwitter-ionic form or neutralised with N-terminal acetylation & C-terminal amidation capping groups.

Molecular Dynamics simulation

In order to evaluate the thermodynamical and structural properties of the polymorph models, unrestrained MD simulation was run using the multi-stage protocol outlined in § 2.5.1. A

Polymorph	Size	N-C Terminus	Amberff99SB	Charmm22-cmap
Class1-P	8x2	free	22ns	10ns
	8x2	capped	25ns	32ns
	16x2	free	40ns	10ns
	16x2	capped	40ns	30ns
Class2-P	8x2	free	22ns	10ns
	8x2	capped	25ns	32ns
	16x2	free	40ns	10ns
	16x2	capped	40ns	30ns
Class6-AP	8x2	free	22ns	10ns
	8x2	capped	25ns	32ns
	16x2	free	40ns	10ns
	16x2	capped	40ns	30ns

Table 4.1: Details of the molecular dynamics simulations run for the three polymorphs with different lengths and terminal capping groups. MD was run with the Amberff99sb forcefield and the Charmm22-cmap forcefield used for validation. The duration of each simulation is listed in nano-seconds.

summary of all the production MD and force-field validation simulations is presented in table 4.1. The root mean square deviation (RMSD) of backbone carbon-α atoms was used to monitor convergence of the MD simulations. This was achieved within 20 to 30 ns for the 8x2 and 16x2 models respectively.

Steered Molecular Dynamics simulation

The SMD protocol for mechanical characterisation of each polymorph is described in detail in § 2.4.1. The final configurations of the fibril models at the end of the MD were used as the starting points for SMD simulations with each model re-solvated in a larger water box to enable extension under force with minimal self interactions. A summary of all the mechanical probing simulations is given in table B.1. The parameters for each pulling mode were as stated previuosly, with the exception of the capped models which required the use of a time-step of 0.5 femto-secs and pulling velocity 0.04 Å/ps.

4.3 Results: Structural properties and thermodynamics

Prior to the mechanical probing SMD simulations, an evaluation of how the packing arrangements of the peptides in the three polymorphs influence the structural and thermodynamic

properties in each fibril was made. A clear understanding of the relevant interactions imposed by the steric zipper classes is vital for the interpretation of the relative mechanical responses amongst the polymorphs. An additional assessment of the effect of the terminal capping groups on these structural and thermodynamic characteristics was also carried out.

Structural comparisons

The various configurations adopted by the fibril models and their structural properties at the conclusion of the unrestrained MD show different trends between the three polymorphs. However, the anti-parallel Class6-AP model is consistently more structurally ordered than either of the two parallel counterparts irregardless of differences in fibril length, capping groups or force-field used. The features of the final configurations of the 16x2 fibrils are depicted in figure 4.1. These visual features are quantitatively expressed in terms of secondary structure content and mean hydrogen bond numbers in figure 4.2. For these measured quantities, the ranking amongst the polymorphs places Class6-AP higher than Class1-P which is in turn higher than Class2-P.

Thermodynamic stability & electrostatics

The enthalpic stability of the polymorphs was assessed through the MM-PBSA method. Table 4.2 lists the individual terms contributing to the enthalpy for both capped & uncapped 16x2 fibrils. The enthalpy differences show that the Class6-AP polymorph is the most thermodynamically stable whilst class2-P is the least for the uncapped fibrils. The trends in the thermodynamic ranking of the uncapped models are consistent with the structural features found between them. Class6-AP for instance has the most amount of beta-strand content and number of stabilising hydrogen bonds (figure 4.2). The introduction of terminal caps however, seems to reduce the enthalpy differences between the polymorphs. The differences between class6-AP and class2-P fibrils become vanishingly small despite huge differences in both structural order and hydrogen bonding occupancies. The effect of the capping groups on the overall stability of the fibrils can be understood from the changes in the the electrostatic character they induce since they protect otherwise solvent exposed charged terminus ends. These capping groups have the ability to significantly alter the steric packing of the hydrophobic core (fig B.1) and hence the overall electrostatics of the fibrils.

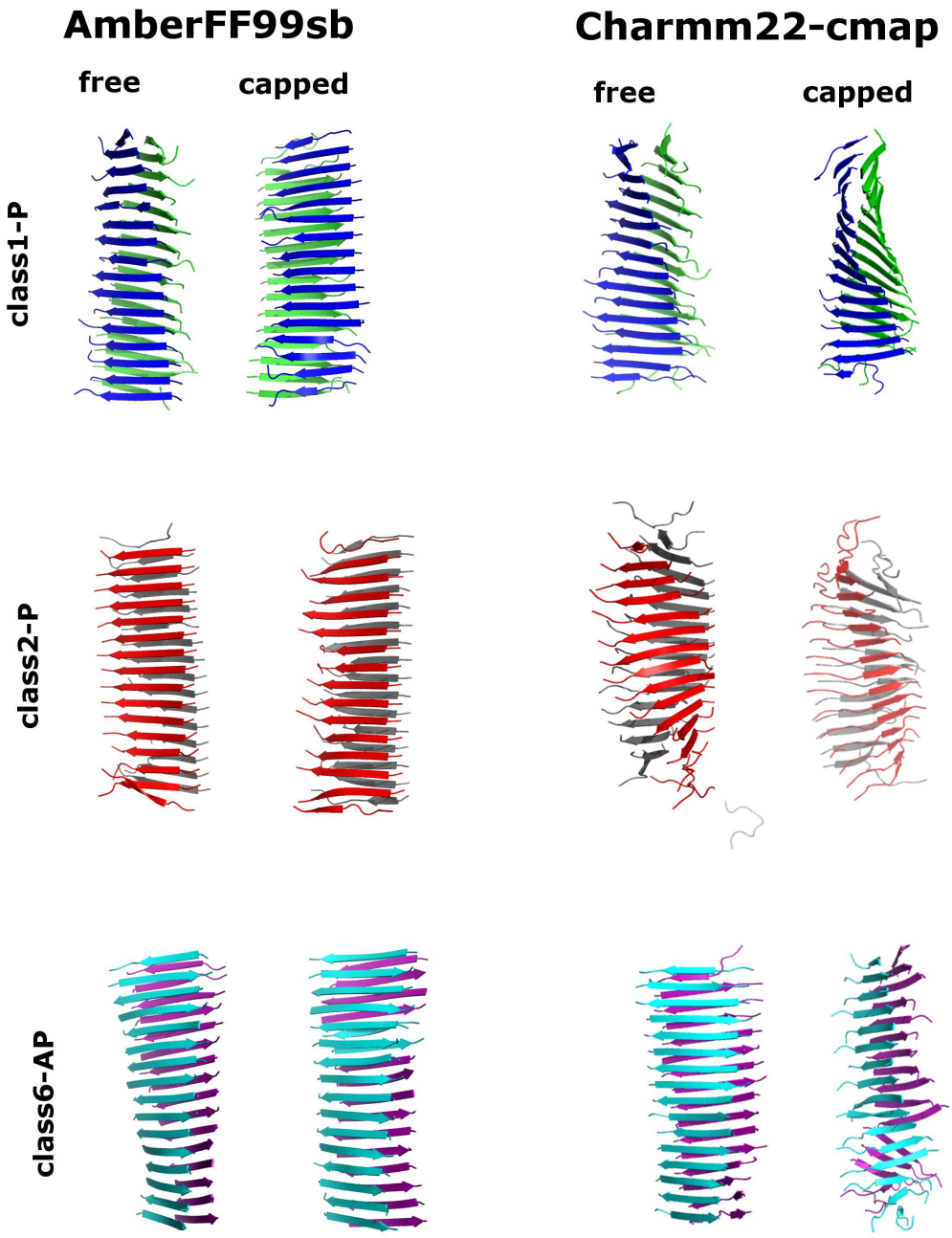

Figure 4.1: Final configurations obtained of the 16x2 fibrils models at the end of unrestrained MD using AmberFF99sb & Charmm22-cmap forcefields. No significant structural differences are observed for the free-terminus models run in the two forcefields. The capped models run in Charmm22-cmap however show increased twisting and disorder relative to those run in AmberFF99sb, with class2-P having more random coiled potions in its peptides

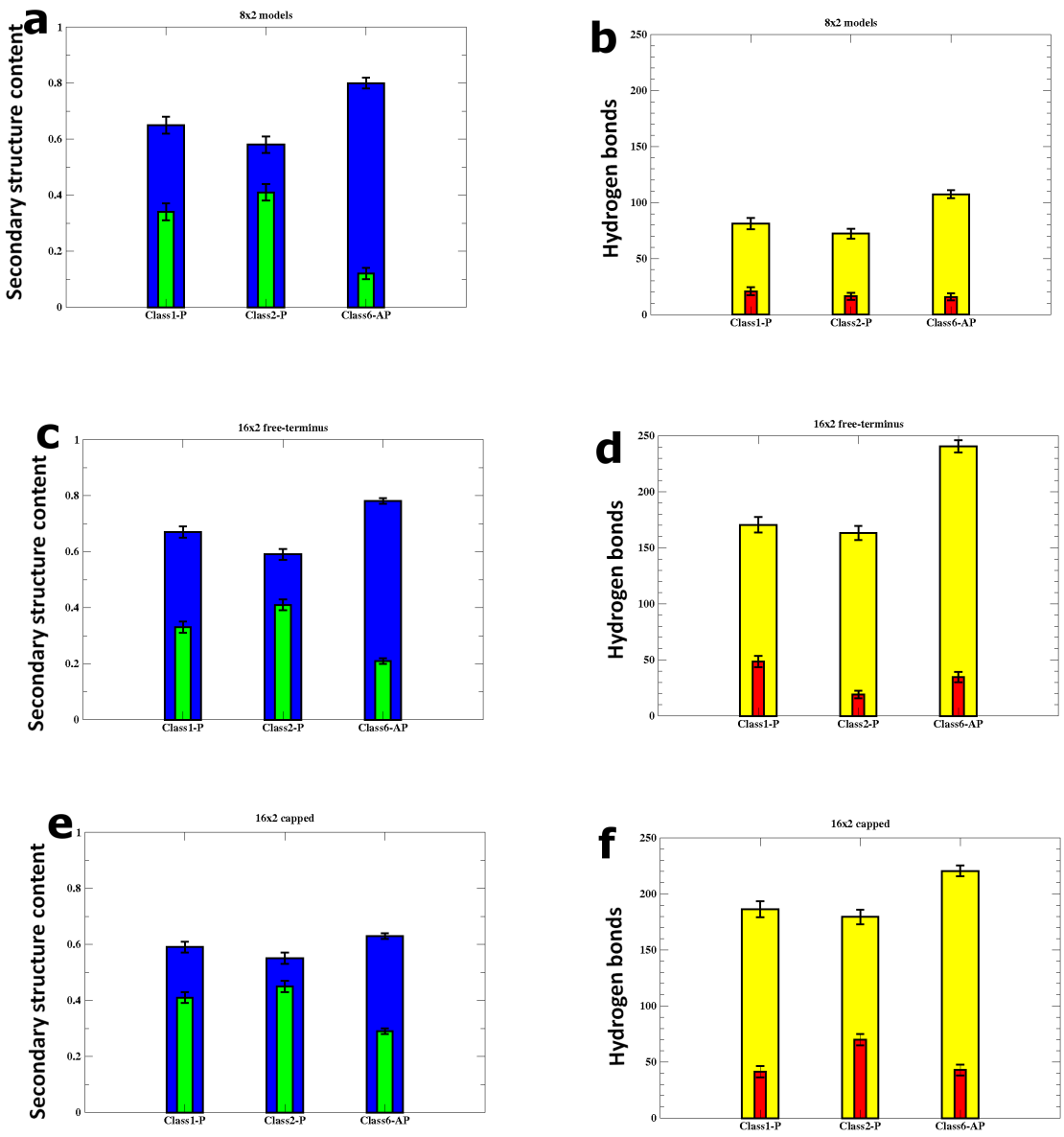

Figure 4.2: Left and right columns show the mean fraction of secondary structure content and hydrogen bond numbers respectively. The details for the 8x2 fibril models are in panels (a) and (b). On the left, blue bars show beta-strand content and green bars show random coil conformations. On the right plots, the mean number of inter-strand backbone (yellow) and side chain (red) hydrogen bonds. Panels (c) and (d) relate to the free terminal ended 16x2 models while panels (e) and (f) are for the capped 16x2 fibrils. The secondary structure content and hydrogen bond analysis is computed from the final 10ns of MD.

model	$\Delta H_{electro}$	ΔH_{vdW}	ΔH_{Solv}	ΔH_{Total}
Class1-P free	-10728.89 ± 178.96	-1142.96 ± 27.35	-4381.56 ± 179.36	-10549.96 ± 51.10
Class2-P free	-2489.12 ± 151.98	-1248.73 ± 25.82	-12410.88 ± 146.13	-10370.50 ± 51.19
Class6-AP free	-12900.62 ± 76.43	-1261.80 ± 26.91	-2298.86 ± 54.64	-10761.36 ± 50.93
Class1-P capped	-10709.91 ± 47.56	-1221.79 ± 24.53	-1554.92 ± 36.28	-6968.85 ± 47.95
Class2-P capped	-10927.08 ± 54.02	-1353.66 ± 28.20	-1402.38 ± 36.14	-7115.22 ± 50.68
Class6-AP capped	-10909.72 ± 42.92	-1394.20 ± 26.41	-1341.87 ± 30.74	-7126.31 ± 49.38

Table 4.2: Contributions to the enthalpy (ΔH_{Total}), calculated using MM-PBSA methodology. Units are in kcal/mol, errors are the standard deviation in the mean. Analysis is from the final 10ns of MD. Electrostatic ($\Delta H_{electro}$) and van der Waals (ΔH_{vdW}) terms calculated from molecular mechanics force field. The solvation free energy (ΔH_{Solv}) is the sum of polar contribution calculated from the Poisson-Boltzmann equation and non polar contribution calculated empirically.

Relatively poor electrostatic energy in the Class2-P uncapped model is a major contributing factor its unfavorable enthalpy (table 4.2). The poor electrostatic energy interactions are in contrast to a very favorable solvation energy for this polymorph. Introduction of the capping groups has a significant impact on these two terms, with a dramatic improvement observed in the electrostatic energy which is coupled to an equally drastic reduction in the favourable solvation energy (table 4.2). The origins of this effect can be understood from a consideration of the orientations of the peptides in each model and the ensuring electrostatic interactions imposed by the geometry.

Figure 4.3 shows the electrostatic potential of the free terminal polymorphs. The class2-P parallel arrangement has peptides oriented in the same direction both within and between the β-sheets. Thus in its zwitter-ionic form regions of similar electrostatic potential directly interfaced resulting in a relatively unfavorable electrostatic energy. Significant electrostatic repulsions occur which are much higher than those of the other two geometries whose peptides are not pointing in the same direction *both* within and between the β-sheets. Conversely, the arrangement in class2-P and the unscreened potential at the terminus end is suited for interactions with the polar solvent molecules as this provide favorable binding sites. This is a possible explanation for the favorable solvation energy compared to class6-AP and class1-P. For all models, the introduction of the terminal caps contributes to relative reductions in the solvation energy and consequently the relative enthalpy too.

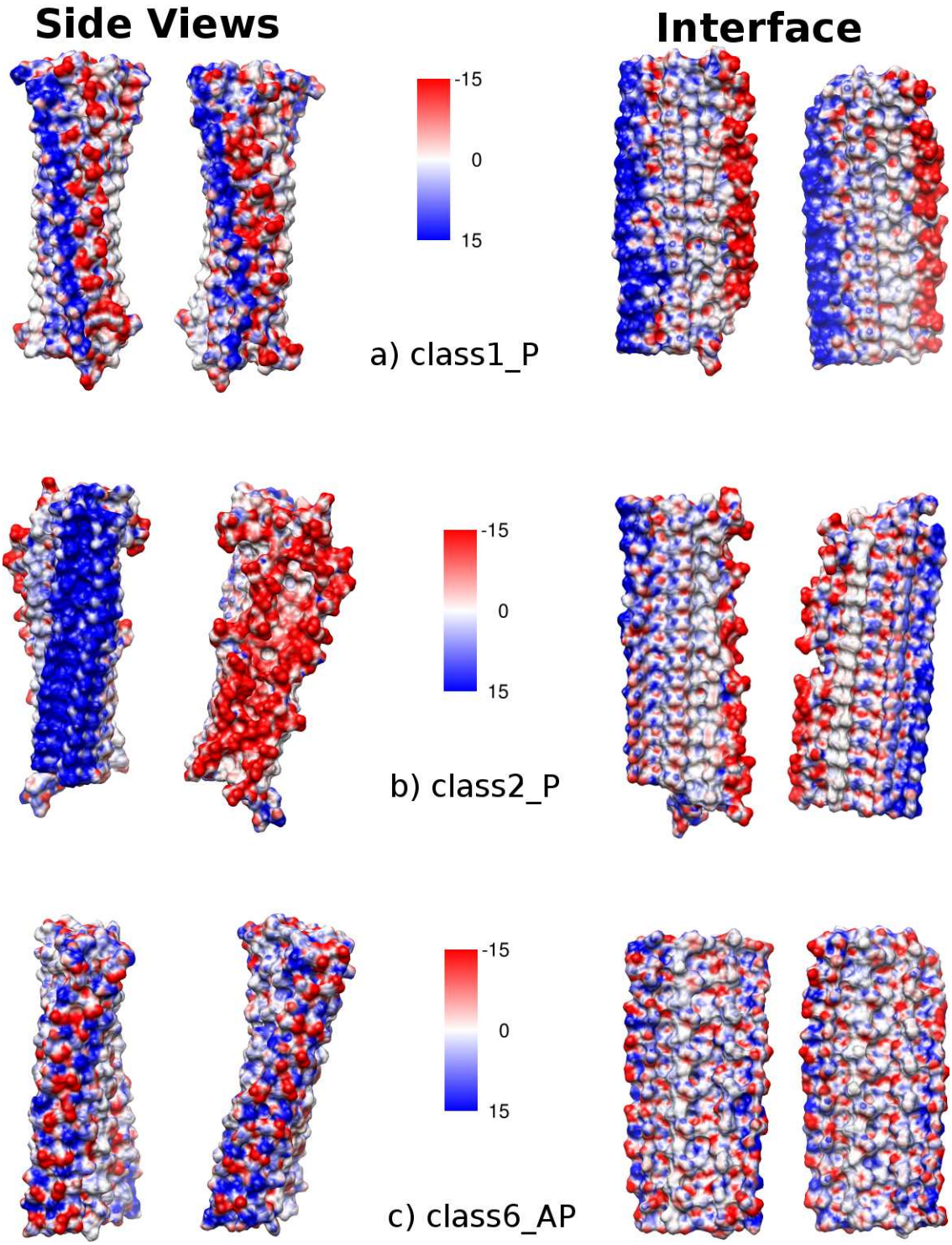

Figure 4.3: Electrostatic potential for the uncapped polymorphs. **Side view:** The fibrils are viewed edge-on with the peptide axis perpendicular to the plane of the page, second image is the perspective of the first image rotated 180° about the fibril axis. **Interface:** The interior of the interacting pair of β-sheets, separated for clarity. The pair mates with the left hand edge of one sheet directly interacting with the right hand edge of the other sheet (and vice versa). Note that although the sheets appear flat in this 2D depiction, they are in fact twisted.

4.4 Results: Mechanical characterisation

Having established how the peptide packing arrangements in each polymorph affect the fibrils structure and stability, the influence of these packing arrangements on the mechanical properties were then examined. The pulling geometries outlined in § 2.4.1 were used to characterise the mechanical response of the fibril polymorphs, with each independent pulling mode simulation repeated four times and the mechanical stability inferred from analysis of peak forces. The discussion of the results below starts with responses of the 8x2 fibrils then moves on to the relative effects of doubling the fibril length and finally the viability of terminal capping as a mechanical modulator.

4.4.1 Responses to anisotropic probing

Figure 4.4 shows the mean peak forces across all four pulling geometries for each of the 8x2 fibrils. As already observed in chapter 3, all three polymorphs demonstrate an anisotropic response to mechanical probing. Similar mean peak forces are recorded amongst the polymorphs when the hydrogen bond networks are probed (shear & stretch). There are however, very distinct responses when the hydrophobic core interactions are probed (peel & slide).

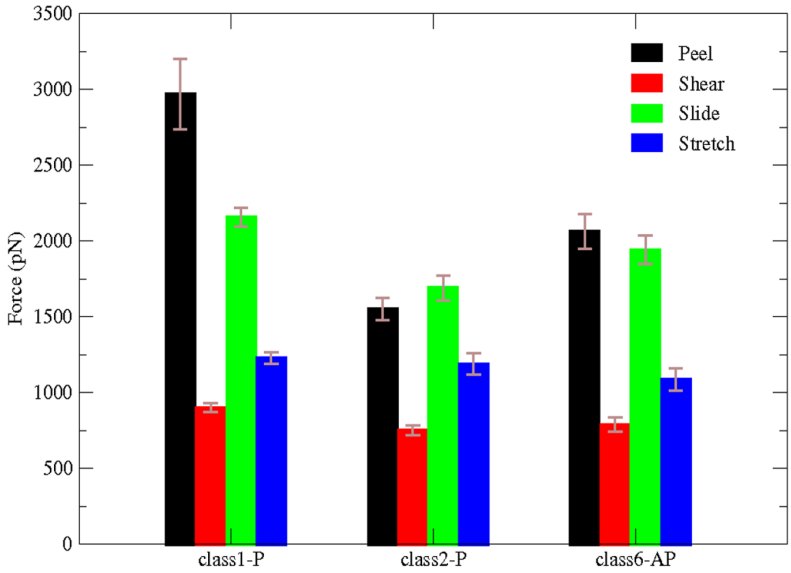

Figure 4.4: Mean peak forces recorded for the uncapped 8x2 fibril polymorphs during each fragmentation mode. Error bars are standard error in the mean calculated from four repeats of each simulation.

Polymorph	$\Delta G_{electro}$	ΔG_{solv}	ΔG_{vdw}	$\Delta G_{Binding}$
Class1-P	-2457.79 ± 4.17	2466.82 ± 4.01	-152.46 ± 0.31	-143.43 ± 0.62
Class2-P	938.98 ± 1.79	-849.46 ± 1.66	-198.60 ± 0.46	-109.08 ± 0.45
Class6-AP	-254.94 ± 1.05	359.97 ± 0.84	-257.77 ± 0.40	-152.74 ± 0.52

Table 4.3: The interactions of the interface between the pair of β-sheets are decomposed into electrostatic ($\Delta G_{electro}$), solvation (ΔG_{solv}) and Van der Waals (ΔG_{vdw}) energy terms which all contribute to the binding free energy ($\Delta G_{Binding}$). Analysis is from 1 ps snapshots of the final 10 ns of unrestrained MD as calculated by the MM-PBSA method. The mean energies are expressed in kcal/mol units, with standard error in the mean.

Hydrophobic core disruption: Peel & Slide

The largest mean peak forces for both hydrophobic core probing modes (peel and slide) are recorded by the Class1-P polymorph. The molecular basis behind the relative ranking in mean peak force between the polymorphs can be understood from examining the inter-sheet interfaces that are affected during the SMD simulation. Both slide and peel modes disrupt the inter-β-sheet interactions by forcing the hydrophobic core residues to be exposed to solvent molecules. For the three polymorphs studied here, a correlation is observed between the mean peak force and the inter-sheet electrostatic interaction energies (table 4.3) that arise due to the unique packing arrangements of the monomer β-strands.

The Class1-P polymorph has the most favourable electrostatic energy between the stacked β-sheets because these are arranged in an anti-parallel configuration, which brings the oppositely charged C- and N-termini close together. However, since Class2-P is in a parallel arrangement both within an individual β-sheet and within the stacked pair, this polymorph has the least favourable electrostatic interaction between the sheets of the three. The fact that the Class 6 polymorph is comprised of anti-parallel β-sheets stacked in a parallel configuration places it intermediate between the other two. The correlation between the peak force and the electrostatic interfacial energy demonstrates how the details imposed by polymorphic arrangements of the peptides in fibril can determine the mechanical characteristics when force is applied in a particular direction.

Hydrogen bond network response: Shear & Stretch

In the two pulling geometries (shear and stretch) that primarily interrogate the hydrogen bond networks, similar mean peak forces were recorded for all three polymorphs. The stretch and shear simulations probe the inter-peptide hydrogen networks in directions parallel and perpendicular to the fibril axis respectively. Figure 4.4 shows that pulling parallel to the hydrogen bond network results in higher peak forces than when pulling across it. This implies that the hydrogen bond network provides a cooperative resistance to the forces applied in the direction of the long axis.

A surprising aspect of these simulations is that the polymorphs record virtually identical mean peak forces when subjected to stretch SMD, in spite of the fact that they contain different numbers of hydrogen bonds (fig 4.2b). Consideration of this and the classical Pauling-Corey hydrogen bonding model would imply that anti-parallel class6-AP model should be much more resistant to these deformations than either of the two parallel fibrils. Moreover, a systematic simulation study of the relationship between thermodynamic stability and the symmetry class of fibrils has shown that in non-Q/N rich sequences, the anti-parallel fibrils tend to be more energetically stable than their parallel counterparts (21).

The molecular basis for the measured result here is however intertwined with the secondary structure elements present at the end of the dynamics in each model prior to SMD. The final configurations of the Class1-P, Class2-P, and Class6-AP structures consist of 34%, 40% and 19% random coil conformations respectively (fig 4.2a), indicating that some native backbone bonding defects are present in all fibril models. These defects can dominate the mechanical response of the fibrils in a particular pulling direction by providing weak points that liable to fracture, as previously seen in chapter 3.

4.4.2 Effects of increased fibril length

Having demonstrated that pulling along the long axis of the fibril by the stretch deformation mode is sensitive to the presence of structural defects within the fibrils, an examination was made into how the normalised peak force changes when the fibril doubles in length from 8 peptides in each pair of stacked β-sheets (8x2 model) to 16 peptides (16x2 model), since there is a greater probability that structural defects will be present in longer fibrils. The

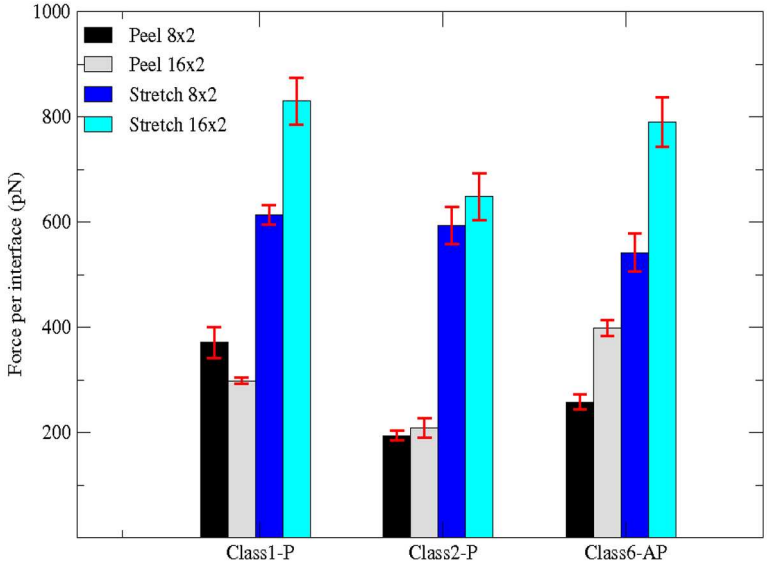

Figure 4.5: Force per interaction interface normalised to compare relative response of 8x2 fibrils vs 16x2 fibrils for peel & stretch SMD. Error bars are standard error in the mean calculated from four repeats.

influence of fibril length on mechanical properties has already been demonstrated through normal mode analysis in conjunction with a coarse-grained elastic network model based on SNNFGAILSS fibrils (113), which showed that the bending rigidity increases up to a critical length. It is however, not possible to assess the importance of defects within such a coarse-grained model. The two sizes of each fibril polymorph are interrogated by the stretch and peel SMD modes that probe the hydrogen bond networks and hydrophobic core interactions respectively. For direct comparison between the 8x2 vs 16x2 fibril lengths, the mean peak force during each SMD is normalised to the number interfaces probed (fig 4.5).

Increased fibril length: Response to Stretch

Figure 4.5 shows all three fibril models register an increase in the mean peak force per interface upon elongation for the stretch mode (blue vs cyan). This is to be expected since the numbers of hydrogen bonds are increased hence their cooperative action in resisting the pulling action of the stretch SMD (parallel to this network) is enhanced. However, the relative gains in the mechanical resistance appear to be unique to each polymorphic model. Doubling the length for Class1-P and Class6-AP models leads to an increase in the mean peak force per interface of 35% and 46% respectively whilst for Class2-P the increase

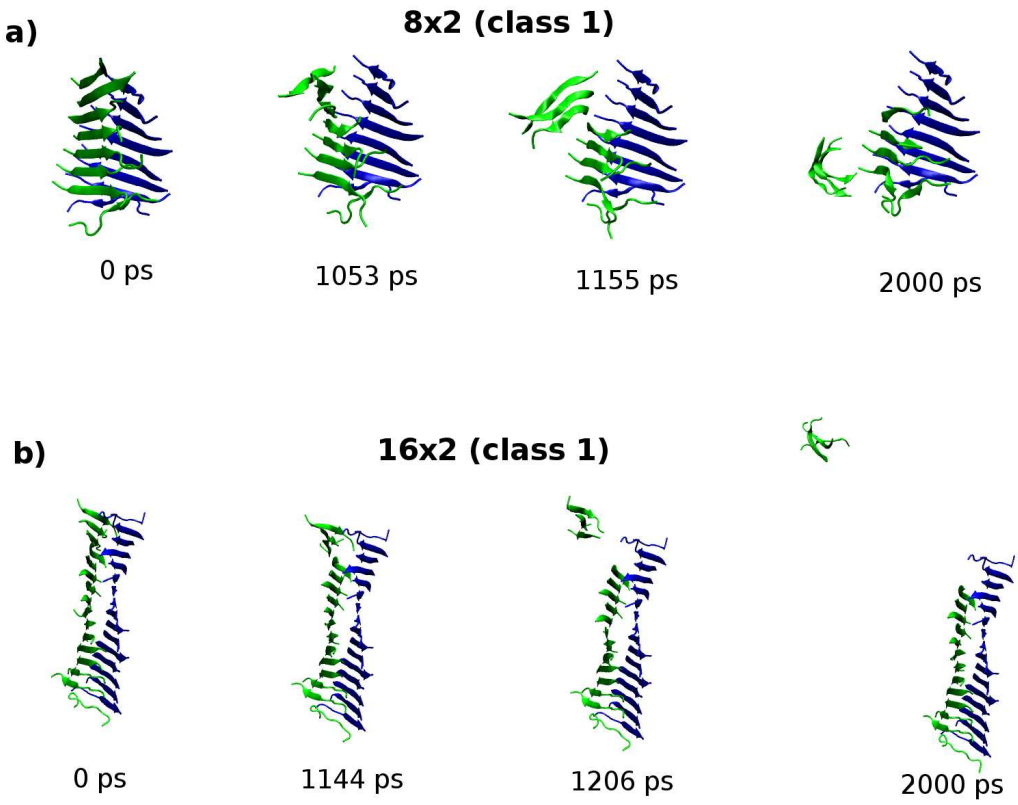

Figure 4.6: Molecular configurations sampled from the peel SMD trajectory for (a) 8x2 and (b) 16x2 fibril models of the class1-P polymorph. The carbon-α atoms in the peptides coloured green are pulled while carbon-α atoms in peptides shown in blue are fixed. In contrast to the 8x2 fibrils, the response of the 16x2 fibrils leads to an exposure of only a small fraction of the hydrophobic surface.

is marginal at 9%. The ranking of peak force per interface amongst the polymorphs is reflected in the total number of backbone and side chain hydrogen bonds present in each model (fig 4.2d).

Increased fibril length: Response to Peel

Figure 4.5 also shows the normalised peak forces per interaction interface for the peel geometry that addresses the hydrophobic core between the β-sheets. It is interesting to note that doubling the fibril length results in very distinct responses among the polymorph models in this direction. The net effect of elongation for Class2-P is a very marginal increase (within the error) of 7% in the peak force per interface while Class6-AP registers a significant 54% increase. Intriguingly, the Class1-P polymorph suffers a reduction of $\approx$19% such that

the normalised force per interface is higher in the case of the shorter 8x2 fibril than the 16x2 model. In contrast to the shorter (8x2) aggregates, applying peel SMD to the 16x2 fibrils results in a fragmentation of only a small fraction of the hydrophobic surface (fig 4.6). Consequently it was not possible to relate the peak force measured to the thermodynamic stability of the hydrophobic interface, because it is not completely disrupted during the deformation. To understand the mechanical robustness of the fibrils, it is necessary to have information about the structure of the fragments that result, as well as the structure of the unperturbed fibrils themselves.

4.4.3 Mechanical modulation by terminal capping

The final strategy explored for modulating the mechanical properties of the fibrils was the addition of terminal capping groups at both ends of the peptide strands which neutralise the charged groups at either ends of the peptide monomers. The capping is in the form of N-terminal acetylation and C-terminal amidation. The thermodynamic and structural analysis in § 4.3 shows that the capped fibrils have significantly distinct properties relative to the uncapped models. Here, the influence of these differences on the relative mechanical response of each polymorph model are examined and the viability of terminal capping as a mechanical modulator evaluated. The mean peak forces recorded through the peel and stretch SMD pulling modes are compared for the 16x2 capped and uncapped polymorphic variants[1].

Peel: Capped vs Free terminus

In the Peel pulling geometry, the net effect of adding capping groups appears to be higher peak forces for Class1-P and Class6-AP whilst no attenuation is observed for Class2-P fibrils, as shown in figure 4.7. The fragmentation mechanisms between capped and uncapped models are also a distinct from each other within a polymorphic state. Figure 4.8 shows a comparison of the evolution of the distances between peptide pairs on opposite β-sheets for the class1-P model during each set of repeat peel SMD simulations. The resistance mechanism for the charged-termini models shows that the separation of the peptide pairs occurs gradually, i.e. the fibrils have ductile characteristics. In stark contrast, the capped

[1]The capped models require the use of different simulation parameters for the time-step and pulling velocity (0.5 femto-secs & 0.04Å/ps). For cross comparability, a new set of peel & stretch SMD simulations were run for the uncapped models.

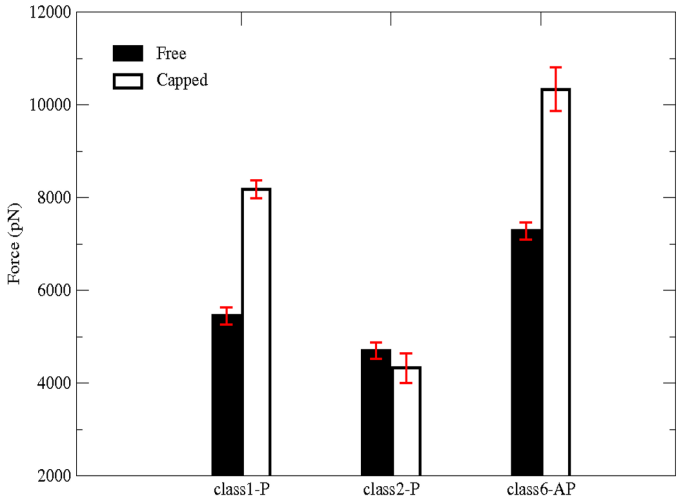

Figure 4.7: Effect of terminal capping on the mean peak force during peel pulling mode. Response of free terminal fibrils (black bars) is contrasted against capped fibrils (white bars). Error bars represent standard deviation in the mean calculated from four repeat simulations.

models undergo significant displacements over a very short period of time; consequently, they are more brittle than the capped counterparts. This could be due the changes in the fibrils electrostatic interactions which is a direct consequence of the terminal capping. In the capped case for example, the fibrils may break suddenly because the interactions stabilising the fibrils (hydrogen bonding and hydrophobic forces) are short ranged in comparison to the long range electrostatic interactions within the uncapped fibrils. This highlights how a relatively simple modification at the terminus end can have a significant impact on the mechanical character of amyloid fibrils formed from short peptide sequences.

Stretch: Capped vs Free terminus

The effects of the capping groups in the Stretch SMD direction on the mean peak forces are somewhat dependent on the structural details of each polymorph prior to the pulling. Figure 4.9 shows the mean peak forces for the capped and uncapped versions of the three polymorphs models during stretch SMD simulations. The trends between the free-terminus polymorphs tend to reflect the numbers of hydrogen bonds (fig 4.2f) in each of these models with class6-AP the most resistant and class2-P the least. However, the rank among the capped models does not seem particularly correlated to the relative density of the hydrogen bond networks, for instance the Class1-P peak force is higher than that of Class6-AP. Fur-

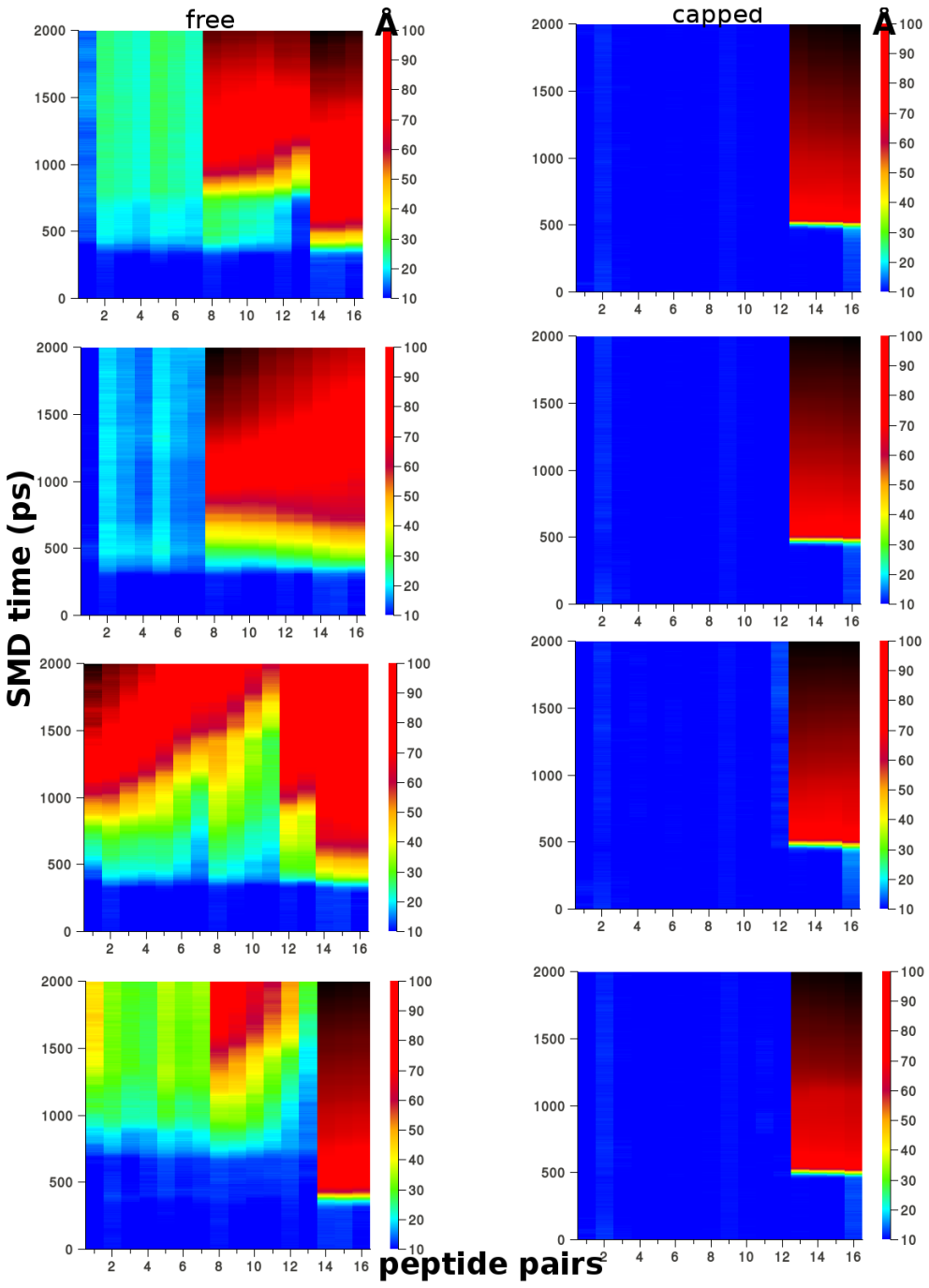

Figure 4.8: Plots showing the displacement between peptide pairs on opposite β-sheets during the peel SMD for the class1-P polymorph model. The free (left) and capped terminal (right) models are each compared from four independent simulations. Colour scale is centre of mass distance between pairs in angstroms (Å), the x-axis is the pair number (total of 16), and the y-axis is time in ps.

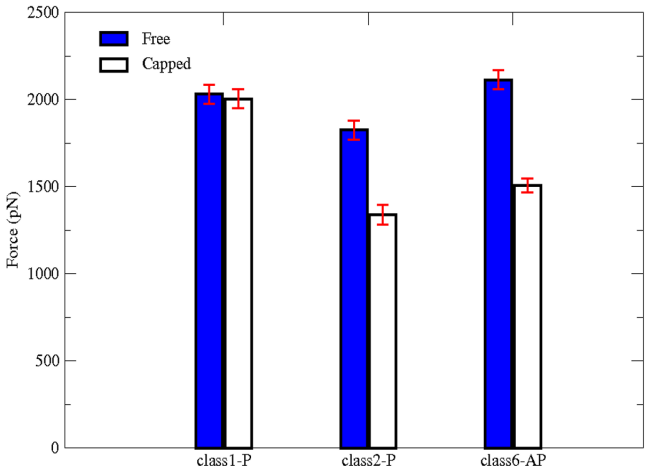

Figure 4.9: Effect of terminal capping on the mean peak force during stretch pulling mode. Response of free terminal fibrils (blue bars) is contrasted against capped fibrils (white bars). Error bars represent standard deviation in the mean calculated from four repeat simulations.

thermore, the capped class2-P and class6-AP models both record lower peak forces relative to the uncapped counterparts.

The molecular basis for the response of the capped and uncapped fibrils to stretch SMD is unique to the details of each polymorph as determined by both the number and the nature of the defects present in the fibril models. Although both the Class1-P and Class2-P polymorphs show an increase in the number of hydrogen bonds when the termini are capped, the quantity of ordered β-sheet secondary structure is reduced, indicating an increased number of defects within the aggregates. Consequently, the mean peak force per interface required to break the fibrils is reduced or remains the same when the termini are capped.

The behaviour of the Class6-AP polymorph, which shows a large reduction in the mean peak force for the capped fibrils, provides a particularly striking example of how the response of fibrils to an applied force can be dominated by the molecular details of the defects present. Figure B.2 shows the starting configuration for the SMD simulations for the capped fibril model. This polymorph developed a substantial crack defect in the one of the paired β-sheets prior to SMD, which substantially reduced the ability of this fibril model to resist stretch relative to its uncapped counterpart.

4.5 Conclusions

Polymorphic packing arrangements:

The SMD simulations used to mechanically probe three polymorphs of fibrils formed from the SNNFGAILSS sequence reveal an anisotropic response to the four pulling geometries used. There is however, a significant difference in the relative responses between the polymorphs when compared across the individual pulling geometries. Moreover, it is possible to switch the fibrils from ductile to brittle by changes to the balance between short and long range electrostatics. The results collectively demonstrate how the mechanical response of fibrils is directly related to the peptide packing arrangements and the number and nature of the defects present within the models.

Mechanical modulation:

The relative mechanical performances of the different sized fibrils produced contrasting results across the three polymorphs. For Class1-P, elongation led to increased mechanical resistance in one direction that was coupled to increased fragility in another. Fibril elongation had a very minimal effect on the response of the Class2-P model whilst Class6-AP saw significant improvements. The insights gained from such simulations could be particularly useful, for example, in the selection or design of a template from which functionalised peptides can self-assemble into conducting nano-wires. Since such a nano-wire structure needs to be as long as possible but not prone to fracture, the Class6-AP polymorph would, on the basis of these simulations, be a suitable choice relative to either Class1-P or Class2-P for the SNNFGAILSS sequence.

The efficacy of terminal capping as a mechanical modulator was shown to be dependent on both the polymorphic arrangement and the pulling geometry of the SMD mode. Significant gains in resistance were obtained for Class1-P & Class6-AP models upon capping while no difference was recorded in the relative responses of the Class2-P in the Peel SMD mode. The utility of the capping strategy could become apparent when an ideal polymorphic arrangement of a specific peptide sequence is desired for a nano-device but the resulting fibrils found to be fragile (or too robust) for the purpose. The simulation data show that for some polymorphic arrangements minor terminal modifications could be used as means to stabilise (or weaken) the mechanical character in a non-trivial way to suit the desired purpose.

Significance of defects in the fibrils:

The small model fibrils investigated in this study contain a high number of structural defects, because they lack the stabilisation from crystal packing within a larger aggregate. Consequently, it is possible to characterise how the nature and presence of such defects influences their mechanical response. These illustrate the general principles that must be considered when evaluating and comparing the mechanical properties of amyloid fibrils containing structural defects. Although a number of fibril species have been reported to contain extremely high degrees of structural order with defects present approximately one every 1000 molecules along the fibril axis (114) nonetheless, experiments probing the mechanical properties of α-synuclein and full length transthyretin using high pressure conditions have shown that their robustness is indeed dominated by the presence of defects within the hydrophobic core (115). In the crystallisation of inorganic substances, such as minerals and ceramics, it is known that the number of defects will depend critically upon how these crystals were grown, with factors such as the rate of growth, the presence of surfaces or impurities, and whether the solution was agitated. If amyloid fibres with bespoke mechanical properties are to be used in nanotechnology, it will be necessary to carefully assess the reproducibility of the experimental conditions used to produce the fibrils, because small changes in the manufacturing could potentially alter the polymorphic form or the number density of defects present, and substantially affect mechanical robustness.

Chapter 5

Engineering defects into fibrils to alter mechanical characteristics

5.1 Overview

Synopsis

This chapter explores the implications of the results presented in chapters 3 and 4 that relate to the ability of defects in the fibril structure to dominate mechanical characteristics in a given pulling direction. In contrast to the preceding chapters however, this work probes mature fibres formed from Amyloid-beta (Aβ) and related variants that lead to early on-set Alzheimer's disease § 1.5. Two separate studies centering on understating and then exploiting the mechanisms of failure in the fibrils are presented. The focus of the first investigation is on how defects that arise at peptide interfaces due to heterologous cross-seeding of the wild-type Aβ fibres with peptides containing physiological mutations, such as those found in Familial Alzheimer's Disease (FAD), can attenuate the mechanical properties in a manner that might be insightful for the design of nano-materials. A systematic examination is made on the relationship between the physical changes imposed on the fibril structure through various degrees of cross-seeding and the mechanical response. The second investigation centers on whether the observed clinical trends in Familial Alzheimer's Disease such as increased toxicity & polymerisation can be directly attributed to the frangibility of the Aβ FAD variants.

Introduction

At present, there is an ever growing list of proposed applications that seek to exploit the desirable properties of self assembling amyloid fibrils in the construction of novel nano-materials and devices (see § 1.3). The diversity of such potential applications require an ability to both understand and control various aspects of the fibrils material properties for efficient design or use. Such endeavours can take advantage of the abundant knowledge already gained from biological studies that have primarily focused on understanding the disease aspects of amyloid. This could include the effects of temperature, solvent choice, pH , salt concentration or agitation on the morphology of fibrils grown *in vitro*. One such effect could be the use of selected fibril fragments to seed the growth of daughter fibrils with drastically different structural characteristics from the parent fibrils. The presence of the seeding material allows for a ready nucleus on which fibril elongation by monomer addition can proceed. The seeding material is usually composed of fragments of preformed homogeneous fibrils that then eliminate the characteristic lag phase in the elongation of the

mature fibrils (see fig 1.2). Depending on the perspective from which amyloid growth is seen, the absence of a lag phase induced by seeding could be viewed as either disastrous or desirable. In amyloid related diseases, this is more likely to be an undesirable consequence as it is largely associated with the early on-set and progression of symptoms. In *in-vitro* nano-material construction processes however, elimination of the lag phase could be a very desirable step as it would likely increase the efficiency of the assembly process.

In some cases, fibril fragments from one peptide sequence are able to act as seeds for the growth of a different amyloidogenic peptide sequence. This is known as 'cross-seeding' and has wide implications for the risk to the health of patients already suffering from one form of amyloidosis. It is possible for instance that a patient of Type II diabetes might then become susceptible to developing a secondary amyloid related condition such as Alzheimer's disease if fragments of hIAPP fibrils can seed the growth of Aβ_{40} fibrils. The very close sequence similarly between the hIAPP and Aβ_{40} particularly makes this a very feasible prospect although *in-vitro* elongation experiments have since suggested that hIAPP fibrils are very poor seeds for Aβ_{40} elongation (116). The reduction in the risk of one amyloid species potentially influencing the formation of another is thought to be a major reason why the *in-vivo* polymerisation of functional amyloid tends to be a highly regulated process. An example in humans would be the biosynthesis of melanin which takes place in specialised organelles since disease related Aβ & α-synuclein fibrils are also able to act as templates for melanin polymerisation (40). Cross-seeding is also a means for the propagation of different amyloid strains. In a recent study of the cross-seeding between bovine insulin fibrils with variant human insulin analogues, the transmission of phenotypic features from the mother seeds to daughter fibrils is observed upon cross-seeding despite the formation of morphologically distinct fibrils (117).

Cross-seeded amyloid growth offers an attractive avenue and means for the modulation of the material properties of the resultant fibrils that could be then exploited when creating novel amyloid based nano-devices. The seeding material can be engineered to induce 'defects' along the fibril structure that could for example switch the fibrils material character from brittle to ductile depending on the desired application. To this end, the computational work presented in this chapter systematically investigates the efficacy of various types of cross-seeding configurations in altering the structural and mechanical characteristics of Aβ

fibrils. A number of models of wild-type fibrils were built containing differing degrees of cross-seed interfacing with heterologous monomers associated with Familial Alzheimer's Disease (FAD). The variant Aβ FAD monomers used are those of the Arctic and Iowa mutations in the WT sequence in positions E22G & D23N respectively. Additional models that act as controls for the investigations undertaken made use of proline substitutions in either the L17P or E22P positions of WT Aβ sequence. An experimental *in-vitro* mutagenic cross-seeding analysis of a number of variant Aβ_{40} peptides (at position V18) with the wild-type (WT) reveals that their cross-seeding susceptibility is closely related to the aggregation propensity of the peptide chain (118).

The main aim of the simulations presented here was to assess whether the chosen cross-seed interfaces that were engineered into the fibrils played a significant role in either enhancing or reducing the measured peak forces relative to the WT. The performance of the WT fibril was also contrasted to homogeneous models made entirely of the variant sequences. A strong incentive for using the Aβ fibrils in this investigation was the sequence similarity with the full length amylin protein from which the SNNFGAILSS fibrils in chapters 3 and 4 are derived. Although amylin fibrils are yet to be structurally determined there is strong evidence to suggest that they also adopt the 3-fold symmetric arrangement exhibited by Aβ (119). This in addition to the sequence similarity naturally makes Aβ a good choice to extend the concepts developed from studies of the short the amylin SNNFGAILSS fragments.

Unlike the pair of β-sheets in the SNNFGAILSS fibrils however, the size and shape of the 3-fold symmetric Aβ_{40} fibres imposes restrictions on options for meaningful SMD pulling geometries that can adequately probe its mechanical properties. The primary SMD protocol employed throughout this chapter is concerned with disrupting the hydrogen bond interactions along the fibril as these are believed to play a pivotal role in the overall mechanical stability. Due to the anisotropic response properties, a secondary pulling geometry independent of the hydrogen bond network is also used to analyse the homogeneous fibril models.

5.2 Construction of Models & Simulation details

The detailed methods for building the various Aβ fiber models are outlined in § 2.3.2. Initial coordinates of a six layered 3-fold symmetric WT Aβ structure determined by ssNMR experiments, were used as a template for constructing a 12-layered fiber made of a total of 1152 protein residues. Modifications to the Aβ sequence in all the U-shaped monomers inspired by two early on-set Familial Alzheimer's disease mutations were then used to make two additional fibril models. The sequence substitutions were in positions E22G (Arctic) and D23N (Iowa). Controls for the investigations undertaken also required the construction of models containing proline substitutions in positions L17P (Proline-L17P) and E22P (Proline-E22P). Both the homogeneous Proline-L17P and Proline-E22P variants form amyloid fibrils that retain the fundamental structural features found in the WT fibrils (120). The starting structures prior to equilibration and dynamics for the 12-layered wild-type, Arctic, Iowa and Proline-L17P models are shown in figure 5.1a.

In addition to these homogeneous fibrils, a number of models in different configuration types of cross-seeding between the variant monomers and WT fibrils were built. The location and nature of the cross-seed interfaces are schematically shown in figure 5.2 and described in the following section. In all, a total of 21 individual (12-layered) Aβ fibers were modeled for the simulation work described in this chapter. It should be noted that all models constructed here assume that the variants also adopt the 3-fold symmetric configuration exhibited by the wild-type fibril. This may not necessarily be an accurate representation as there is evidence from EM images of amyloid morphology for instance that shows Arctic fibrils have larger diameters relative to WT (121) while Iowa fibrils exist in highly polymorphic states (17). Nonetheless, in the absence of experimentally determined structures for these fibrils, the WT configuration is a justifiably reasonable starting point to explore the relative mechanical effects induced by the various sequence substitutions.

Engineered defect sites

Sections of the 12-layered wild-type fibril were replaced with Aβ monomers of either Arctic, Iowa, Proline-L17P or Proline-E22P variants to form cross-seeded interfaces. The locations of the replacement layers are schematically depicted in figure 5.2. The cross-seeding replacements are applied to all three U-shaped monomers in the affected layers of the fibril.

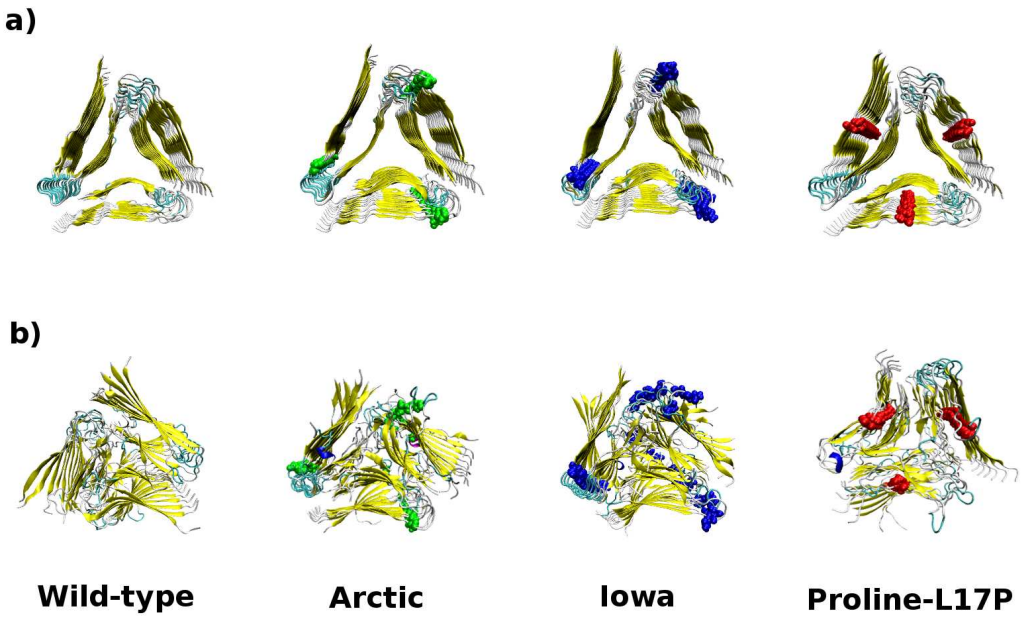

Figure 5.1: (a) The initial and (b) final structures after unrestrained MD simulation for wild-type, Arctic, Iowa & Proline-L17P homogeneous fibril models. The coloured spheres represent the location of the sequence substitutions across the fibril monomers: green is E22G (Arctic), blue is D23N (Iowa) and red is L17P.

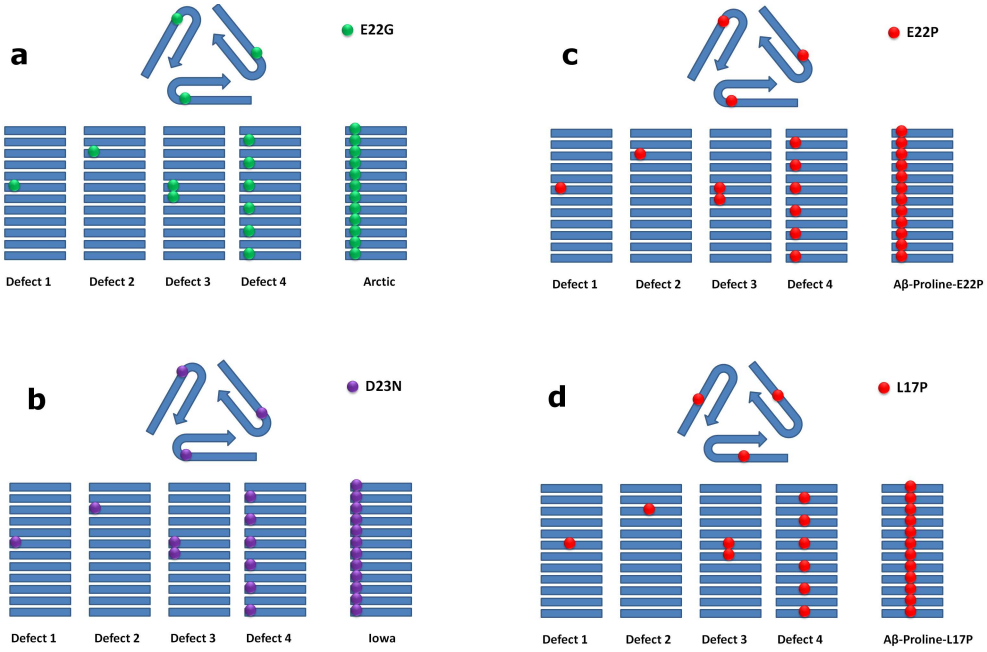

Figure 5.2: Location of the cross-seed layers in each model. Four unique configurations (Defect 1 to 4) were constructed using E22G, D23N, L17P & E22P variants in addition to the homogeneous fibrils.

The cross-seed sites fall into four unique configurations which are namely, Defect 1, Defect 2, Defect 3 and Defect 4 (fig 5.2). In Defect 1, a single layer of sequence modified monomers is used to replace the middle layer of the wild-type fibre. In the Defect 2 configuration, the cross-seed replacement is placed near one end of the fibre and is located on the third layer from the top. The placement for Defect 3 is located in the middle of the fibril but has two interfaced modified layers replacing the wild type monomers. Finally, in Defect 4, half the WT layers are replaced in an alternating manner with the modified monomers. The configuration of each scheme is designed to assess how the location and number of the cross-seed interface sites attenuate the mechanical characteristics relative to the wild-type model.

Molecular Dynamics

The energy minimisation, equilibration and production MD phases for the various Aβ models followed the simulation protocol described in § 2.5.1. Each fibril was explicitly solvated in TIP3P water molecules with periodic boundary conditions employed. The combined protein and water environment system consisted of ≈140 000 atoms for each model. The production MD simulations were run until convergence in the RMSD of backbone Carbon-α atoms had been attained. For these cross-seeded models, convergence occurred within 30-40 ns of unrestrained MD. The results for the RMSD, Secondary structure content, enthalpy and hydrogen bonding for the cross-seed models, calculated from the final 10 ns of the converged MD simulations, are shown in Appendix C.

SMD force application

Tensile pulling: Once the production MD simulations had fulfilled the convergence criterion, the fibrils were then subjected to tensile pulling SMD simulations. The implementation of the pulling geometry for the Aβ_{40} fibre takes into account the triangular 3-fold symmetry and is a modification of the 'Stretch-SMD' mode (§ 2.4.1). The steering forces are applied along the same direction as the hydrogen bond network and the fibre is pulled from one end along its length until reaches its tensile limit. The modified protocol requires the backbone carbon-α atoms of peptides on one end of the fibril ('bottom') to be held fixed whilst the center of mass of the backbone carbon-α atoms at the other end ('top') of the fibril are pulled linearly. A schematic representation of the tensile SMD pulling geometry is

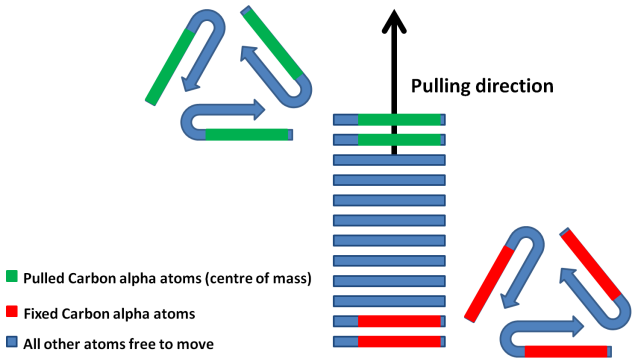

Figure 5.3: Schematic representation of the tensile pulling SMD geometry. The center of mass of carbon-α atoms in portions of monomers marked in green is pulled in the direction indicated by the arrow whilst the carbon-α in portions marked in red are held fixed. All other atom types are free to move unrestrained.

depicted in figure 5.3. The fixed carbon-α atoms are chosen from residues on the 'bottom' two layers of the 12-layered structure located in the first β-strand region prior to the loop in each monomer. Conversely, the pulled carbon-α atoms are chosen from the 'top' two layers, also employing similar residue selection. The SMD parameters are; a pulling velocity of 0.04 Å/ps, spring constant of 500pN/Å and a time-step of 0.05 femto-seconds with a simulation duration of 2 ns. The SMD probing simulations were repeated four times for each fibril system.

5.3 Insights from heterologous cross-seed defect interfaces

The mean peak force measured from four repeat simulations for each cross-seeded model relative to the WT fibril are shown in figure 5.4. The range of peak forces reveals how both the location within the fibril and chemical character of the cross-seed variant peptide can have a significant impact on either enhancing or reducing the peak force relative to the WT. This gives hints of changes in the fundamental mechanical characteristics of the fibrils as a direct consequence of the presence of the modified residues.

With the exception of the control Proline-L17P variant, all models register a relative increase in peak force for the single cross-seed site located in the middle of the fibril (D1). However, all models then see a drop for the D2 site, for which the cross-seed interface is located nearer to the pulled end of the fibril. The Arctic variant in the D3 configuration reg-

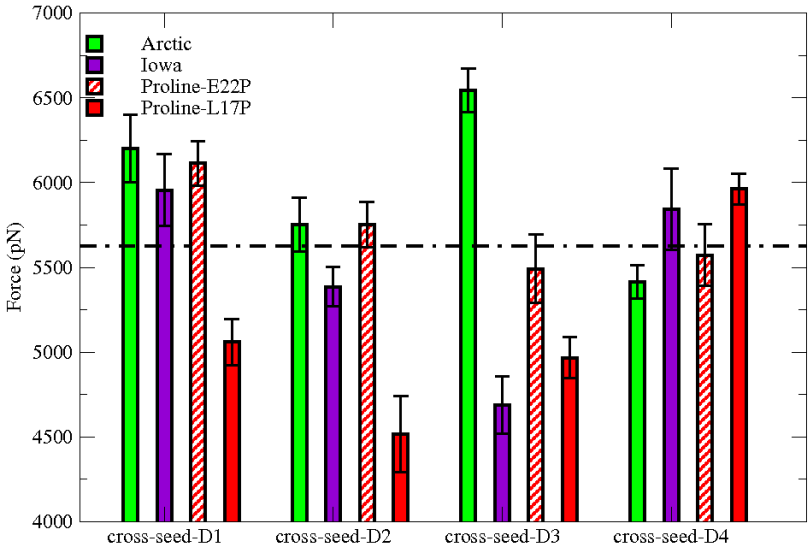

Figure 5.4: Mean peak forces recorded for the tensile pulling geometry for each cross-seeded variant model. The mean peak force obtained for the wild-type fibril, shown as dotted line, serves as reference point to highlight the degree of attenuation caused by the presence of the cross-seed monomers . Error bars are standard error in the mean from four independent SMD simulations.

isters the highest mean peak force which is significantly larger than that of WT. The mean peak forces in figure 5.4 suggest that the various integration of the modified peptides into the WT fibril can have measurable effects on the mechanical response. This computational approach could be adopted in the design strategies of novel amyloid based nano-materials as a viable means of attenuating the mechanical character of the fibril to match the desired application.

The attenuation of the mean peak force relative to the WT model (fig 5.4) can be understood from monitoring the sequence of events that occur during the fibril breakage simulations. Visual inspection and post-processing analysis of the SMD trajectories, in particular, provide insights into the relationship of between failure points relative to the location of cross-seed interfaces and the corresponding attenuation in mean peak force. A general trend from these observations is that in the cases where lower peak forces are obtained, the breakage of the fibril occurs at the cross-seed interface. Conversely, for cases with higher mean peak forces than WT, fibril breakage does not involve the cross-seed interface.

An example of this phenomenon is shown in the comparative time evolution of the Arctic and Proline-L17P cross-seed variants in the the D1 configuration during four repeat SMD

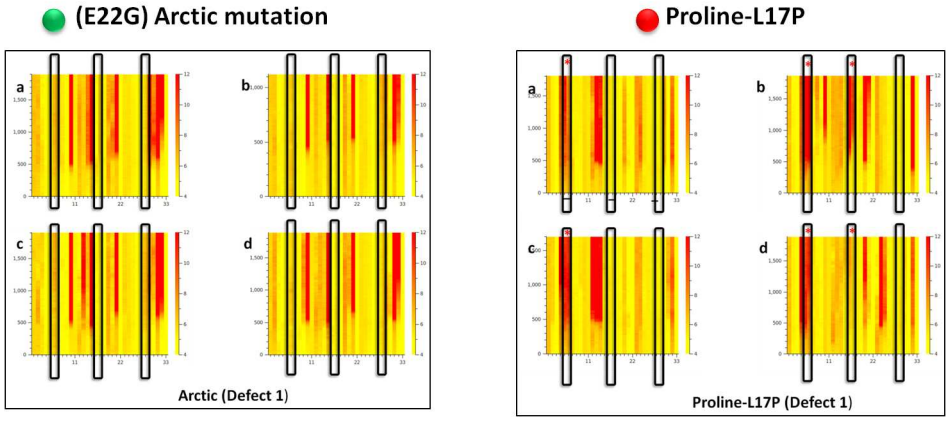

Figure 5.5: Plots of the displacement of centre of mass distance (angstroms) between adjacent peptide layer pairs (x-axis) as a function of SMD time (y-axis) in pico seconds. Left panel has the evolution of four independent simulations (a-d) for the Arctic variant while the right panel is for the Proline-L17P variant, both in the D1 cross-seed configuration. The black rectangles in each plot highlight the location of the cross-seed interfaces in the 3-fold symmetric fibrils. The red asterisk in the black rectangles denote breakages that occur at the cross-seed interfaces.

simulations (fig 5.5). The D1 Arctic variant breaks at locations that do not involve the cross-seed interface, with a correspondingly high peak force relative to WT (fig 5.4). The breakage locations for D1 Proline-L17P variant however, occur at the cross-seed interfaces implying that these act as weak points hence the correspondingly low peak force relative to WT. Analysis of the other model systems also yield similar trends and relationships to those in figure 5.5.

The relationship between the observed fibril fracture points and the cross-seed interface was thus put on a quantitative footing by defining a dimensionless parameter, η. This is the ratio between the average number of observed breaks occurring at the cross-seed sites, n, in the four repeat simulations and total number of cross-seed sites, N, available on the three faces of the 3-fold symmetric fibril.

$$\eta = \frac{\langle n \rangle}{N} \tag{5.1}$$

The value of η has a maximum of 1 if breakage is observed at all cross-seed sites in all the three faces of the fibril and a minimum of 0 if the fibril breaks at any other location that does not include the cross-seed interfaces. The calculation of η for a set of repeat simulations proceeds from the analysis of inter-peptide displacement plots such as those in

Configuration	Aβ variant	$\langle\eta\rangle$	Interface break	$\langle Force\rangle$ pN
D1				
	Arctic	0	×	6201.58 ± 108.13
	Iowa	0	×	5955.95 ± 210.36
	Proline-E22P	0	×	6113.51 ± 132.59
	Proline-L17P	0.5	✓	5058.92 ± 136.60
D2				
	Arctic	0.75	✓	5752.00 ± 157.55
	Iowa	0.33	✓	5386.13 ± 114.51
	Proline-E22P	0.67	✓	5751.86 ± 135.46
	Proline-L17P	0.92	✓	4515.58 ± 225.59
D3				
	Arctic	0	×	6543.53 ± 130.47
	Iowa	0.5	✓	4687.41 ± 168.52
	Proline-E22P	0.08	✓	5491.83 ± 201.83
	Proline-L17P	0.75	✓	4966.21 ± 121.37
D4†				
	Arctic	-	✓	5415.43 ± 98.590
	Iowa	-	✓	5843.20 ± 240.08
	Proline-E22P	-	✓	5571.87 ± 181.51
	Proline-L17P	-	✓	5961.40 ± 91.100

Table 5.1: Summary of the fracture locations relative to the cross-seed interface. The quantity η for each configuration is averaged from four repeat simulations. The (✓) symbol denotes the breakage of a fibril at a cross-seed site while (×) denotes breaks that do not occur at the cross-seed interfaces. The peak forces obtained for each configuration can be contrasted to WT, 5625.21 ± 124.66 pN. †Note that the in the D4 configuration every WT layer is always interfaced with a variant layer.

figure 5.5. For each cross-seed configuration, the magnitude of mean peak force relative to WT and corresponding value of η are summarised in table 5.1.

In the D1 cross-seed configuration all models, with the exception of the control Proline-E17P, record a value of $\eta = 0$. This implies that despite these models undergoing some mechanical failure during the probing, the point of breakage in the fibril is nonetheless not correlated to the cross-seed interface since the fibrils show no propensity to break at these locations. This is in stark contrast to the control Proline-L17P variant which scores a value of $\eta = 0.5$ implying that the cross-seed interfaces act as defect sites at which the fibril

breaks. The trend of the Proline-L17P variant outscoring the other models and consistently breaking at the cross-seed interfaces is also observed in the other configurations (D2 and D3).

An interesting comparison can be made between the performances of the fibril models containing proline residues (Proline-L17P vs Proline-E22P). The L17P variants have the sequence substitution in the middle of the primary beta strand region of the U-shaped Aβ monomers. Incorporating this variant into the WT layers leads to a general defect like response when mechanically probed. The E22P sequence substitution meanwhile is situated in the loop region of the U-shaped Aβ monomer, much like most FAD related mutations. Incorporation of this variant into the WT fibrils generally leads to a mechanical stabilisation effect similar to that found in the Arctic variants (E22G). The differences between the responses of the variants with proline residues (L17P & E22P) highlight how both the chemistry and location of the substituted residue can play a role in attenuating the mechanical characteristics.

In the D2 configuration, where the cross-seed interface is located nearer to the pulled end of the fibrils, all models score η values greater than 0. This is correlated to peak forces that are either lower than or similar to the WT for the models that were previously mechanically stabilised in the D1 configuration (Arctic, Iowa & Proline-E22P). This indicates that for this probing direction, placement of the cross-seed interface close to the pulled fibril end had a net destabilising effect on the mechanical response in all cases.

In the D3 configuration, where the cross-seeding is similar to D1 but involves two successive layers of interfaced variant monomers, a contrasting set of responses are observed across the models. The Arctic variant has a similar response to that obtained in the D1 case with $\eta = 0$ and also a higher mean peak force relative to the WT. The Iowa variant however, sees a significant drop in peak force that is coupled to a value $\eta = 0.5$. The Proline-E22P model has a small $\eta = 0.08$ value coupled to peak force that is similar to that of WT. Proline-E17P scores $\eta = 0.75$ with an expected lower mean peak force relative to the wild-type model. Note that for the final cross-seed configuration, D4, every WT layer is always interfaced with a variant layer hence an η evaluation would not provide any meaningful insights.

The results from the sets of simulations as summarised in table 5.1 collectively demonstrate how the location of a sequence substitution in the peptide monomer and the cross-seeding placement relative to length of the wild-type fibril influences the mechanical characteristics in the probed directions.

5.4 Insights from homogeneous Aβ variants

In order to interpret the results presented in the previous section, a fundamental understanding of the structural changes and key interactions that are imposed by the variant monomers relative to WT is needed. In this section, homogenous fibril models made up exclusively of either Arctic, Iowa or Proline-L17P monomers were built and compared for structural, thermodynamic and mechanical properties. Studying these models might also provide further insights on the reasons why fibril formed from Familial Alzheimer's Disease mutants, such as Arctic and Iowa, have increased fibrillation kinetics and toxicity relative to WT.

5.4.1 Simulation Details

The construction of the homogeneous variants and their subsequent simulation details are identical to those already outlined in § 5.2. At the conclusion of the MD simulations, mechanical probing was done by the use of two distinct pulling geometries since the fibril responses are known to be anisotropic. The first of these is the tensile pulling SMD outlined earlier in this chapter. The second pulling geometry is in a direction that is independent of the hydrogen bond network and accounts for interactions between the three individual proto-fibrils that make up the 3-fold symmetric Aβ fibre. This pulling geometry is designed to probe the interactions between the FAD proto-fibrils by measuring the peak force required to extract one such proto-fibril from the whole fibre. Figure 5.6 shows a schematic representation of the pulling mode. The protocol requires the center of mass of backbone carbon-α atoms on one outer β-sheet face of a selected proto-fibril to be pulled in a direction perpendicular the fibre long axis whilst the carbon-α atoms the outer β-sheet faces of the other two proto-fibrils are held fixed.

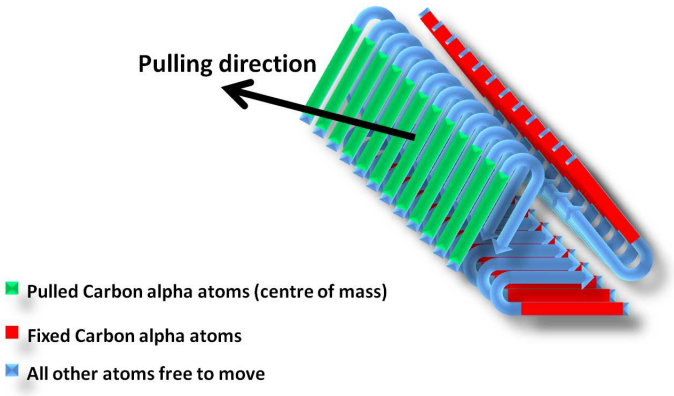

Figure 5.6: Schematic representation of SMD protocol designed to extract a single proto-fibril from the 3-fold symmetric Aβ fibre. The center of mass of carbon-α atoms in portions of monomers marked in green is pulled in the direction indicated by the arrow whilst the carbon-α in portions marked in red are held fixed. All other atom types are free to move unrestrained.

5.4.2 Structural properties and Thermodynamic Stability

Structural differences

The configuration of the homogeneous wild-type, Arctic, Iowa and Proline-L17P fibrils at the start and conclusion of the production MD phase are shown in figures 5.1a and 5.1b respectively. All the models adopt a twisted morphology relative to their initial structures. A summary of the structural properties measured in the final 10 ns of MD simulation is presented in table 5.2. Secondary structure analysis reveals that the Arctic fibril has less beta-strand content relative to either of the WT or Iowa counterparts. Within the bounds of error, the three models also have similar backbone hydrogen bond numbers, however the Iowa fibril has the least number of side-chain hydrogen bonds. Meanwhile, the structural trends found in the control Proline-L17P fibril are as to be expected. The homogeneous Proline-L17P fibril predictably has the highest content of non-beta-strand randomly oriented structural elements, $\approx 60\%$, coupled with the least number of backbone hydrogen bonds in any model.

Thermodynamic Stability

As a means of comparing the thermodynamic stability across the homogeneous fibril models, the MM-PBSA methodology was employed to compute enthalpy differences relative to the WT structure in the final 10 ns of converged MD trajectories (fig 5.7). One of the

Variant	RMSD	SS_{Beta}	SS_{Random}	$Hbond_{Back}$	$Hbond_{Side}$
Wild-type	1.06 ± 0.13	0.48 ± 0.01	0.48 ± 0.01	522.81 ± 9.91	44.33 ± 5.35
Arctic	1.22 ± 0.15	0.46 ± 0.01	0.50 ± 0.01	516.66 ± 9.53	42.10 ± 4.80
Iowa	1.42 ± 0.25	0.48 ± 0.01	0.48 ± 0.01	500.92 ± 10.91	28.48 ± 4.73
Aβ-L17P	1.32 ± 0.26	0.36 ± 0.01	0.61 ± 0.01	428.40 ± 9.71	48.02 ± 5.61

Table 5.2: Structural properties of the fibrils evaluated from snapshots in the final 10 ns of unrestrained MD. The summary includes mean value of the RMSD of backbone carbon-α atoms, fraction of secondary structure content (SS_{Beta} & SS_{Random} for beta strands and randomly orientated strands respectively) the and mean number of hydrogen bonds ($Hbond_{Back}$ & $Hbond_{Side}$ for backbone and side-chains respectively).

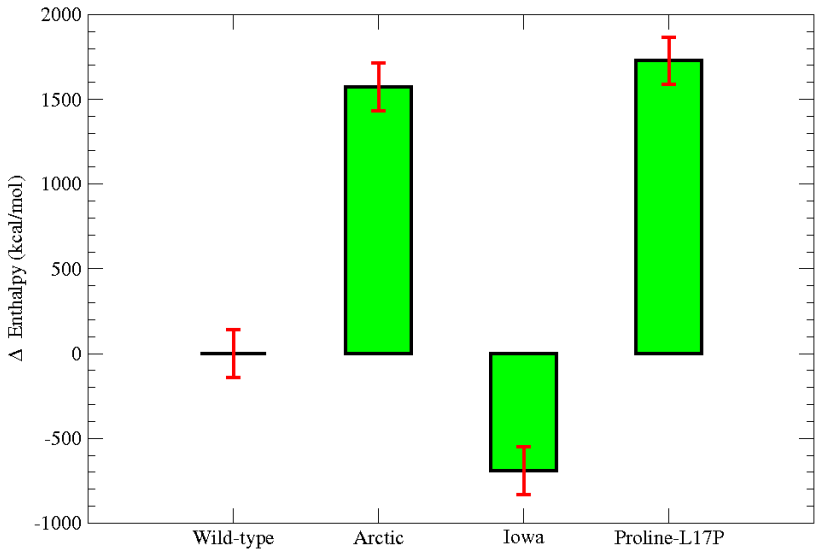

Figure 5.7: The differences in configurational enthalpy for the homogeneous fibrils plotted relative to the wild-type model.

most striking aspects of these calculations is that the Iowa fibril appears more enthalpically favourable than the wild-type. This is particularly surprising if one considers the effect of the D23N Iowa substitution. In the wild-type $A\beta_{(9-40)}$ sequence (fig 5.8), the negatively charged aspartic acid residue (D23) forms a salt-bridge with the positively charged lysine residue (K28). The presence the D23-K28 salt-bridge is known to provide an overall stabilising effect in Aβ fibrils (70; 122; 123). Since the Iowa substitution (D23N) acts to replace the aspartic acid residue with an uncharged asparagine, the resulting loss of the favourable salt-bridges throughout the length of the fibril should intuitively correspond to loss in stability relative to the WT. However, the relative enthalpy analysis for the Iowa fibril

shows a counter-view (fig 5.7). The Arctic and proline-L17P models are both significantly less enthalpically favourable in comparison to the WT model .

A clearer understanding of the calculated rank in enthalpy, especially in the case of the Iowa fibril, is reached when one closely examines the protonation states of the titratable residues and overall electrostatic interactions in each fibril model. The charge state for the titratable residues as assigned by the ProPka package are depicted in figure 5.9. A close inspection of figure 5.9 shows that each model has a multitude of columns of stacked layers of charged residues which are able to participate in both short and long-range electrostatic interactions. For the Iowa substitution (D23N), removal of the short range salt-bridge interactions may at first glance have an apparent destabilising effect when viewed at a local level. The individual energy terms contributing to the enthalpy calculations in figure 5.7 show however, that the Iowa fibril has the most favorable electrostatic interactions of any model. This suggests that the long range interactions between the different columns of charged residues around the fibrils (fig 5.9) become more favourable and provide a stabilising effect when the columns of negatively charged aspartic acid residues in WT are replaced with the uncharged asparagine layers as in the Iowa fibril. This has direct consequences for the mechanical characteristics, which are discussed further in § 5.4.3.

.

5.4.3 Mechanical response, Breakage events & Fibril fragments

The results discussed in this section, from the two independent pulling geometries used to probe the homogeneous models, again demonstrate the anisotropic nature of mechanical resistance in amyloid fibrils. The dominance of hydrogen bond interaction on the mechanical response in one pulling direction are contrasted to the dominance of electrostatic interactions in a perpendicular direction. A close examination of the breakage events and resultant fibril fragments left at the conclusion of the SMD provide of useful insights into the mechanical characteristics that might then explain the propensity for increased toxicity and enhanced fibril growth kinetics typically associated with FAD.

SMD results: Hydrogen bond numbers vs Long range electrostatics

The mean peak forces recorded from the tensile pulling SMD geometry, averaged from four repeat simulations, are summarised in figure 5.10. The results show that WT and Arctic

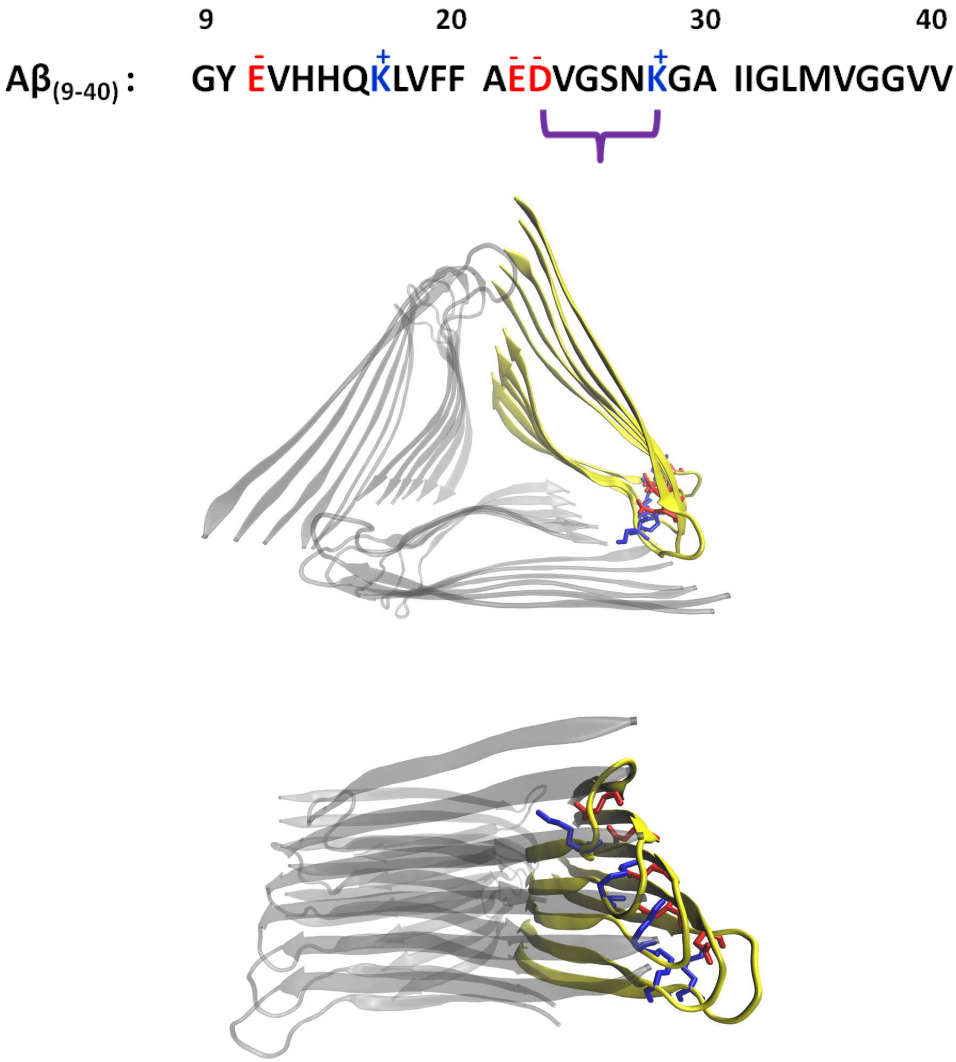

Figure 5.8: **SEQUENCE:** The residues that carry a net charge in the WT A$\beta_{(9-40)}$ sequence are highlighted in blue (positive) and red (negative). The aspartic acid residue in position 23 (D23) of the sequence is known to form a stabilising salt-bridge with the lysine in position 28 (K28) as represented by the purple link. **MODELS:** The location of salt bridges in the fibril layers of the WT ssNMR structure are shown for one proto-filament, for clarity, marked in yellow (lysine in blue and aspartic acid in red) in the top and side views of the model.

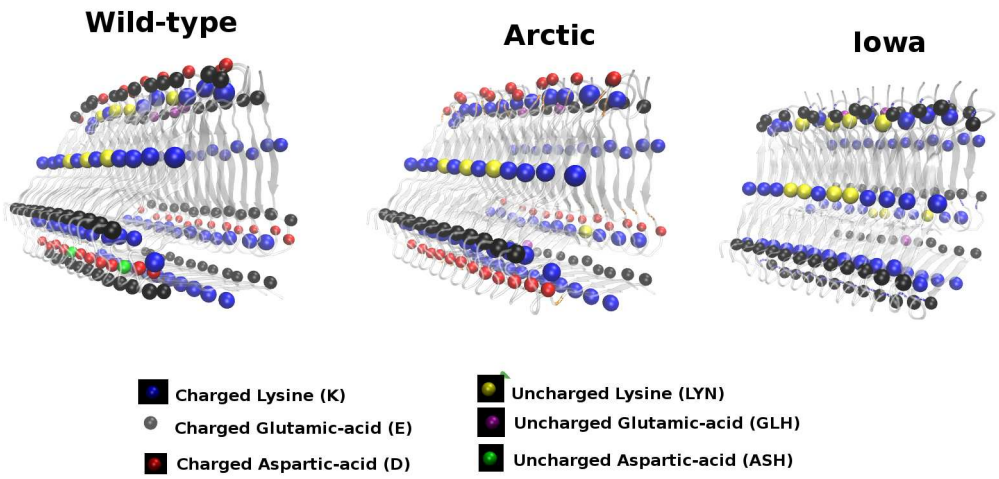

Figure 5.9: Assigned protonation states for the titratable residues in the WT, Arctic and Iowa fibrils respectively (pH = 7.5). The 3-fold symmetric fibril is shown in grey for clarity whilst the titratable residues are represented by the coloured balls whose charge state is denoted by the key.

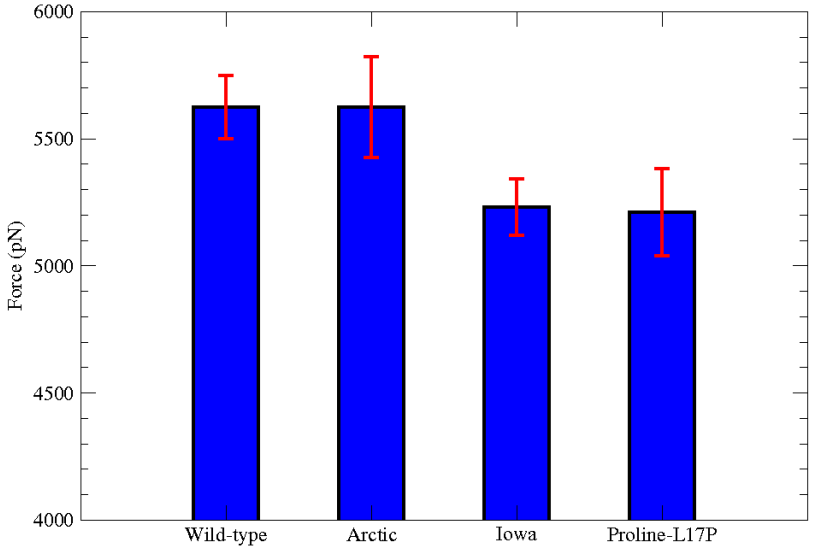

Figure 5.10: Mean peak force recorded from four repeat simulation of the tensile pulling SMD geometry for the homogeneous fibril models.

fibrils register similar peak forces in the tensile pulling simulations which are higher than those recorded by either Iowa or Proline-L17P fibrils. As this pulling geometry primarily addresses the hydrogen bond networks, it is not surprising that the trends in the rank of peak forces appear to reflect that of the total numbers of hydrogen bonds (table 5.2) in each fibril system. WT and Arctic models have the highest numbers of side-chain and backbone hydrogen bonds which are then able to cooperatively resist the SMD pulling in this direction with correspondingly high mean peak forces. The Iowa and E22P substitutions meanwhile result in a reduction of total number of hydrogen bonds hence these fibrils record lower mean peak forces relative to the WT. In as much as the hydrogen bond numbers clearly play an important role in the mechanical resistance of the fibrils in this pulling mode, they don't however fully account for different fragmentation mechanisms and trajectories followed by each fibril even in the cases where similar mean peak forces are recorded.

Although the WT and Arctic fibrils record virtually identical mean peak forces in the tensile pulling mode (5625.21 $\pm$ 124.66 & 5625.01 $\pm$ 199.70 respectively) the manner in which they resist the steering forces are quite distinct. Figure 5.11a shows that despite similar peak forces being reached, the details in the force-time profiles are different. The drop after peak force for the WT is relatively steep in comparison to that of the Arctic fibril which is more gradual. Figures 5.11b and 5.11c graphically demonstrate how the Arctic fibril suffers fractures at significantly more locations along its length relative to the WT fibril despite the similarity in measured peak forces. Furthermore, figure 5.12 shows the very distinct snapshots of the time evolution of WT and Arctic fibrils during the tensile pulling simulations.

The final structures and the detached fragments at the conclusion of the SMD are not the same for WT and Arctic fibrils (figure 5.12). The resistance mechanisms during SMD probing of the Arctic fibril involves peptides at the fracture sites being drawn out in a 'spaghetti-like' fashion. This is in stark contrast to the WT where a 'clean' break is achieved at the fracture site. This behaviour is very similar to that observed in § 4.4.3 where direct changes to the SNNFGAILSS fibril electrostatics resulted in the mechanical resistance mechanisms moving from a ductile to a more brittle like response (see fig 4.8). As in that case, the resultant changes to the electrostatics of Aβ fibrils due to the presence of FAD inspired substitutions here might also be playing a major role in determining the distinct

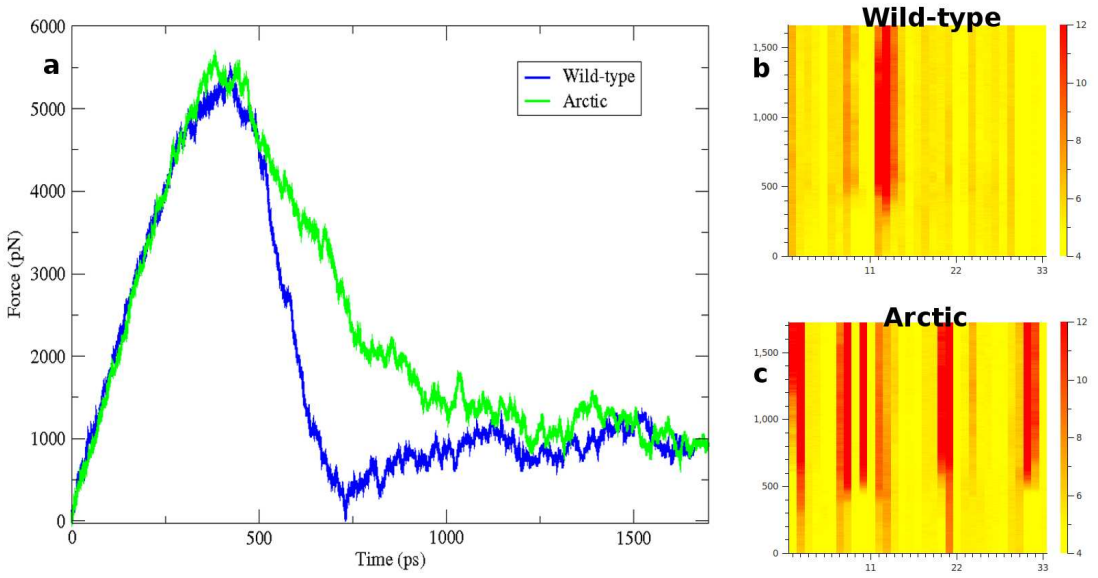

Figure 5.11: a) Representative force-time profiles for selected WT and Arctic fibrils during the tensile pulling mode. The fracture sites corresponding to the force profiles during the SMD are depicted in panel b) for WT and in panel c) for Arctic. The X-axis in these plot represents the distance between adjacent layers of monomers in the fibril while the Y-axis is SMD time in pico seconds. As there are 12 layers and 3-faces in each fibril model, the graduations on the X-axis are 1-11 for layers in the first face, 12-12 for layers in the second face and 13-33 for layers in the third face. The colour scale starts at 4.8 Å (yellow) and saturates at distances equal to or greater than 12 Å (red).

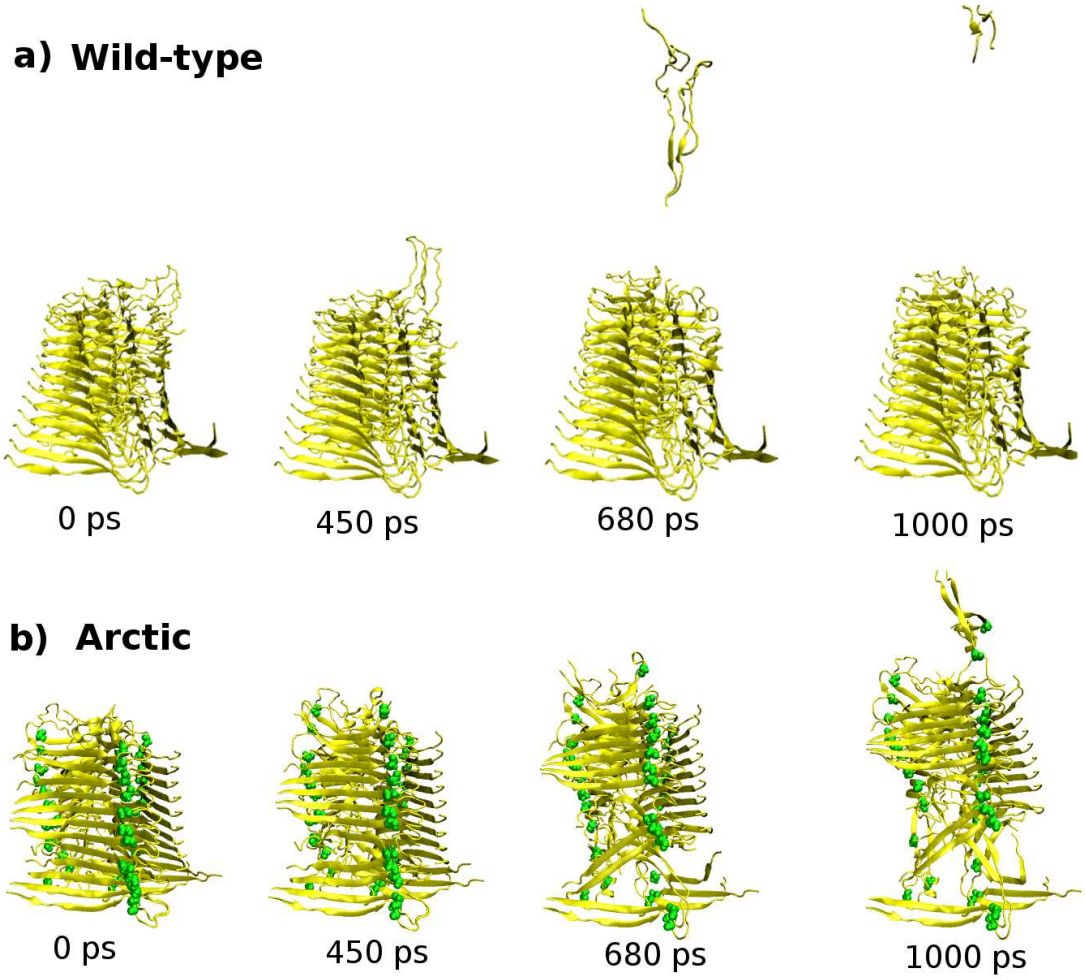

Figure 5.12: Representative snapshots of the time evolution of the a) WT and b) Arctic fibrils during tensile pulling SMD. The WT fibril suffers a clean break at the fracture site while Arctic is drawn out along its length when resisting the steering forces.

Homogeneous model	Proto-fibril Extraction $\langle pN \rangle$
Wild-type	4417.13 $\pm$ 32.14
Arctic	4612.97 $\pm$ 86.43
Iowa	4666.40 $\pm$ 118.00

Table 5.3: Mean peak forces recorded for proto-fibril extraction SMD averaged from four repeat simulations in each case (quoted with the standard error in the mean).

resistance mechanisms. As such, the electrostatic contributions to the mechanical resistance are directly addressed by the second SMD pulling mode with the results summarised in table 5.3. As this pulling geometry probes the non-bonded interactions between the three proto-fibrils that make each 3-fold symmetric fibril, the electrostatic characteristics in each model (fig 5.9) appear to be closely related to trends in mean peak forces (table 5.3). For instance, the Iowa fibril has the most favorable electrostatic energy and it is also the most resistant to the forced removal of one of its three proto-fibrils. Meanwhile, less mechanical resistance is encountered in WT fibril due to the long range net repulsive interactions between its three proto-fibrils in this pulling direction (fig 5.9).

The results from the two independent pulling geometries show that not only is the measured response in each fibril anisotropic but that the relative performance between the models also changes depending on the pulling direction. This coupled with the distinct fracture mechanisms in figure 5.12 again highlight how the deliberate selection of appropriate sequence substitutions during nano-technological design processes may be employed to induce a desired material property such either brittle or ductile behaviour in an amyloid inspired device.

The characteristics of the newly exposed fibril ends at the fracture sites and the structural details of the detached fragments obtained at the conclusion of SMD simulations (e.g fig 5.12) are particularly interesting as these could provide clues as to why the FAD Arctic mutation for instance leads to a more cytotoxic fibril that has increased fibrillation kinetics relative to WT. The interactions of the fractured fibril at the newly formed ends with cell membranes or other monomers might play a role in determining both the degree of toxicity

and the rate of polymerisation respectively. This assertion is supported by a very recent 3D cryoelectron tomography visualisation study of β_2-microglobulin fibrils interacting with lipid membranes which reveals that the 'ends' of the shorter mechanically fragmented fibrils enhance toxicity by directly attaching to then distorting and disrupting the membranes (33). It is therefore important to model and structurally characterise these fibril ends prior to and after their fragmentation.

Detailed analysis of the fractured fibril ends would require new simulations using significantly more layers of Aβ monomers than the 12 used here before any definitive conclusions about their role could be made since fibrils are typically long structures in the μm-range but the simulations discussed here only probe a relatively short segment of a typical fibril length due to computational cost. Nonetheless, figure 5.12 still offers some tantalising glimpses into a potential explanation for the clinically observed trends of FAD fibrils which should be pursued further in future work.

5.5 Conclusions

The main purpose of this investigation was to carry out a systematic evaluation of how defects that are deliberately engineered into model amyloid fibrils affect their mechanical character and whether such strategies could be adopted in the design or construction of amyloid inspired nano-materials. The choice of Aβ fibrils in this work and the subsequent use of sequence variants associated with familial Alzheimer's disease additionally meant that possible links between the frangibility of the models and their enhanced toxic effects could be investigated.

The heterologous interfaces between wild-type Aβ fibrils and the sequence variants derived from either Arctic, Iowa, E17P or E22P substitutions all demonstrated a measurable ability to significantly enhance or reduce the mechanical stability by attenuating the mean peak force recorded for the homogeneous WT fibril. The degree of mean peak force attenuation was directly dependent on the location of the heterologous interface relative to the pulled fibril end, the location (and chemistry) of the amino-acid substitution along the U-shaped monomer and also on the number of cross-interfaced sites along the fibril. In the cases where lower peak forces relative to WT were recorded, the interfaces acted as defect

sites since the fibrils always fractured at these locations. This was especially true for the control E17P fibril configurations with proline substitution in the middle of the beta-sheet region of the Aβ monomers as opposed to those in the loop region like Arctic, Iowa and E22P which saw mixed responses in resistance relative to the WT. Conversely, in the cases where higher peak forces were recorded, the heterologous interfaces were not involved in the fibril fracture suggesting they provided a mechanically stabilising effect for those configurations. The different responses could be traced to the changes in the electrostatic interactions since all the substitutions involved the replacement of a charged titratable side chain with a neutral one.

The role of the electrostatic interactions in attenuating the fibril mechanical characteristics was further investigated through the use of homogeneous models comprised entirely of variant monomers. It was found that despite losing multiple stabilising salt-bridges through the D23N substitution, the Iowa fibril was still significantly more enthalpically stable than WT largely due to the highly favourable long range electrostatic interactions associated with this change as a whole around the highly charged fibril. Changes in the fibril electrostatics have been shown to switch fibrils from a ductile to a more brittle mechanical response in SNNFGAILSS fibrils (Chapter 4) and a similar behaviour in the fracture mechanisms of the homogeneous Aβ fibril was observed in this study. The structure of the distinct fibril fragments obtained due to the differences in these breakage mechanisms at the conclusion of the SMD might, with further analysis, provide clues as to why fibrils associated with FAD have enhanced toxicity or increased polymerisation kinetics.

The two independent pulling geometries used here produced an anisotropic response within each fibril model and the relative rank in mechanical stability amongst the different fibrils was dependent on how the pulling mode impacted an intrinsic property such as hydrogen bond density or electrostatics along that direction. As such, the relative rank between the homogeneous fibrils saw WT record the highest mean peak force for the tensile probing simulations by virtue of having the most hydrogen bonds whilst Iowa recorded the highest mean peak force in the second SMD mode by virtue of having the most favourable electrostatic interactions between its three proto-fibrils.

Overall, these simulations demonstrate how engineering defects into amyloid fibrils through

cross-seed interfaces can be used as an effective means for modulating their mechanical properties particularly in the construction of tailored nano-materials and nano-devices. The careful selection of appropriate sequence substitutions can produce a more malleable or rigid amyloid based nano-material as desired. The results of these simulations also present useful lines of enquiry for future work in understanding the origins of enhanced toxicity and polymerisation kinetics found in FAD related fibrils based on the differences in the fibril fracture mechanisms.

Chapter 6

Conclusions & Future Work

6.1 Conclusions

The main aims of the work presented in this thesis were to determine the factors that are instrumental in the mechanical resilience generically exhibited by all amyloid fibril systems and then examine ways in which they may be modulated. Computational models were mechanically probed through the application of steering force designed to disrupt the stabilising fibril interactions from different directions. The work presented in chapters 3, 4 and 5 each investigate different aspects that were thought to be of significance to the observed mechanical stability of these systems. The conclusions drawn from these studies are discussed in detail at the end of each relevant chapter. Here, a summary of each of these findings is given.

Role of peptide sequence in mechanical resistance of hIAPP$_{20-29}$ protofibrils

In chapter 3, SMD simulations were used to probe the mechanical response of a series of amyloid fibril models built from peptide sequences chosen to have properties intermediate between the highly amyloidogenic human amylin fragment SNNFGAILSS and its non-amyloidogenic rat counterpart. The four pulling geometries used were sufficiently sensitive to discriminate the responses between the fibrils even in cases with minor sequence modifications. The simulations reveal that the force required to induce mechanical failure depends on the direction of the applied stress and upon the degree of structural order present in the β-sheet assemblies, which in turn depends on the peptide sequence. The mechanical responses were highly anisotropic, with the wild-type fibrils consistently recording higher mean peak forces relative to the other variant models. The dense hydrogen bond networks exhibited a cooperative resistance when the pulling direction ran parallel to them. However, a key finding was that defects due to losses in inter-peptide backbone hydrogen bonding were able to dominate the responses as they acted as weak points prone to fragmentation on the application of force. The results have wider implications for the importance of sequence-dependent mechanical properties such as the role of breakage events in cytotoxicity and on fibrillation growth kinetics due to seeding.

Modulating mechanical failure in polymorphic fibril arrangements

In chapter 4, an examination was made into how the different steric packing arrangements of peptides found in amyloid polymorphs contribute to their mechanical stability. A series of SMD simulations were performed to mechanically probe three polymorphs of SNNFGAILSS fibrils and the responses contrasted at different lengths & terminal capping modifications. Although the polymorphs all displayed an anisotropic response to the pulling geometries used, there was however, a significant difference in the relative responses between the polymorphs when compared across the individual pulling geometries. The mechanical responses of fibrils were found to be directly related to the details of the arrangements of peptides in the β-sheets, with the observed strength in each polymorph dependent on how the pulling geometry impacts on the intrinsic properties that arise due to the peptide packing and also on the number of defects present within the models. The effectiveness of terminal capping as a mechanical modulator also varied amongst the three polymorphs. It was possible in some cases to change mechanical behaviour from ductile to brittle thus demonstrating that minor inexpensive terminal modifications could be potentially used as means to attenuate the mechanical characteristics in a non-trivial way for nanotechnology inspired applications.

Engineering defects into fibrils to alter mechanical characteristics

Chapter 5 explored whether heterologous cross-seeding interfaces in mature Aβ amyloid fibrils could be exploited to attenuate the mechanical characteristics in a manner that might be useful in the design and modulation of nano-materials. This work was partly inspired by results in chapters 3 and 4 which highlighted how strongly structural defects are able to dominate the fragmentation mechanism in the fibrils. Comparison of the mean peak forces extracted from tensile pulling SMD show measurable differences in the mechanical responses of the wild-type relative to the cross-seeded fibril models. There was a clear correlation between the location of the cross-seed interface along the fibril length and the degree of peak force attenuation relative to the wild-type Aβ fibril. Some cross-seed configurations enhanced mechanical resistance whilst others were able to significantly reduce it thus making this a viable strategy for mechanical modulation in amyloid based nano-material design. Changes in the electrostatic characteristics, due to the sequence substitutions, were able to switch the fibrils fracture mechanisms from ductile to malleable behaviour.

6.2 Future Work

The results of the work presented in chapters 3, 4 and 5 of this thesis all offer up a number of possible avenues for future investigations. This section proposes additional research that might not only enhance the findings in those chapters but also offer broader insights into the fragmentation process of amyloid fibrils. A common trend prevalent in all the investigations described here was the role of structural defects in dominating the mechanical behaviour of the otherwise highly ordered fibril aggregates. Further work is needed to hierarchically rank the importance and prevalence of different types of structural defect in terms of how they then influence the overall fracture mechanisms of the fibrils.

Clues as to why fibril fragmentation leads to enhanced cyto-toxicity and fibrillation kinetics can possibly be found from analysis of the structure of the fibril fragments since their surface interactions are likely responsible for these characteristics. Due to computational costs associated with the full atomistic representation however, only relatively short lengths of fibrils were mechanically probed in this work. This makes analysis and structural characterisation of the fibril fragments after SMD, particularly their edges, difficult to interpret as the breakage of a much longer fibril model is likely to give a more accurate representation of the details at the fracture site than a short fibril would. It is therefore instructive to model significantly longer fibrils than presented here in order to assess what surface characteristics on the edges of their fragments are responsible for their ability to induce increased cyto-toxicity or fibrillation kinetics.

Obtaining thermodynamic information, such as changes in the free energy landscape as the mechanical probing simulations progress, would enhance the understanding of key interactions involved in the fracture mechanisms of the fibrils. This would require the use of alternative computational approaches like the introduction of a level of coarse-graining that would then enable the generation of statistically significant repeat SMD simulations for PMF reconstruction via the Jarzynski relation. Another approach might be to run a combined series of SMD simulations at multiple temperatures using multiple pulling speeds in order to asses how the mechanical strength of the fibrils might be determined by the choreography of unfolding events. Complementary information to the work presented here might also be extracted from running the simulations using the constant force SMD method.

In chapter 3, the rat-IAPP inspired sequence substitutions were made only on the parallel SNNFGAILSS polymorph. However, the ssNMR experiments find that this polymorph coexists in thermal equilibrium with an anti-parallel counterpart. It would be prudent to assess whether identical substitutions on the anti-parallel polymorph motif also yield similar trends in the rank of mechanical resistance.

In chapter 4, the work on three SNNFGAILSS polymorphs could be expanded to include the other five possible Eisenberg steric zipper classes of the peptide sequence. This would allow for wider conclusions to be drawn on the role played by the peptide packing arrangements in determining the mechanical characteristics of the fibrils. Additionally, the trends from fibrils of other short peptide sequences could be assessed and incorporated to give a more general picture. These could include fibrils formed from sequences such as GNNQQNY, SSTSSAA, VEALYL and GGVVA whose thermodynamic stability across all eight steric classes has been determined (21).

In chapter 5, the ssNMR derived Aβ_{9-40} fibril model has the first eight residues missing in each monomer since these could not be resolved in the structural determination experiments. Although these residue are not thought to be in a beta-strand conformation, their interactions could nonetheless play a crucial role in the mechanical stability of the fibrils, especially given that they are predominately composed of charged amino-acid groups. Thus for an accurate understanding of the mechanical resilience of Aβ fibrils, future investigation should consider modelling the fibrils with the missing residues rationally added.

A final proposal for future work stemming from chapter 5 is the design of additional pulling geometries more suited to the 3-fold symmetric structure found Aβ fibrils that are able to probe the fibrils through bending and twisting motions. These new simulations could be designed by adapting the constant torque SMD method through careful scripting in the NAMD software package. This will provide an opportunity to interrogate the stabilising inter-peptide hydrogen bond interactions in a different manner than could be achieved by the pulling geometries used for the homogeneous FAD associated fibrils.

Appendix A

Force Response for sequence variants:

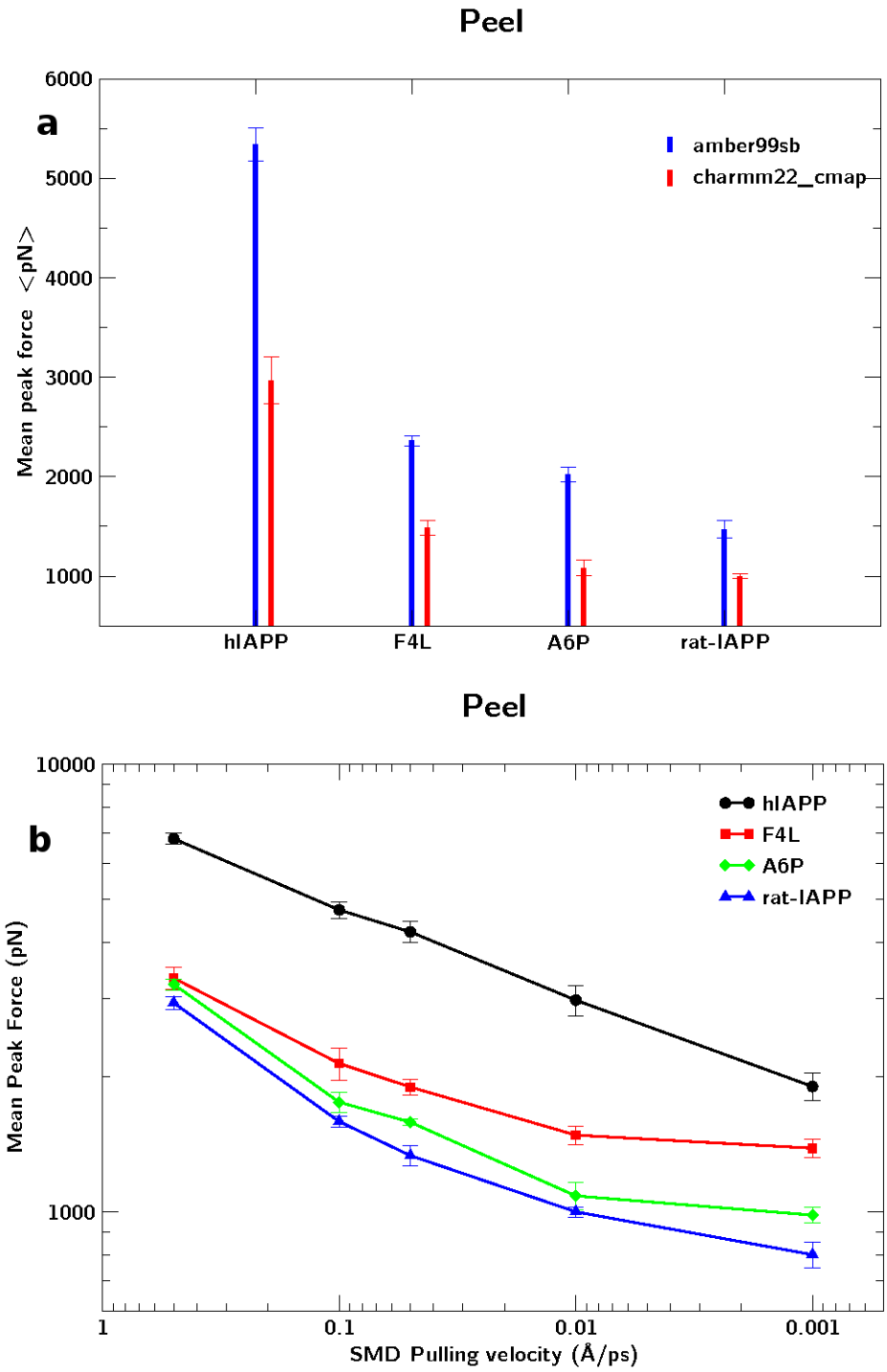

Figure A.1: (a) Comparison of the peel SMD simulation performed using the two different forcefields. Error bars are from averaging the mean peak force from four repeat simulations in each case. (b) The peel SMD simulations run at five different pulling velocities with four repeats used in each case.

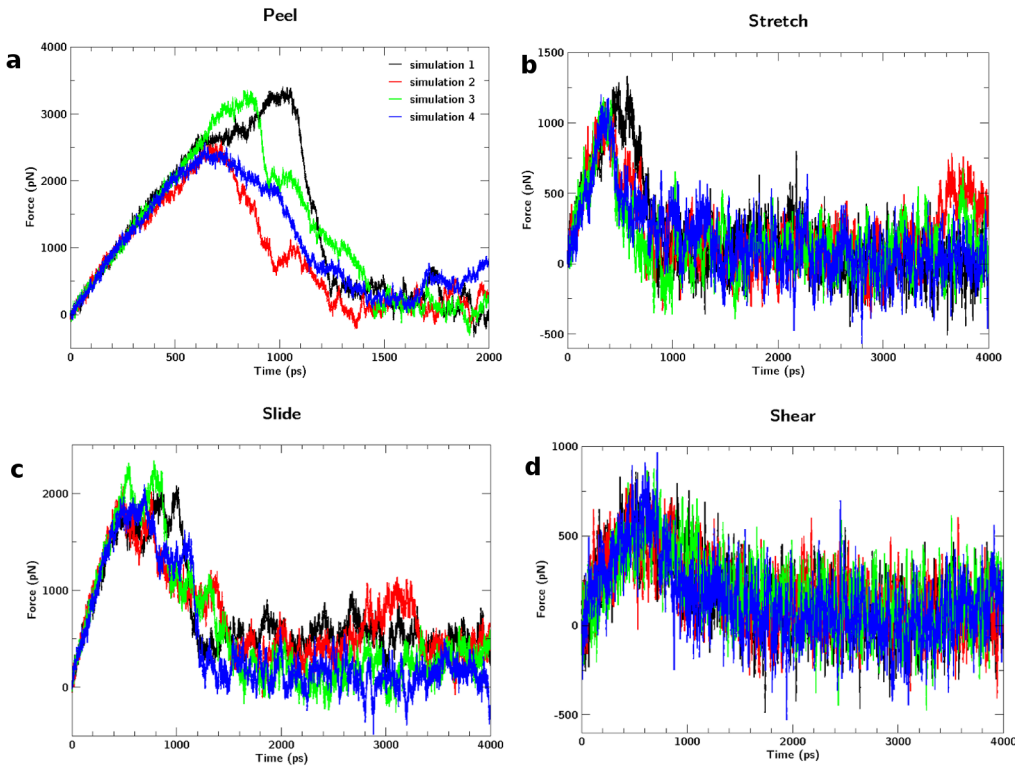

Figure A.2: Force time profiles for the wild-type sequence fibrils as the four different deformations are applied. The colours black, red, green and blue represent four repeat trajectories carried out for each type of deformation. (a) Peel (b) Stretch (c) Slide (d) Shear.

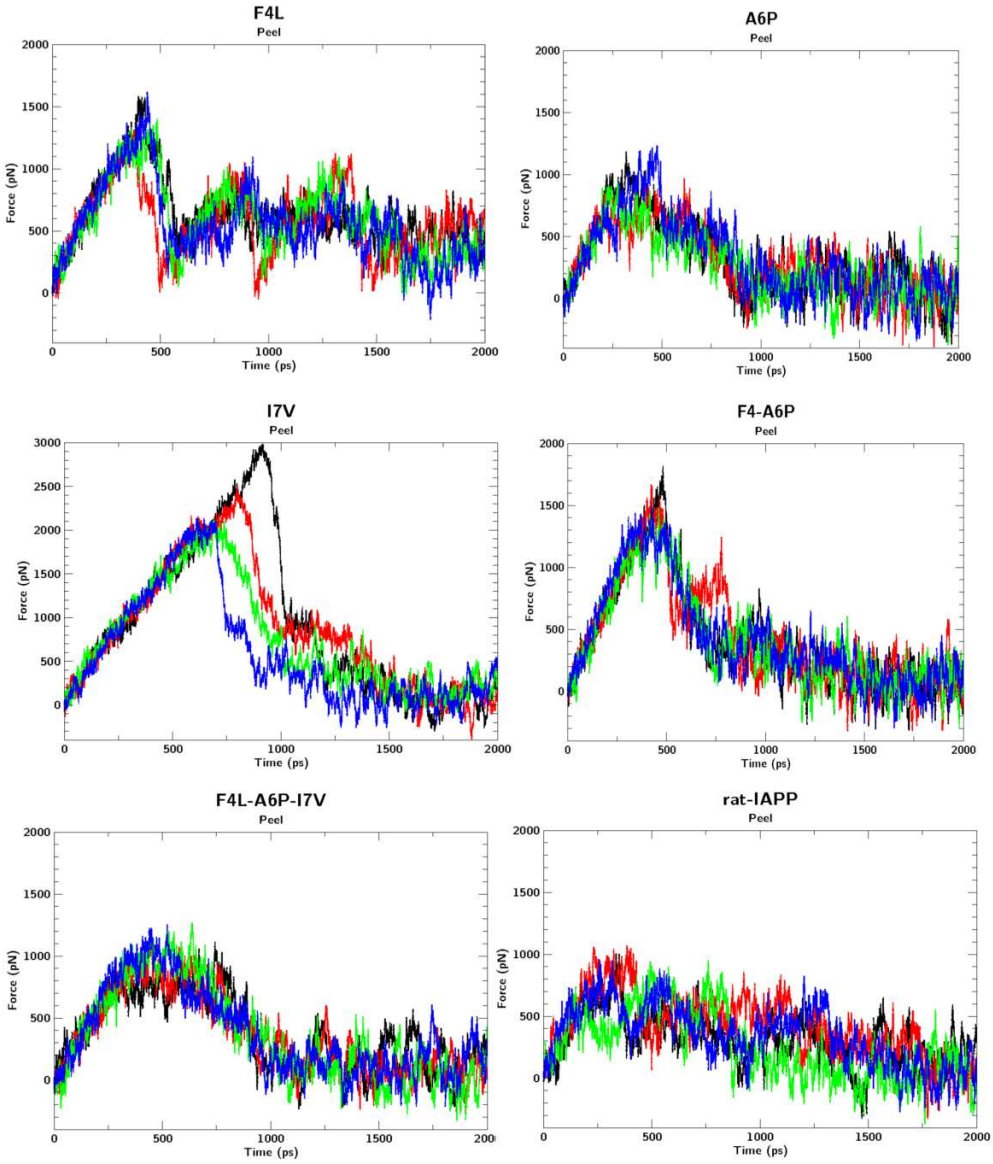

Figure A.3: Force-time profiles during peel simulations for the six model fibril variants. The red, blue, black and green graphs in each plot represent each of the four repeat simulations.

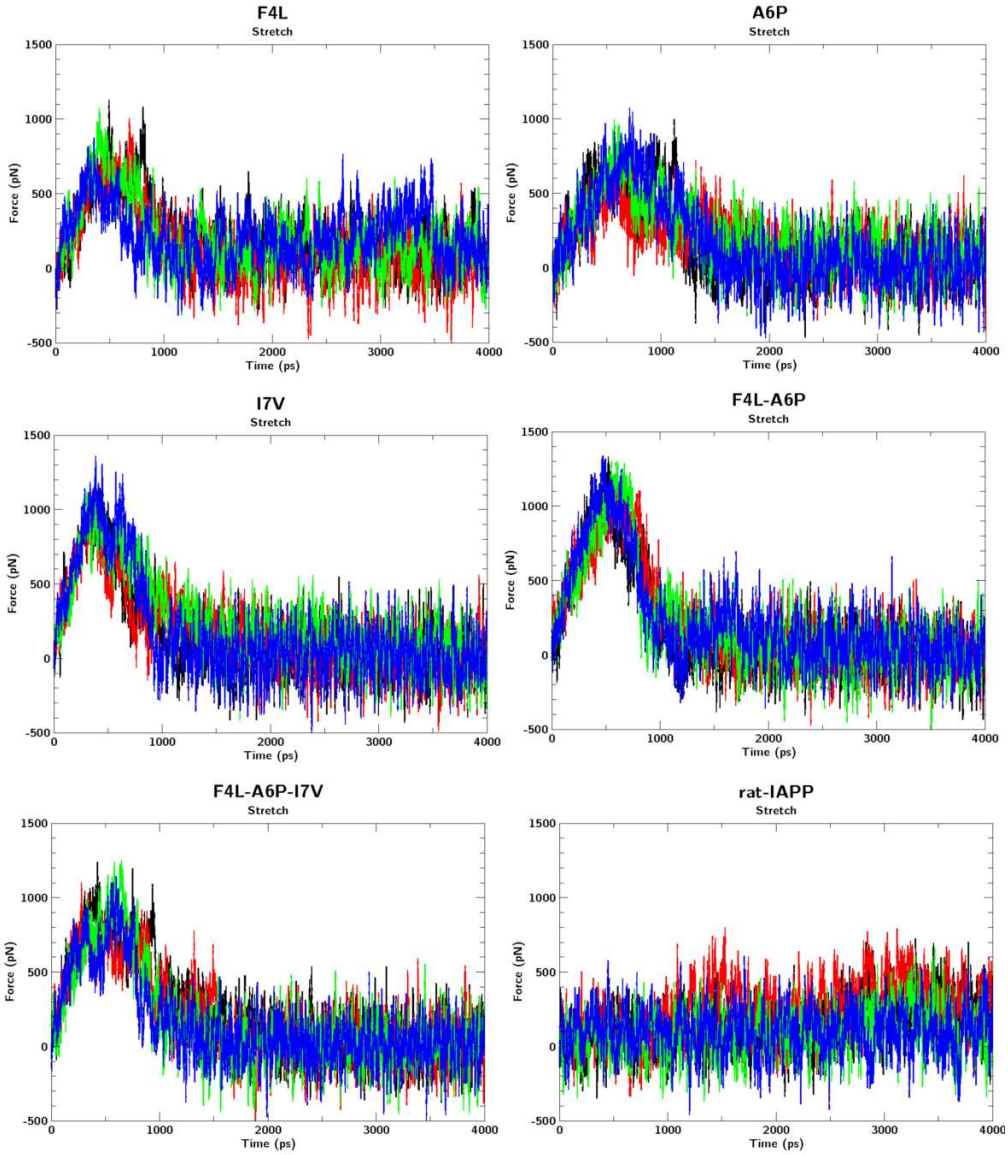

Figure A.4: Force-time profiles during stretch simulations for each of the six model fibril variants. The red, blue, black and green graphs in each plot represent each of the four repeat simulations.

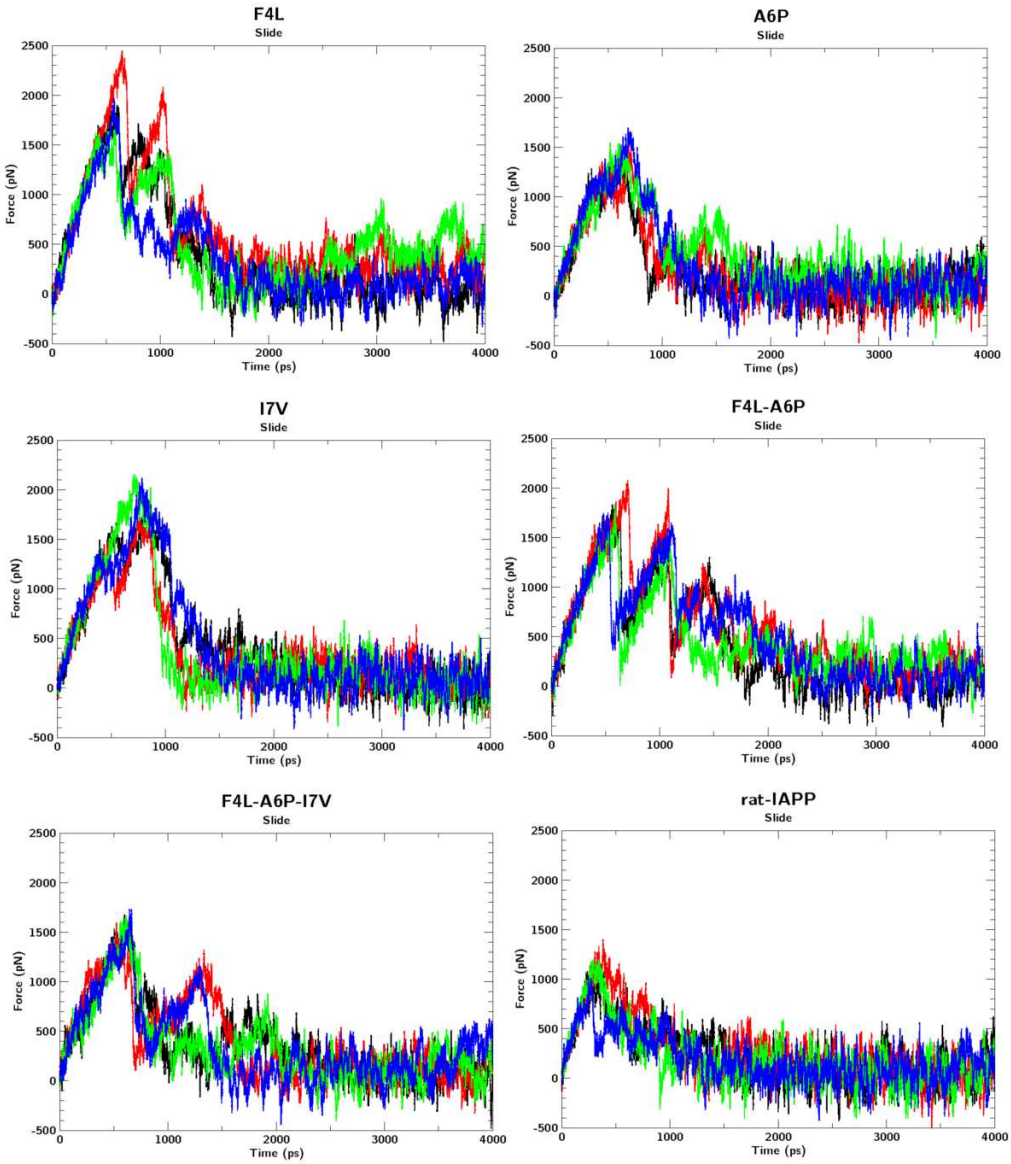

Figure A.5: Force-time profiles during slide simulations for the six variants. The red, blue, black and green graphs in each plot represent the four repeat simulations

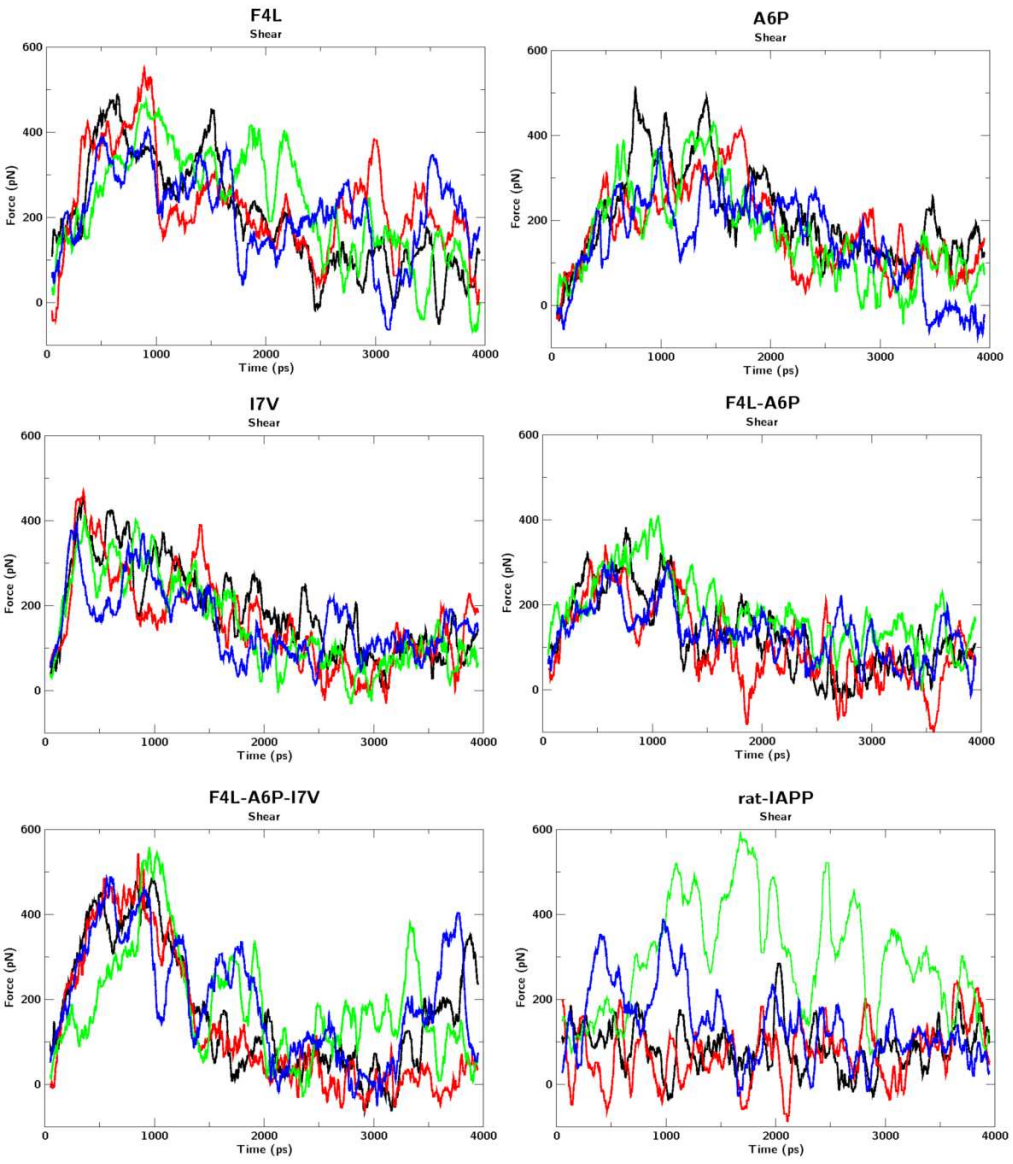

Figure A.6: Force-time profiles during shear simulations for the six variants. Due to noise in the signal, the force profiles for the shear simulations have been post-processed with a moving median filter considering periods of 100 ps. The red, blue, black and green graphs in each plot represent the four repeat simulations.

Appendix B

Details for fibril polymorphs:

136 SMD simulations:

Polymorph	Size	N-C Terminus	Deformation	# of Sims
Class1-P	8x2	free	Peel, Stretch, Slide, Shear	16
	8x2	capped	Peel, Stretch, Slide, Shear	16
	16x2	free	Peel, Stretch	8
	16x2	capped	Peel, Stretch	8
*	16x2	free	Peel, Stretch	8
Class2-P	8x2	free	Peel, Stretch, Slide, Shear	16
	8x2	capped	-	-
	16x2	free	Peel, Stretch	8
	16x2	capped	Peel, Stretch	8
*	16x2	free	Peel, Stretch	8
Class6-AP	8x2	free	Peel, Stretch, Slide, Shear	16
	8x2	capped	Peel, Stretch, Slide, Shear	16
	16x2	free	Peel, Stretch	8
	16x2	capped	Peel, Stretch	8
*	16x2	free	Peel, Stretch	8

Table B.1: Summary of SMD simulations to mechanically probe the fibril models run in NAMD with Charmm22-cmap forcefield. The constant velocity method was used in all cases with a pulling velocity of 0.01Å/ps for the free terminus models whilst the capped models used 0.04Å/ps . For direct comparison, the free terminus models marked with * also used a pulling velocity of 0.04Å/ps.

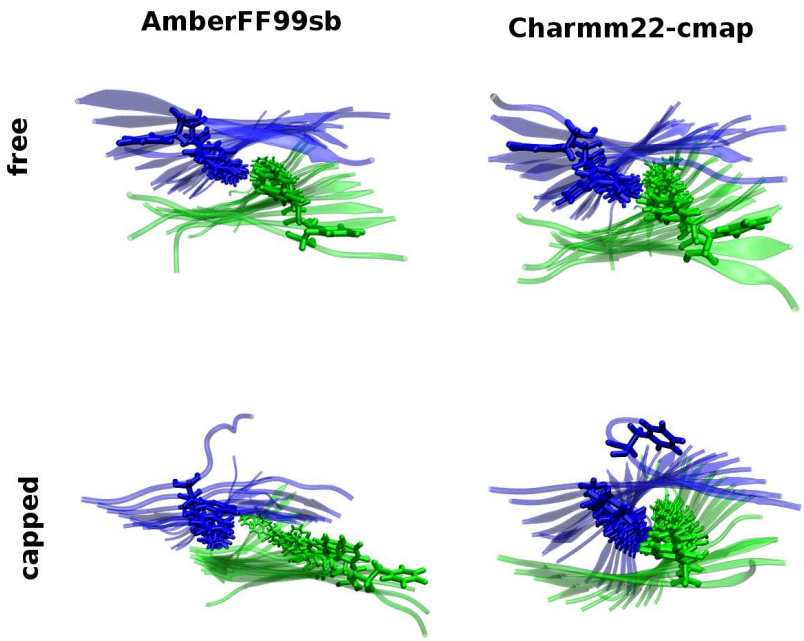

Figure B.1: Class1-P fibril models viewed from above the peptide axis. Positions and alignment of phenylalanine in the hydrophobic core are used as an indicator of the relative packing adopted by the peptides in the two forcefields. Similar packing is observed between the free models. For capped models, the fibril run in AmberFF99sb model has a noticeable shift between its β-sheets as the Phe residues are further apart whereas the capped fibril run in Charmm22-cmap has a more twisted appearance.

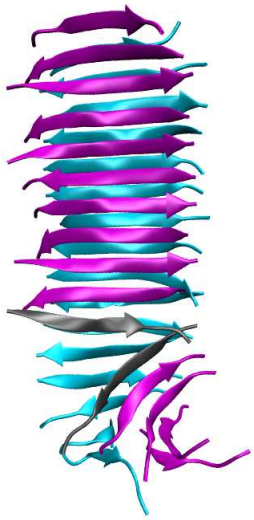

Figure B.2: Molecular conformation of the capped Class6-AP fibril prior to SMD simulation. The misaligned peptides strands on the purple β-sheet, highlighted in grey, are significantly splayed and have reduced inter-peptide hydrogen bonding. This structural defect site is the first point of failure for the leading to the reduced peak force for this fibril model.

Appendix C

Cross-seeded fibril properties:

The results for the RMSD, Secondary structure content, enthalpy and backbone hydrogen bonding for the cross-seed models, calculated from the final 10 ns of the converged MD simulations are presented in plots below. The nomenclature used in the plots is as follows:

- Homogeneous Arctic fibril: Arctic
  Arctic cross-seed D1 configuration: a-D1-cross-seed ...etc
- Homogeneous Iowa fibril: Iowa
  Iowa cross-seed D1 configuration: i-D1-cross-seed...etc
- Homogeneous Proline-L17P fibril: Proline-L17P/Physical-PRO
  Proline-L17P cross-seed D1 configuration: p-D1-cross-seed...etc

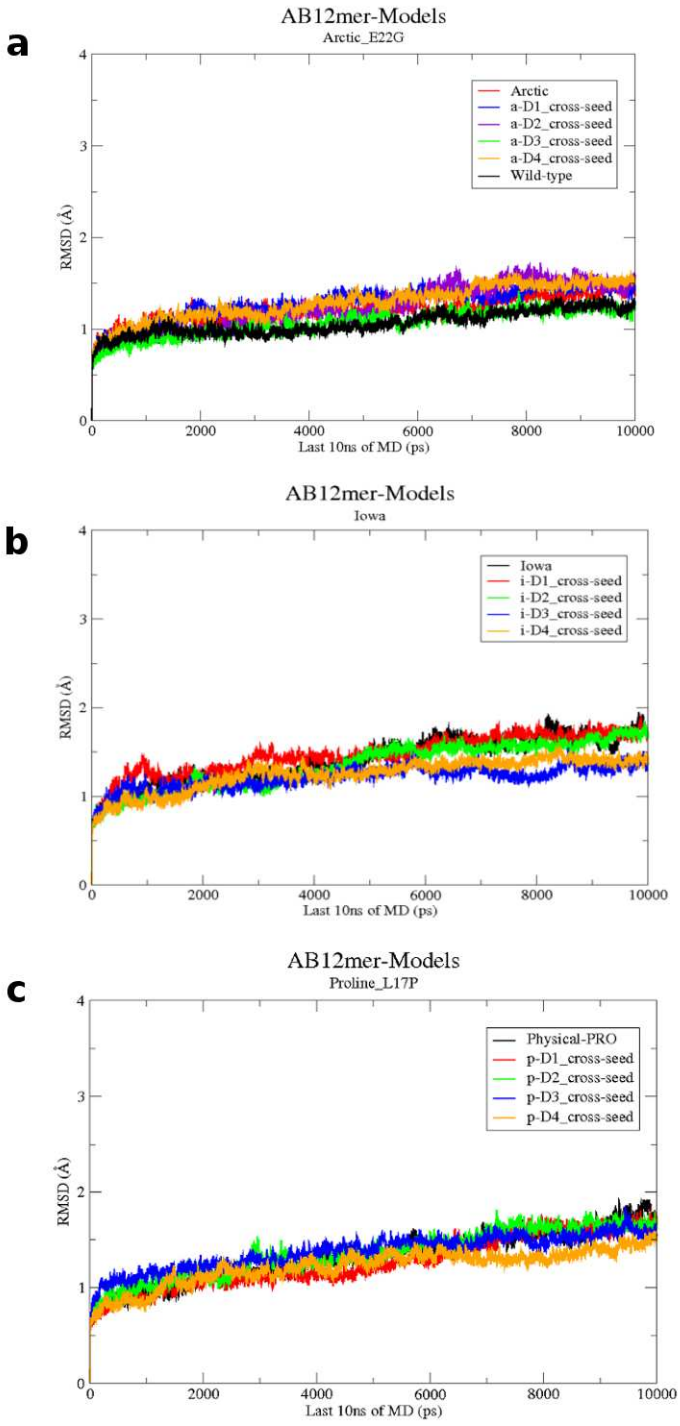

Figure C.1: RMSD values for the final 10 ns of ≈ 40 ns unrestrained molecular dynamics. (a) Wild-type and Arctic cross-seeds, (b) Iowa cross-seeds and (c) Proline-L17P cross-seeds.

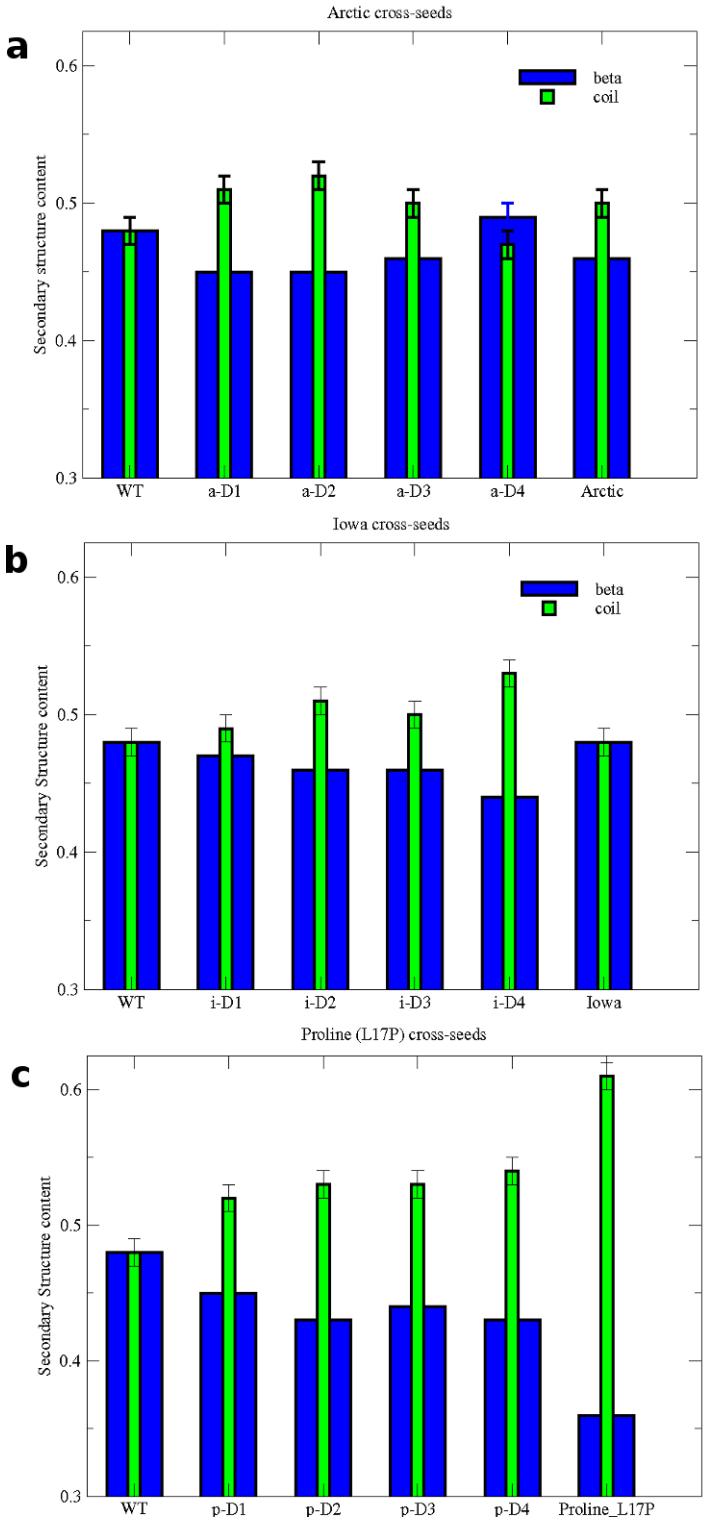

Figure C.2: Mean secondary structure content values for the final 10 ns of $\approx$ 40 ns unrestrained molecular dynamics. (a) Wild-type and Arctic cross-seeds, (b) Iowa cross-seeds and (c) Proline-L17P cross-seeds.

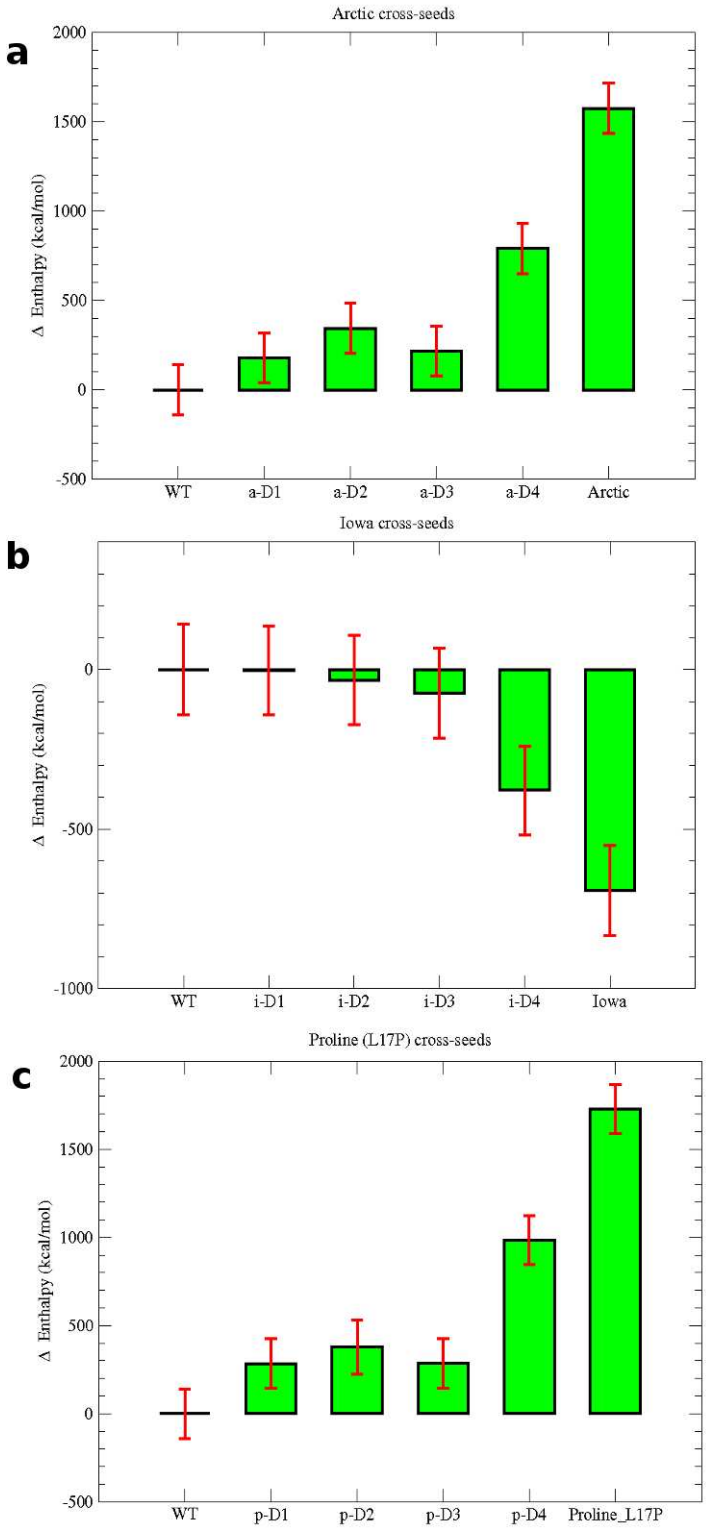

Figure C.3: Mean enthalpy for the final 10 ns of ≈ 40 ns unrestrained molecular dynamics. (a) Wild-type and Arctic cross-seeds, (b) Iowa cross-seeds and (c) Proline-L17P cross-seeds.

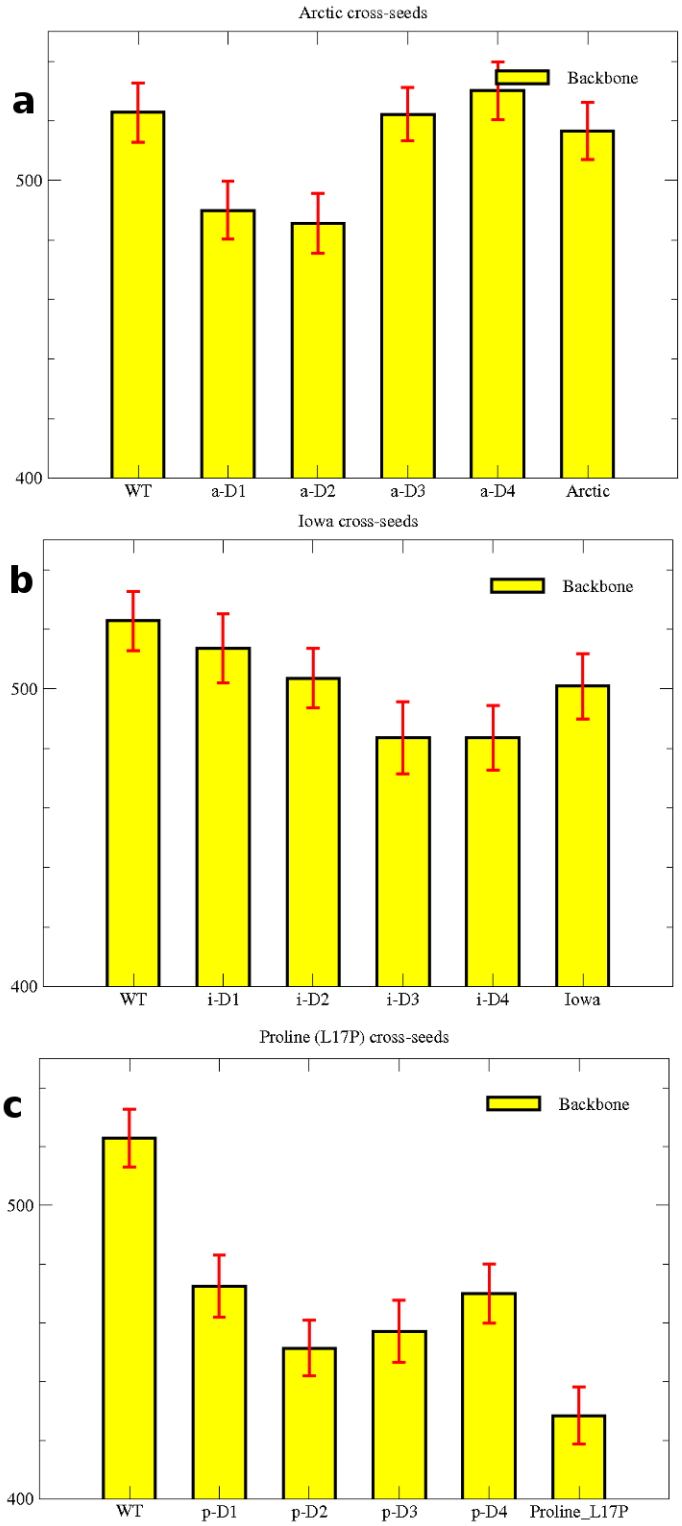

Figure C.4: Mean number of backbone hydrogen bonds for the final 10 ns of $\approx$ 40 ns unrestrained molecular dynamics.(a) Wild-type and Arctic cross-seeds, (b) Iowa cross-seeds and (c) Proline-L17P cross-seeds.

Bibliography

[1] Henry Stewart Talks. *Protein folding, aggregation and design : concepts, experiments, theories and mechanisms.* electronic resource, 2006. 2, 3, 6, 7, 8

[2] R. M Murphy and A.M Tsai. *Misbehaving proteins: protein (mis)folding, aggregation, and stability.* Springer,New York, 2006. 2, 4, 5

[3] T.R Jahn and S.E Radford. Folding versus aggregation: Polypeptide conformations on competing pathways. *Archives of Biochemistry and Biophysics*, 469(1):100–117, 2008. 2, 3, 4, 5, 7, 8, 52

[4] R.H Pain. *Mechanisms of protein folding*. Oxford University Press,2nd ed, 2000. 2, 4, 7, 52

[5] C M. Dobson. Protein folding and misfolding. *Nature*, 426(6968):884–890, 2003. 3, 10

[6] S Auer, C.M Dobson, and Vendruscolo M. Characterization of the nucleation barriers for protien aggregation and amyloid formation. *HFSP Journal*, 1(2):137–146, 2007. 3, 5

[7] F Chiti and C.M Dobson. Protein misfolding, functional amyloid and human disease. *Annu. Rev. Biochem*, 75:333–366, 2006. 4, 10, 11, 18

[8] C M. Dobson. The structural basis of protein folding and its links with human disease. *Philos Trans R Soc Lond B Biol Sci*, 35(1406):133–145, 2001. 4, 10

[9] A Aguzzi and T O'Connor. Protein aggregation diseases: pathogenicity and therapeutic perspectives. *Nat Rev Drug Discov*, 9:237–248, 2010. 4, 11

[10] S.R Collins, Adam Douglass, R.D Vale, and J.S Weissman. Mechanism of prion propagation: Amyloid growth occurs by monomer addition. *PLoS Biol*, 2(10):e321, 09 2004. 5, 6

[11] W Xue, S.W Homans, and S.E Radford. Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. *Proceedings of the National Academy of Sciences*, 105(26):8926–8931, 2008. 5

[12] J Shorter and S Lindquist. Hsp104 catalyzes formation and elimination of self-replicating sup35 prion conformers. *Science*, 304(5678):1793–1797, 2004. 6

[13] D Eisenberg and M Jucker. The amyloid state of proteins in human diseases. *Cell*, 148(6):1188–1203, 2012. 6, 11

[14] K.E Marshall and L.C Serpell. Structural integrity of beta-sheet assembly. *Biochem Soc Trans.*, 37(4):671–676, 2009. 7, 8

[15] S.K Maji, L Wang, J Greenwald, and R Riek. Structure-activity relationship of amyloid fibrils. *FEBS letters.*, 583:2610–2617, 2009. 7

[16] J Madine, E Jack, P G Stockley, S.E Radford, L.C Serpell, and D.A Middleton. Structural insights into the polymorphism of amyloid-like fibrils formed by region 20-29 of amylin revealed by solid-state nmr and x-ray fibre diffraction. *J. Am. Chem. Soc.*, 130:14990–15001., 2008. 8, 10, 35, 83, 84

[17] R Tycko, K L Sciarretta, P.R.O.J Orgel, and S C Meredith. Evidence for novel b-sheet structures in iowa mutant b-amyloid fibrils. *Biochemistry*, 48:6072–6084, 2009. 8, 106

[18] A Aguzzi, M Heikenwalder, and M Polymenidou. Insights into prion strains and neurotoxicity. *Nat Rev Mol Cell Biol.*, 8(7):552–561, 2007. 8

[19] R Nelson, M.R Sawaya, M Balbirnie, A.O Madsen, C Riekel, R Grothe, and D Eisenberg. Structure of the cross-b spine of amyloid like fibrils. *NATURE*, 435, 2005. 9, 53

[20] M. R. Sawaya, S. Sambashivan, R. Nelson, M. I. Ivanova, S. A. Sievers, M. I. Apostol, M. J. Thompson, M. Balbirnie, J. J. W. Wiltzius, H. T. McFarlane, A. O. Madsen,

C. Riekel, and D Eisenberg. Atomic structures of amyloid cross-b spines reveal varied steric zippers. *NATURE*, 447, 2007. 9, 53

[21] J.T Berryman, S.E Radford, and S.A Harris. Systematic examination of polymorphism in amyloid fibrils by molecular-dynamics simulation. *Biophys J*, 100:2234–2242, 2011. 10, 93, 131

[22] J.T Nielsen, M Bjerring, M.D Jeppesen, R.O Pedersen, J.M Pedersen, K.L Hein, T Vosegaard, T Skrydstrup, D.E Otzen, and N.C Nielsen. Unique identification of supramolecular structures in amyloid fibrils by solid-state nmr. *Angew.Chem.Int. Ed.*, 48:2118–2121, 2009. 10, 83

[23] M Fandrich and C M. Dobson. The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *The EMBO Journal*, 21:5682–5690, 2002. 10, 52

[24] C M. Dobson. Protein misfolding, evolution and disease. *Trends in Biochemical Sciences*, 24(9):329 – 332, 1999. 10

[25] D.M Walsh, I Klyubin, J.V Fadeeva, M.J Rowan, and D.J Selkoe. Amyloid-beta oligomers: their production, toxicity and therapeutic inhibition. *Biochem Soc Trans*, 30:552–557, 2002. 11

[26] C.G Glabe. Structural classification of toxic amyloid oligomers. *Journal of Biological Chemistry*, 283(44):29639–29643, 2008. 11

[27] L Pieri, K Madiona, L Bousset, and R Melki. Fibrillar a-synuclein and huntingtin exon 1 assemblies are toxic to the cells. *Biophysical journal*, 102(12):2894–2905, 2012. 11

[28] Y.J Lee, R Savtchenko, V.G Ostapchenko, N Makarava, and I.V Baskakov. Molecular structure of amyloid fibrils controls the relationship between fibrillar size and toxicity. *PLoS ONE*, 6:e20244, 2011. 11, 79

[29] I.C Martins, I Kuperstein, H Wilkinson, E Maes, M Vanbrabant, W Jonckheere, P Van Gelder, D Hartmann, R D'Hooge, B De Strooper, J Schymkowitz, and F Rousseau. Lipids revert inert abeta amyloid fibrils to neurotoxic protofibrils that affect learning in mice. *EMBO J*, 27(1):224–233, 2008. 11

[30] M. F. M Engel, L Khemtemourian, C.C Kleijer, H.J.D Meeldijk, J Jacobs, A.J Verkleij, B de Kruijff, J.A Killian, and J.W.M Hoppener. Membrane damage by human islet amyloid polypeptide through fibril growth at the membrane. *Proceedings of the National Academy of Sciences*, 105(16):6033–6038, 2008. 11

[31] J.L Khemtemourian, A.J Killian, W Hoppener, and M.F.M Engel. Recent insights in islet amyloid polypeptide-induced membrane disruption and its role in -cell death in type 2 diabetes mellitus. *Experimental Diabetes Research*, 2008:421287, 2008. 11, 18

[32] W Xue, A.L Hellewell, W.S Gosal, S.W Homans, E.W Hewitt, and S.E Radford. Fibril Fragmentation Enhances Amyloid Cytotoxicity. *Journal of Biological Chemistry*, 284(49):34272–34282, 2009. 11, 79

[33] L Milanesi, T Sheynis, W-F Xue, E V Orlova, A L Hellewell, R Jelinek, E W Hewitt, S E Radford, and H R Saibil. Direct three-dimensional visualization of membrane disruption by amyloid fibrils. *Proceedings of the National Academy of Sciences*, 109(50):20455–20460, 2012. 11, 124

[34] D.M Fowler, A.V Koulov, W.E Balch, and J.W Kelly. Functional amyloid- from bacteria to humans. *Trends Biochem Sci*, 32(5):217–224, 2007. 12

[35] M.M Barnhart and M.R Chapman. Curli biogenesis and function. *Annu. Rev. Microbiol*, 60, 2006. 12

[36] U Slotta, S Hess, K Spiess, T Stromer, L Serpell, and T. Scheibel. Spider silk and amyloid fibrils: a structural comparison. *Macromol. Biosci.*, 7, 2007. 12

[37] V.A Iconomidou, G Vriend, and S.J Hamodrakas. Amyloids protect the silkmoth oocyte and embryo. *FEBS Letters*, 479(3):141–145, 2000. 12

[38] J.E Podrabsky, J.F Carpenter, and S.C Hand. Survival of water stress in annual fish embryos: dehydration avoidance and egg envelope amyloid fibers. *Am J. Physiol - Regul. Integr Comp Physiol*, 280(1):R123–R131, 2001. 12

[39] S Keten, Z Xu, B Ihle, and M.J Buehler. Nanoconfinement controls stiffness, strength and mechanical toughness of b-sheet crystals in silk. *Nat Mater*, 9(4):359–367, 2010. 13

[40] D.M Fowler, A.V Koulov, C Alory-Jost, M.S Marks, William E Balch, and J.W Kelly. Functional amyloid formation within mammalian tissue. *PLoS Biol*, 4(1):e6, 11 2005. 12, 13, 104

[41] Gras S.L. Amyloid fibrils: From disease to design. new biomaterial applications for self-assembling cross-b fibrils. *Aust. J. Chem.*, 60:333–342, 2007. 13, 14, 83

[42] S Mankar, A Anoop, S Sen, and S Maji. Nanomaterials: amyloids reflect their brighter side. *Nano Reviews*, 2, 2011. 14, 83

[43] T Scheibel, R Parthasarathy, G Sawicki, X Lin, H Jaeger, and S.L Lindquist. Conducting nanowires built by controlled self-assembly of amyloid fibers and selective metal deposition. *Proceedings of the National Academy of Sciences*, 100(8):4527–4532, 2003. 14, 83

[44] S.K Maji, D Schubert, C Rivier, S Lee, J.E Rivier, and R Riek. Amyloid as a depot for the formulation of long-acting drugs. *PLoS Biol*, 6(2):e17, 2008. 14, 83

[45] M Ahn, S Kang, H.J Koo, J Lee, Y Lee, and S.R Paik. Nanoporous protein matrix made of amyloid fibrils of b2-microglobulin. *Biotechnology Progress*, 26(6):1759–1764, 2010. 14

[46] T.C Holmes, S de Lacalle, X Su, G Liu, A Rich, and S Zhang. Extensive neurite outgrowth and active synapse formation on self-assembling peptide scaffolds. *Proceedings of the National Academy of Sciences*, 97(12):6728–6733, 2000. 14, 83

[47] A Horii, X Wang, F Gelain, and S Zhang. Biological designer self-assembling peptide nanofiber scaffolds significantly enhance osteoblast proliferation, differentiation and 3-d migration. *PLoS ONE*, 2(2):e190, 2007. 14, 83

[48] G Bhak, S Lee, J.W Park, S Cho, and S.R Paik. Amyloid hydrogel derived from curly protein fibrils of a-synuclein. *Biomaterials*, 31(23):5986–5995, 2010. 14, 83

[49] T.P.J Knowles, T.W Oppenheim, A.K Buell, D.Y Chirgadze, and M.E Welland. Nanostructured films from hierarchical self-assembly of amyloidogenic proteins. *Nat Nano*, 5(3):204–207, 2010. 14

[50] T.P.J Knowles and M.J Buehler. Nanomechanics of functional and pathological amyloid materials. *Nat Nano*, 6(8):469–479, 2011. 14, 15

[51] J.F Smith, T.P.J Knowles, C.M Dobson, C.E MacPhee, and M.E Welland. Characterization of the nanoscale properties of individual amyloid fibrils. *Proceedings of the National Academy of Sciences*, 103(43):15806–15811, 2006. 16

[52] A.S Mostaert, R Crockett, G Kearn, I Cherny, E Gazit, L.C Serpell, and S.P Jarvis. Mechanically functional amyloid fibrils in the adhesive of a marine invertebrate as revealed by raman spectroscopy and atomic force microscopy. *Archives of Histology and Cytology*, 72(4+5):199–207, 2009. 16

[53] A Mostaert, M Higgins, T Fukuma, F Rindi, and S Jarvis. Nanoscale mechanical characterisation of amyloid fibrils discovered in a natural adhesive. *Journal of Biological Physics*, 32:393–401, 2006. 16

[54] S Guo and B.B Akhremitchev. Packing density and structural heterogeneity of insulin amyloid fibrils measured by afm nanoindentation. *Biomacromolecules*, 7(5):1630–1636, 2006. 16

[55] L.L del Mercato, G Maruccio, P.P Pompa, B Bochicchio, A.M. Tamburro, R Cingolani, and R Rinaldi. Amyloid-like fibrils in elastin-related polypeptides: Structural characterization and elastic properties. *Biomacromolecules*, 9(3):796–803, 2008. 16

[56] K Nitzan, L Adler-Abramovich, D Barlam, R.Z Shneck, E Gazit, and I Rousso. Self-assembled peptide nanotubes are uniquely rigid bioinspired supramolecular structures. *Nano Letters*, 5(7):1343–1346, 2005. 16

[57] J. F. Graveland-Bikker, I. A. T. Schaap, C. F. Schmidt, and C. G. de Kruif. Structural and mechanical study of a self-assembling protein nanotube. *Nano Letters*, 6(4):616–621, 2006. 16

[58] T P Knowles, A W Fitzpatrick, S Meehan, H R Mott, M Vendruscolo, C M Dobson, and M E Welland. Role of intermolecular forces in defining material properties of protein. *Science*, 318(5858):1900–1903, 2007. 16, 79

[59] F Meersman, R.Q Cabrera, P.F McMillan, and V Dmitriev. Structural and mechanical properties of ttr105-115 amyloid fibrils fromcompression experiments. *Biophys J*, 100(1):193–197, 2011. 16

[60] E.P Raman, T Takeda, and D.K Klimnov. Mechanical unbinding of ab peptides from amyloid fibrils. *Journal of Molecular Biology*, 373(3):785–800, 2007. 16

[61] R Paparcone, S Keten, and M.J Buehler. Atomistic simulation of nanomechanical properties of alzheimer's ab(1-40) amyloid fibrils under compressive and tensile loading. *Journal of Biomechanics*, 43(6):1196–1201, 2010. 17

[62] R Paparcone and M J Buehler. Failure of ab(1-40) amyloid fibrils under tensile loading. *Biomaterials*, 32(13):3367–3374, 2011. 17

[63] R Paparcone, M A. Pires, and M J. Buehler. Mutations alter the geometry and mechanical properties of alzheimers ab(140) amyloid fibrils. *Biochemistry*, 49(41):8967–8977, 2010. 17

[64] B Isralewitz, M Gao, and K Schulten. Steered molecular dynamics and mechanical functions of proteins. *Current Opinion in Structural Biology*, 11(2):224 – 230, 2001. 17, 29, 43

[65] M Sotomayor and K Schulten. Single-Molecule Experiments in Vitro and in Silico. *Science*, 316(5828):1144–1148, 2007. 17, 43

[66] Kapurniotu A. Amyloidogenicity and cytotoxicity of islet amyloid polypeptide. *Biopolymers*, 60(6):438–59, 2001. 17, 18

[67] P Westermark, U Engstrm, K H Johnson, G T Westermark, and C Betsholtz. Islet amyloid polypeptide: pinpointing amino acid residues linked to amyloid fibril formation. *Proceedings of the National Academy of Sciences*, 87(13):5036–5040, 1990. 18, 54, 59, 62, 64

[68] J Davis and W.E Van Nostrand. Enhanced pathologic properties of dutch-type mutant amyloid beta-protein. *Proceedings of the National Academy of Sciences*, 93(7):2996–3000, 1996. 18

[69] A.T Petkova, W Yau, and R Tycko. Experimental constraints on quaternary structure in alzheimer's b-amyloid fibrils. *Biochemistry*, 45(2):498–512, 2006. 18

[70] A.K Paravastu, R.D Leapman, W Yau, and R Tycko. Molecular structural basis for polymorphism in alzheimer's b-amyloid fibrils. *Proceedings of the National Academy of Sciences*, 105(47):18349–18354, 2008. 18, 19, 36, 116

[71] D Scheuner, C Eckman, M Jensen, X Song, M Citron, N Suzuki, T D Bird, J Hardy, M Hutton, W Kukull, E Larson, E Levy-Lahad, M Viitanen, E Peskind, P Poorkaj, G Schellenberg, R Tanzi, W Wasco, L Lannfelt, D Selkoe, and S. Younkin. Secreted amyloid beta-protein similar to that in the senile plaques of alzheimer's disease is increased in vivo by the presenilin 1 and 2 and app mutations linked to familial alzheimer's disease. *Nat Med*, 2(8):864–870, 1996. 20

[72] L Miravalle, T Tokuda, R Chiarle, G Giaccone, O Bugiani, F Tagliavini, B Frangione, and J Ghiso. Substitutions at codon 22 of alzheimer's ab peptide induce diverse conformational changes and apoptotic effects in human cerebral endothelial cells. *Journal of Biological Chemistry*, 275(35):27110–27116, 2000. 20

[73] J. A McCammon, B. R. Gelin, and M Karplus. Dynamics of folded proteins. *Nature*, 267:585 – 590, 1977. 23

[74] T Schlick, R Collepardo-Guevara, L.A Halvorsen, S Jung, and X Xiao. Biomolecular modeling and simulation: a field coming of age. *Quarterly Reviews of Biophysics*, 44(02):191–228, 2011. 23

[75] A. R Leach. *Molecular modelling principles and applications.* Pearson Education Limited, 2001. 23

[76] T Schlick. *Molecular modelling and simulation: An interdisciplinary guide.* Springer Science, 2001. 23

[77] J. A Mccammon and S. C Harvey. *Dynamics of proteins and nucleic acids.* Cambridge University Press, 1987. 23

[78] J.C Phillips, R Braun, W Wang, J Gumbart, E Tajkhorshid, E Villa, C Chipot, R D Skeel, L Kale, and K. Schulten. Scalable molecular dynamics with namd. *Journal of Computational Chemistry*, 26:1781–1802, 2005. 30, 47

[79] C Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78:2690 – 2693, 1997. 32, 33

[80] C Jarzynski. Rare events and the convergence of exponentially averaged work values. *Phys. Rev. E 73*, 2006. 33

[81] D.D. L Minh and J.A Mccammon. Springs and speeds in free energy reconstruction from irreversible single-molecule pulling experiments. *The journal of physical chemistry. B.*, 112(19), 2008. 33, 44

[82] S Park and K Schulten. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.*, 120:5946–5961, 2004. 33

[83] L Young, H Ndlovu, TW Knapman, SA Harris, S E Radford, and Ashcroft AE. Monitoring oligomer formation from self-aggregating amylin peptides using esi-ims-ms. *International journal for ion mobility spectrometry*, pages 1–11, 2013. 35

[84] J C Gordon, J B Myers, T Folta, V Shoja, L S Heath, and A Onufriev. H++: a server for estimating pkas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.*, 33(Web Server issue):W368–W371, 2005. 37

[85] H Li, A D Robertson, and J H Jensen. Very fast empirical prediction and rationalization of protein pka values. *Proteins*, 61:704–721, 2005. 37

[86] M Sikora, J.I Sulkowska, and M Cieplak. Mechanical strength of 17134 model proteins and cysteine slipknots. *PLoS Comput Biol*, 5:e1000547, 2009. 38

[87] M Sikora and M Cieplak. Mechanical stability of multidomain proteins and novel mechanical clamps. *Proteins*, 79:1786–1799, 2011. 38

[88] P Lukasz, S Mateusz, N Wieslaw, and C Marek. Molecular jamming: The cystine slipknot mechanical clamp in all-atom simulations. *J. Chem. Phys*, 134:085102, 2011. 38

[89] M Sikora, J.I Sulkowska, B.S Witkowski, and M Cieplak. Bsdb: the biomolecule stretching database. *Nucleic Acids Research*, 39(suppl 1):D443–D450, 2011. 39

[90] S Park, F Khalili-Araghi, E Tajkhorshid, and K Schulten. Free energy calculation from steered molecular dynamics simulations using jarzynski's equality. *The Journal of Chemical Physics*, 119(6):3559–3566, 2003. 43, 44

[91] D.A. Case, T.A. Darden, T.E. Cheatham III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, K.M. Merz, D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, and D.H. Mathews. *AMBER 9, University of California, San Francisco.*, 2006. 45

[92] W Kabsch and C Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22(12), 1983 Dec. 48

[93] W. Humphrey, A. Dalke, and K. Schulten. Vmd - visual molecular dynamics. *J. Molec. Graphics*, 14, 1996. 48, 49, 54

[94] M.C Lawrence and P.M Colman. Shape complementarity at protein/protein interfaces. *Journal of Molecular Biology*, 234(4):946 – 950, 1993. 49

[95] N.A Baker, D Sept, S Joseph, M.J Holst, and J.A McCammon. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America*, 98(18):10037–10041, 2001. 49

[96] H Ndlovu, A.E Ashcroft, S.E Radford, and S.A Harris. Effect of Sequence Variation on the Mechanical Response of Amyloid Fibrils Probed by Steered Molecular Dynamics Simulation. *Biophysical journal*, 102:587–596, 2012. 52

[97] F Chiti, M Stefani, N Taddei, G Ramponi, and Dobson C.M. Rationalisation of mutational effects on protein aggregation rates. *Nature*, 424:805–808, 2003. 52

[98] S Sakagashira, T Sanke, T Hanabusa, H Shimomura, S Ohagi, K Y Kumagaye, K Nakajima, and K Nanjo. Missense mutation of amylin gene (S20G) in Japanese NIDDM patients. *Diabetes*, 45(9):1279–1281, 1996. 53, 79

[99] S Sakagashira, H.J Hiddinga, K Tateishi, T Sanke, T Hanabusa, K Nanjo, and N.L Eberhardt. S20G Mutant Amylin Exhibits Increased in Vitro Amyloidogenicity and Increased Intracellular Cytotoxicity Compared to Wild-Type Amylin. *Am J Pathol*, 157(6):2101–2109, 2000. 53, 79

[100] P Marek, A Abedini, B Song, M Kanungo, M.E Johnson, R Gupta, W Zaman, S.S Wong, and D.P Raleigh. Aromatic interactions are not required for amyloid fibril formation by islet amyloid polypeptide but do influence the rate of fibril formation and fibril morphology. *Biochemistry*, 46(11):3255–3261, 2007. 54

[101] S.M. Tracz, A Abedini, M Driscoll, and D.P Raleigh. Role of aromatic interactions in amyloid formation by peptides derived from human amylin. *Biochemistry*, 43(50):15901–15908, 2004. 54, 59

[102] D.F Moriarty and D.P Raleigh. Effects of sequential proline substitutions on amyloid formation by human amylin20-29. *Biochemistry*, 38(6):1811–1818, 1999. 54

[103] H Dietz, F Berkemeier, M Bertz, and M Rief. Anisotropic deformation response of single protein molecules. *Proceedings of the National Academy of Sciences*, 103(34):12724–12728, 2006. 71

[104] D J Brockwell, E Paci, R C Zinober, P D Olmsted, D A Smith, R N Perham, and S E Radford. Pulling geometry defines the mechanical resistance of a beta-sheet protein. *Nat Struct Mol Biol*, 10(9):731–737, 2003. 71

[105] M Carrion-Vazquez, H Li, H Lu, P E Marszalek, A F Oberhauser, and J M Fernandez. The mechanical stability of ubiquitin is linkage dependent. *Nat Struct Mol Biol*, 10(9):738–743, 2003. 71

[106] A.M Stoneham. *Theory of defects in solids: electronic structure of defects in insulators and semiconductors*. Monograph on the Physics and Chemistry of Materials. Oxford University Press, 1975. 78

[107] M Tanaka, S R Collins, B H Toyama, and J S Weissman. The physical basis of how prion conformations determine strain phenotypes. *Nature*, 442(7102):585–589, 2006. 79

[108] D Zanuy, F Rodriguez-Ropero, N Haspel, J Zheng, R Nussinov, and C Aleman. Stability of tubular structures based on beta-helical proteins: self-assembled versus polymerized nanoconstructs and wild-type versus mutated sequences. *Biomacromolecules*, 8(10):3135–3146, 2007. 80

[109] B H Toyama, M J Kelly, and J S Weissman. The structural basis of yeast prion strain variants. *Nature*, 449:233–237, 2007. 82

[110] J. L. Jimenez, E.J Nettleton, and H.R. Saibil. The protofilament structure of insulin amyloid fibrils. *PROC. NATL. ACAD. SCI. USA*, 99:9196–9201, 2002. 82

[111] R. Krishnan and S.L. Lindquist. Structural insights into a yeast prion illuminate nucleation and strain diversity. *Nature*, 435:765–772, 2005. 82

[112] L.J. Domigan, J.P. Healy, S.J. Meade, R.J. Blaikie, and J.A. Gerrard. Controlling the dimensions of amyloid fibrils: Towards homogeneous components for bionanotechnology. *Biopolymers*, 97:123–133, 2011. 84

[113] G. Yoon, J. Kwak, J.I. Kim, S. Na, and K. Eom. Mechanical characterization of amyloid fibrils using coarse-grained normal mode analysis. *Adv. Funct. Mater*, 21:3454–3463, 2011. 94

[114] T P. J. Knowles, J F. Smith, G.L Devlin, C M. Dobson, and M E. Welland. Analysis of structural order in amyloid fibrils. *Nanotechnology*, 18:044031, 2007. 101

[115] D Foguel, M.C. Suarez, A.D. Ferrao-Gonzales, T.C. Porto, L Palmieri, C.M. Einsiedler, L.R. Andrade, H.A. Lashuel, P.T. Lansbury, J.W. Kelly, and J.L. Silva. Dissociation of amyloid fibrils of a-synuclein and transthyretin by pressure reveals their reversible nature and the formation of water-excluded cavities. *Proc. Natl. Acad. Sci. USA*, 100:9831–9836, 2003. 101

[116] B O'Nuallain, A D Williams, P Westermark, and R Wetzel. Seeding specificity in amyloid growth induced by heterologous fibrils. *Journal of Biological Chemistry*, 279(17):17490–17499, 2004. 104

[117] W Surmacz-Chwedoruk, H Nieznanska, S Wojcik, and W Dzwolak. Cross seeding of fibrils from two types of insulin induces new amyloid strains. *Biochemistry*, 51:9460–9469, 2012. 104

[118] A Peim, P Hortschansky, T Christopeit, V Schroeckh, W Richter, and M Fandrich. Muatagenic exploration of the cross seeding and fibrillation propensity of alzheimers b-amyloid peptide variants. *Protein Science*, 15:1801–1805, 2006. 105

[119] J Zhao, X Yu, G Liang, and J Zheng. Heterogeneous triangular structures of human islet amyloid polypeptide (amylin) with internal hydrophobic cavity and external wrapping morphology reveal the polymorphic nature of amyloid fibrils. *Biomacromolecules*, 12:1781–1794, 2011. 105

[120] A. D. Williams, E. Portelius, I. Kheterpal, J. Guo, K. D. Cook, Y. Xu, and R. Wetzel. Mapping ab amyloid fibril secondary structure using scanning proline mutagenesis. *J. Mol. Biol*, 335:833–842, 2004. 106

[121] C Nilsberth, C B Westlind-Danielsson A, Eckman, M M Condron, K Axelman, C Forsell, C Stenh, J Luthman, D B Teplow, S G Younkin, J Naslund, and L Lannfelt. The arctic mutation (e693g causes alzheimer's disease by enhanced ab proto-fibril formation. *Nature neuroscience*, 4:887–893, 2001. 106

[122] A T Petkova, Y Ishii, J J Balbach, O N Antzutkin, R D Leapman, F Delaglio, and R Tycko. A structural model for alzheimer's b-amyloid fibrils based on experimental constraints from ssnmr. *PNAS*, 99:16742–16747, 2002. 116

[123] N V Buchete and G Hummer. Structure and dynamics of parallel b-sheets, hydrophobic core and loops in alzheimer's ab fibrils. *Biophys J*, 97:3032–3039, 2007. 116