Effect of *De Novo* Peptide Properties on Self-Assembling Large Amyloid Fibers

Caitlin Weigand Rippner

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science

In

Biological Systems Engineering

Justin R. Barone

Scott H. Renneckar

Chenming (Mike) Zhang

April 23, 2013

Blacksburg, VA

Keywords: amyloid, peptide, self-assembly, glutamine repeats, FTIR, AFM

Effect of *De Novo* Peptide Properties on Self-Assembling Large Amyloid Fibers

Caitlin Weigand Rippner

ABSTRACT

Amyloid aggregation involves the spontaneous formation of fibers from misfolded proteins. This process requires low energy input, results in robust fibers, and is thus of interest from a materials manufacturing perspective. The effect of glutamine content and hydrophobicity of template peptides on amyloid aggregation of a template-peptide system involving myoglobin was studied at near-physiological conditions by Fourier transform infrared spectroscopy, atomic force microscopy, field emission scanning electron microscopy, and nanoindentation. Hydrophobic interactions were found to be important for controlled hierarchical fiber growth via a cooperative mechanism, with the largest effect in myoglobin mixtures. Hydrophobic packing increased for most systems as aggregation progressed. The largest changes in structure occurred upon drying. When myoglobin was present with the highest glutamine-containing template (P7), the high glutamine peptide was not effective as a template, since it appeared to prefer self-catalysis. A low level of glutamine in some unordered templates was insufficient for amyloid development. However, templating was more important in glutamine-free templates mixed with myoglobin, which formed fibers with a surprisingly high elastic modulus. This may have been due to template patterning. Nanoindentation results confirmed that glutamine blocks were not necessary for strong intermolecular interactions and cooperative fibril formation.

ACKNOWLEDGEMENTS

I am grateful to my family, my friends, and my advisor for supporting me. Thank you to the engineers who inspire me to keep coming back to this work.

A number of people provided invaluable assistance to me during my Master's degree. I thank my advisor Justin Barone for being a gracious host, for conducting lively weekly group meetings, and for edifying talks about family and career. Thank you to Scott Renneckar and Mike Zhang for serving on my committee. I'm grateful to my fellow graduate students who trained me, assisted in my experiments, and provided many interesting discussions along the way. In particular, Devin Ridgley, Betsy Claunch, and Parker Lee have been instrumental to my success. Thank you to Liz Collins, Sara Senyondo, and Lele Kimball for discussions on the human aspect of graduate school. Thanks also to Bob Wright for traveling to Blacksburg during his free time to train me on a previous project.

Lastly, thank you to my husband Devin for going on this adventure with me.

CONTENTS

All images belong to the author unless otherwise noted.

FIGURES

TABLES

EQUATIONS

CHAPTER 1: Introduction and Literature Review

Nature encodes in biological molecules a host of mechanisms that strengthen, protect, and power life on the molecular scale. Natural protein fibers such as keratin, elastin, collagen, fibroin, actin, tubulin, and viral spikes show a hierarchical organization that permits matching of the survival needs of the organism to fiber properties on many length scales. These characteristics inspire the search for engineered biomaterials that can provide useful variations on known natural fibers. Self-assembled amyloid proteins are one example of a protein-based material with interesting properties that can be controlled by amino acid sequence. An inspiring example of the potential for controllable materials comes from spider silk fibers: protein composition and properties are altered by genetic splicing in response to the environment.[1] The potential for encoding specific material properties into biological structures at the level of genetic expression is just beginning to be tapped.

1.1 Project Objectives

1.1.1 Objective 1

- Synthesize template peptides of varying hydrophobicity, α helix content, and glutamine content and mix them with myoglobin to define the important features of the template peptide in spontaneous large fiber formation.

1.1.2 Objective 2

- Directly measure the interaction between template and adder proteins.

1.1.3 Objective 3

- Measure the mechanical properties of spontaneously formed large fibers.

1.2 Literature Review

1.2.1 Protein Folding and Aggregation

1.1.1.1 Amyloid Proteins

In general, protein aggregation is an important phenomena that impacts processing of medical, food and material products.[2] Amyloid aggregation is a type of protein aggregation that causes a number of pathological medical conditions, known together as protein misfolding diseases.[3] Protein misfolding diseases all share the common feature of dysfunctional, aggregated proteins that can be harmful or lethal over time. A few examples are Huntington's disease, Alzheimer's disease, type II diabetes, and bovine spongiform encephalopathy.[3] Such amyloid fiber aggregates form when a protein unfolds and hydrogen bonds to another unfolded protein segment to form a beta sheet. A few proteins can form oligomeric or globular beta sheet aggregates, which are the suspected pathogens in protein misfolding diseases. In other cases, beta sheets can continue to stack laterally and vertically to produce anisotropic fibrils. In these fibrils, the protein chain axis is perpendicular to the fibril axis, forming what is known as the "cross-β" structure.[4–6] Amyloids are durable and resistant to macrophage and enzymatic degradation.

Functional protein fibers based on the beta sheet structure are also important to the survival of a number of organisms. Some of these functional fibers are formed in the cross-β configuration, such as the adhesive secreted by barnacles for attachment to marine surfaces,[7] bacterial curli fimbriae structures used for attachment to biofilms and to inert surfaces,[8] glowworm threads used as nests and for catching prey, lacewing fly egg stalk silks used to

suspend and protect eggs from cannibalism,[9,10] a structural substrate for melanin deposition in mammals,[11] fungal hydrophobins,[12] and more.[13–15] Typical functional fibers in nature are formed with beta sheets aligned parallel to the fiber axis, such as high strength spider dragline silks[16] and fish proteins that gel upon freezing and thawing.[17] This is from the applied deformation used to form the fiber, whereas amyloids form spontaneously. The orientation of β-sheets in fibrous proteins may also be transformed by *ex situ* processing.[18,19]

1.1.1.2 Protein Folding and Misfolding

In disease states, protein aggregation occurs when functional proteins unfold and refold into a biologically inactive 3-dimensional structure. However, the primary purpose of functional amyloids seems to be to form the stable fiber that results, so they are not misfolded since the biological preference for fibers may dominate over other forms. Many wild type protein structures are only marginally stable in their natural state, and usually require a specific set of temperature, pH, and buffer conditions. Amyloid folding seems to go against many of these rules, since it occurs over a range of temperatures, pH, and salt concentrations, and the resulting fiber can be extremely stable in its fully assembled form. In the past, it was generally observed that proteins are optimized to have maximal biological activity in their specific environment. For example, stomach enzymes work best at low pH.

When proteins fold, stability derives from the balance between enthalpy and entropy. The net effect is the free energy of folding. Enthalpy describes the energy of favorable noncovalent interactions within the polypeptide chain, and is maximized with tighter packing and higher interaction frequency.[20] Entropy describes the opposing energy required to create order when folding occurs. From the second law of thermodynamics, the natural tendency is for disorder to

increase, which correlates with a higher number of possible protein conformations. However, when one particular protein conformation dominates, order is increased because the flexibility of the protein is limited. For the most part, stable proteins have enthalpic gains greater than their entropic losses, but the magnitude of the difference is often small, on the order of 1-3 hydrogen bonds.[20,21] Some folded states with sub-optimal free energy may be observed most frequently due to high kinetic accessibility.[22] The amyloid state is highly ordered, and thus requires either a large array of very favorable noncovalent enthalpic interactions or favorable kinetics. Since amyloid formation may now be considered a general part of the aging process, fast kinetics are not expected.[23] Therefore, the methods employed were intended to observe the thermodynamic aspects of aggregation: does it aggregate, or not?

1.3 Experimental Design

1.3.1 Peptide Design

The self-assembly of large amyloid fibers has been observed. The fibers have morphological features and physical properties similar to natural β-sheet containing fibers like silk. Such fibers had not been observed previously. Experimental evidence supports a multi-protein process where at least two proteins with specific but different properties work cooperatively to self-assemble the large amyloid fiber. In the hypothesized process, a hydrophobic "template" peptide forms a folded cross-β structure which enables hiding of hydrophobic amino acid side groups in the spaces between the β-sheets. The template peptide sequence has hydrophobic amino acids next to one another and thus has some hydrophobic groups exposed to the water on the outer faces of the folded β-sheets. The "adder" peptide or protein is α-helical and more hydrophilic than the template peptide. The more hydrophilic adder peptide or protein is stable in aqueous

solution and does not individually undergo conformation change or aggregate at the experimental conditions as measured with FT-IR spectroscopy.[24,25] However, in the presence of the template, the hydrophobic groups on the α-helices of the adder prefer the exposed hydrophobic groups of the template, undergo an α to β transition, and "add" into the template as measured with Fourier transform-infrared (FT-IR) spectroscopy.[25,26] Through AFM and SEM imaging, the progression of large amyloid fiber aggregation can be defined in terms of four morphological stages. The initial work was performed with a protein hydrolysate as the template while adder proteins were varied. In this work, pure engineered peptides were used as templates while keeping the adder protein constant. The goal of the present work was to quantify the effect of varying the template peptide on large amyloid fiber formation.

The synthetic peptides studied have a sequence of 20-22 amino acids. The sequences were varied to simulate key segments for aggregation, focusing on hydrophobicity and glutamine (Q) content. For all experiments, equine myoglobin (My, 154AA, Uniprot P02144) was used for "adding" purposes. In wild-type myoglobin at biological conditions, fibrous aggregation is not known to occur, so any observed aggregation is expected to result from cooperation with a template.[27–29] Myoglobin is a well-studied protein with a 3D structure consisting in 8 α helices surrounding a very hydrophobic core with a central iron-containing heme group.[30]

Sequences of the experimental peptides are shown in Table 1.1. Gd20KK was designed from the 3-22 tryptic hydrolysate of gliadin, a protein from wheat gluten and the suspected template in the crude gliadin hydrolysate used as the template in previous work. CB4 is the N-terminal sequence of a *Megabalanus rosa* peptide that is characterized by insolubility and abundant predicted β sheet structure.[31] The P4, P7, and P4 analog (P4-AN) templates were

designed to represent extremes of secondary structure, hydrophobicity, and Q content. P4 was expected to be primarily β sheet with high hydrophobicity, while P7 was predicted to be primarily α helical with lower hydrophobicity. P4-AN has equivalent hydrophobicity and predicted β sheet structure to P4, but glutamine residues were replaced by polar residues, including some with charge. The bioinformatics indices PSIPRED and ExPASy were used to predict peptide properties in amino acid sequence design.[32–34]

Table 1.1 Template peptide properties.

Identity	Sequence	Q (%)	GRAVY	%α	%β	pI	Mol. Wt (g/mol)
Gd20KK	MKTFLILALLAIVATTATTAKK	0	1.820	90	0	10.3	2320
CB4	SLGGVVAYLQLANIQQAVFISRI	13	1.022	0	83	8.5	2461
P4	LVLVQQQQLVLVQQQQLVLV	40	1.0	0	90	5.52	2317
P4-AN	LVLVKEHELVLVNHKHLVLV	0	1.005	0	90	7.03	2332
P7	QQQQQQLVLVLVQQQQQQQ	70	-1.25	90	0	5.52	2317

1.3.2 Basic Research Approach

In this study, we mixed the peptides and myoglobin in water, adjusted to pH 8, incubated at 37°C, and monitored the liquid for aggregation behavior over a period of 20 to 35 days. This experimental timeline lines up very well with the time scales observed in the past for complete sedimentation of polyglutamine repeats.[35,36] To characterize the template-adder interaction, time-resolved FTIR was used to measure the alpha helix to beta sheet transition and hydrophobic packing. Secondary structure was calculated using an accepted Amide I peak deconvolution technique. Spectroscopic studies of hydrated protein structure paired well with morphological studies using microscopy techniques. Atomic force microscopy (AFM) imaging was done at discrete points during the incubation period to provide a developmental snapshot of the hydrated

aggregate. The morphology of the fully assembled final product was observed using field emission scanning electron microscopy (FESEM). Dried fiber chemistry and elastic strength were studied using solid FTIR and nanoindentation.

1.4 Factors Affecting Aggregation

1.4.1 Precedent for Synthetic Peptide Studies

In other studies of synthetic peptides, many relied on high temperature, shear, charge effects, or denaturing pH for fibrillation to occur.[37–40] Those techniques were avoided in order to focus on biologically relevant conditions to specifically note the effects of the 2 proteins only. Most other studies of synthetic peptides also limited the peptide length to 4-10 amino acids, which coincides with the typical length of single β sheets that stack in amyloid aggregates.[37,41–43] The length of the peptides used here requires that some folding interactions must occur due to a thermodynamic need to increase stability in water.

1.4.2 Peptide Sequence Determines Structure and Function

This research, like much of the work preceding it, aimed to further explain the link between amino acid sequence and protein behavior.[21] The side chain chemistry and arrangement of the amino acids themselves were expected to have a great contribution to the final amyloid structures observed. Interactions that stabilized the growing fibrils are described in the following chapters.

The important structure-determining characteristics of the template peptides were glutamine repeats (P4 and P7), hydrophobicity (all peptides), the presence of charged amino acids (Gd20KK, CB4, P4-AN), and amphiphilicity (all peptides). Minor contributions may have been made by aromatic cation-π and π-π interactions (Gd20KK, CB4, P4 analog).[44] The

hydrophilic regions of the de novo peptides consist in either a block of glutamine residues (P4 and P7) or randomly selected polar amino acids that resulted in identical secondary structure predictions (P4 analog). See Table 1.1 for the exact template sequences.

1.4.2.1 Hypothesized Role of Glutamine Repeats

Glutamine repeats are an important factor in prion disease development, such as in Huntington's disease, and are also highly conserved in some functional amyloids.[45,46] A higher number of glutamine repeats correlates to faster progression of disease symptoms in affected individuals. In studies of *Pseudomonas* genomics, codons for glutamine are highly conserved in species that display amyloid curli proteins which enable life-sustaining functions wherein microbes can flocculate and adhere to surfaces. Thus, the presence of Q in the protein amino acid sequence is associated with amyloid fiber formation from a variety of proteins. The glutamine composition of the templates tested in this research (P4, P7) is far below known disease thresholds, so the cooperative amyloids studied were therefore of primary interest for their materials properties.[45] However, *in vitro* studies show that Q repeats of less than 35 residues can form amyloid fibrils.[47] Poly-glutamine repeats have formed the basis for many engineered peptide systems and they continue to shed light on the impact of side chain chemistry on amyloid protein aggregation processes and mechanical properties.[42,45,48]

From X-ray diffraction studies, the polyglutamine structure shows tightly packed side chains locking together with a high frequency of hydrogen bonds.[49] Because of the structural regularity and chemical complementarity of glutamine to itself and to the peptide backbone, these repeats cause a large stabilizing influence on β sheet structures in water. Each propionamide moiety of the side chain has two oppositely-charged dipoles, which form two

possible hydrogen bonds per pair of interacting glutamines. Glutamine can also form a regular, repeating pattern of two hydrogen bonds per Q residue with the peptide backbone at the C=O and N-H groups.[50] Although this mode of interaction is thought to be critical to the stacking of amyloid tapes into larger structures, the presence of repeating Q side chains enables monomeric addition to either the backbone or to the side chain face of a growing fiber, so it is especially flexible. This flexibility may explain why Q-based aggregation is thermodynamically favored. However, Q blocks exert their effect on the larger system, so it follows that the sequence context for Q blocks has a large impact on stability, solubility, and aggregation.[51] Other amino acids such as asparagine (N), serine (S), and threonine (T) in Gd20KK, CB4, and P4-AN can also form enthalpically favorable hydrogen bonds, but N has similar structure to Q and may therefore interact in an analogous way to Q in CB4 and P4-AN templates.

1.4.2.2 Hypothesized Role of Hydrophobicity

Hydrophobicity is expected to be a major driving factor for small peptide-templated aggregation since it is widely influential in protein folding and unfolding. In the templates tested, Leucine (L) and valine (V) are the primary contributors to calculated hydrophobic tendency (positive GRAVY), but methylene skeletons of amino acids with longer side chains enable additional hydrophobic interactions.[52] L and V are expected to have a greater propensity to form β sheets compared to alanine (A) and glycine (G),[53] perhaps due to size and shape factors.[43] However, charged, polar lysine (K) has also been found to lend thermodynamic stabilization in inter-sheet interactions with L via hydrophobic interactions of the methylene skeleton. Lysine can also interact with aromatic residues with its ε-methylene group.[44]

Hydrophobic interactions are also prominent in therapeutic efforts to prevent and treat

protein misfolding diseases. In the peptides associated with Alzheimer's disease, Aβ40 and Aβ42, hydrophobic interactions are key to toxicity and its antithesis, therapeutic treatment. Hydrophobic interactions are also critical to the function of a flexible, ligand-binding molecular tweezer designed to bind specifically to the butylene moiety of the side chain of K. In the molecular tweezer, hydrophobic interactions combine with amino group electrostatic interactions to bind the lysine target between two flexible arms.[54] Such molecular tweezers illustrate the power of utilizing side chain chemistry in designing systems to control protein behavior. Also in Aβ research, antibodies with grafted hydrophobic Aβ peptides on the complementarity-determining regions interact in a sequence-specific way with Aβ and inhibit amyloid formation and toxicity.[55] Hydrophobicity is important to self-assembly of fibers in the presence or absence of a disease context. Increased hydrophobicity leads to a transfer of fibrillogenesis after 20 days in a transplant of 6 amino acids from a fibril-generating protein to a non-fibrillating protein.[56]

Hydrophobicity has an impact on the overall protein folding process by biological surface interactions and protein-ligand binding as well as functional protein binding energy.[57] For example, in protein synthesis in the endoplasmic reticulum, quality control of nascent proteins by chaperone molecules is done via interactions localized to hydrophobic peptide surfaces.[58] Lipid membranes provide an important context for determining the role of hydrophobic cellular surfaces in protein folding and aggregation.[59]

1.4.2.3 Hypothesized Role of Ionic Interactions

The aqueous solvent had minimal ionic strength, so ionic interactions between amino acid side chain moieties were expected to play a more dominant role in inter-protein associations involved in fiber assembly. In mixtures with Gd20KK and Q-free P4-AN, K may be an important factor in

fibrillation due to its charge and ability to interact with glutamic acid (E) and other negatively charged and electron-rich side chains via Coulombic interactions. In studies of K to A point substitutions in Aβ, loss of ionic bonding ability resulted in a change of morphology to shorter, fatter fibrils with lower toxicity.[60] Lysine also participates in critical salt bridges with aspartic acid (D) in the quenched hydrogen/deuterium exchange NMR-derived 3D model of the β sheet structure of residues 17-42 of Aβ(1-42).[61] Both K and histidine (H, in the sequence of P4-AN) have a charged ammonium ion that may either form salt bridges or enable stabilizing cation-π interactions with aromatic residues in My, such as tryptophan (Y) or phenylalanine (F).[44]

1.4.2.4 Hypothesized Role of Amphiphilicity

Alternating hydrophobic and hydrophilic primary structure is a feature of some peptides that gives them a particular amphiphilicity that may enable interactions with fundamentally different protein domains. Amphiphilicity may also lend proteins enhanced thermal stability.[62] In one example from Alzheimer's disease, fibrillar Aβ derives from an amphiphilic fragment of the transmembrane amyloid precursor protein (APP) that has been abnormally cleaved by two different enzymes.[63] Another similar example is the mammalian structural protein Pmel17, which provides a substrate for melanin deposition.[64] This glycoprotein becomes fibrillogenic when enzymatically cleaved, leaving a fragment containing both a transmembrane sequence and a cytoplasmic sequence, similar to APP.[63,65] Such amphiphilic fragments are both potent fibrillators, but they also share an innate ability to insert into amphiphilic lamellae. Other functional amyloids display a similar pattern, with hydrophobic residues playing an important role in aggregation. The barnacle protein Mrcp-100k employs alternating hydrophobic and hydrophilic amino acids in a strong, insoluble protein-based marine surface adhesive.[7] The self-

aggregating yeast protein zuotin, known to bind to Z-DNA, contains a repetitive sequence in which hydrophobic residues alternate with oppositely charged hydrophilic residues.[66] Peptides from the monomeric elastin precursor tropoelastin contain alternating hydrophobic and crosslinking domains that display distinct roles, wherein hydrophobicity promotes self-aggregation and protein extensibility.[41,67,68]

1.4.3 Hypothesized Self-Assembly Process

The exact molecular mechanism for self-assembly of amyloid fibers is not yet established.[69] Researchers have modeled the kinetics in different ways, reviewed elsewhere.[70] Most mechanisms are commonly dependent on amyloid nucleation, similar to crystallization and other regulated biological processes.[8,69,71,72] In a general aggregation process, a small number of single proteins in an abnormal, energetically disfavored state form aggregated oligomers at a very slow rate until assembly has proceeded so much that solubility is no longer favorable.[71] The insoluble fibers that result are responsible for both strong mechanical properties and neurological dysfunction found in disease states. The cooperative template-adder assembly process proposed in the present research may be able to explain assembly of small catalytic peptide fragments into a large fiber.[73]

There are many examples of cooperative protein-peptide interactions that behave in a similar way to the present system. In α-lytic protease, the stable unfolded protein is induced to fold to the native form only in the presence of a shorter prosegment. The mechanism in that case is a kinetic mechanism in which the prosegment significantly accelerates the rate-limiting step.[74] In curli fimbriae assembly in gram-negative bacteria, cooperative assembly was demonstrated. Hierarchical curli assembly proceeds when a small nucleator peptide, CsgB, acts on a larger

protein subunit, CsgA.[8] *In vivo*, this effect is mediated by 4 to 5 chaperone proteins which coordinate the amyloid attachment between the membrane surface and the environment.[15] Interestingly, the nucleator protein CsgB is not specific to bacterial species. Nucleators expressed in one species cause curli assembly and attachment in their species homologs, which indicates that only certain structural elements are necessary for aggregation. Complete DNA and protein sequence homology may not be required. There also exists a disease-related protein pair that has hydrophobic patterning and size characteristics analogous to the present system: Huntingtin protein Htt^{NT} and Htt-associated protein 1, HAP1.[75]

In amyloid folding, a hierarchy of self-assembly is followed by some fibers. Hierarchical construction is present in many naturally-occurring structures and it controls properties from the smallest to the largest scales. Humans depend on a hierarchical structure for hard tissues, such as bone; elastic tissues, such as ligaments; and soft tissues, such as skin. In protein solutions, if conditions are right, peptides can self-associate to form structures starting with individual molecules with dimensions measured in angstroms, up to nanometers, that associate and build to micrometer scale, and even to millimeter scale on drying.[76–78] After proteins are expressed and folded, hierarchical self-assembly provides a stepwise path for construction of rigid, organized, degradation-resistant fibers.[79] In the present system, the cooperative mechanism may not be specific to the "adder" protein sequence. Therefore, it may be possible to generate amyloids with a range of material properties over varying length scales by pairing an appropriate designed template with families of similar proteins.

References

1. Craig, C. L. *et al.* Evidence for Diet Effects on the Composition of Silk Proteins Produced by Spiders. *Mol Biol Evol* **17**, 1904-1913 (2000).
2. Law, A. J. R. & Leaver, J. Effect of pH on the Thermal Denaturation of Whey Proteins in Milk. *J. Agric. Food Chem.* **48**, 672-679 (2000).
3. Chiti, F. & Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease - Annual Review of Biochemistry, 75(1):333. *Annual Review of Biochemistry* **75**, 333-366 (2006).
4. Eanes, E. D. & Glenner, G. G. X-Ray Diffraction Studies on Amyloid Filaments. *J Histochem Cytochem* **16**, 673-677 (1968).
5. Geddes, A. J., Parker, K. D., Atkins, E. D. T. & Beighton, E. 'Cross-β' conformation in proteins. *Journal of Molecular Biology* **32**, 343-358 (1968).
6. Sunde, M. *et al.* Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of Molecular Biology* **273**, 729-739 (1997).
7. Kamino, K. *et al.* Barnacle Cement Proteins: Importance of Disulfide Bonds in their Insolubility. *J. Biol. Chem.* **275**, 27360-27365 (2000).
8. Chapman, M. R. *et al.* Role of Escherichia coli Curli Operons in Directing Amyloid Fiber Formation. *Science* **295**, 851-855 (2002).
9. Parker, K. D. & Rudall, K. M. The Silk of the Egg-Stalk of the Green Lace-Wing Fly: Structure of the Silk of Chrysopa Egg-stalks. *Nature* **179**, 905-906 (1957).
10. Weisman, S. *et al.* Fifty years later: The sequence, structure and function of lacewing cross-beta silk. *Journal of Structural Biology* **168**, 467-475 (2009).
11. Fowler, D. M. *et al.* Functional Amyloid Formation within Mammalian Tissue. *PLoS Biol* **4**,

(2006).

12. Wösten, H. A. & de Vocht, M. L. Hydrophobins, the fungal coat unravelled. *Biochim. Biophys. Acta* **1469**, 79-86 (2000).
13. Iconomidou, V. A. *et al.* Amyloid fibril formation propensity is inherent into the hexapeptide tandemly repeating sequence of the central domain of silkmoth chorion proteins of the A-family. *J. Struct. Biol.* **156**, 480-488 (2006).
14. Shewmaker, F., McGlinchey, R. P. & Wickner, R. B. Structural Insights into Functional and Pathological Amyloid. *J Biol Chem* **286**, 16533-16540 (2011).
15. Otzen, D. Functional amyloid. *Prion* **4**, 256-264 (2010).
16. Hayashi, C. Y., Shipley, N. H. & Lewis, R. V. Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. *International Journal of Biological Macromolecules* **24**, 271-275 (1999).
17. Graether, S. P., Slupsky, C. M. & Sykes, B. D. Freezing of a Fish Antifreeze Protein Results in Amyloid Fibril Formation. *Biophys J* **84**, 552-557 (2003).
18. Rudall, K. M. Silk and other cocoon proteins. *Comparative Biochemistry: A Comprehensive Treatise. Constituents of Life Part B* **IV**, (1962).
19. Bousset, L., Briki, F., Doucet, J. & Melki, R. The native-like conformation of Ure2p in fibrils assembled under physiologically relevant conditions switches to an amyloid-like conformation upon heat-treatment of the fibrils. *Journal of Structural Biology* **141**, 132-142 (2003).
20. Branden, C. & Tooze, J. Folding and Flexibility. *Introduction to Protein Structure* (1999).
21. Anfinsen, C. B. Principles that Govern the Folding of Protein Chains. *Science* **181**, 223-230 (1973).

22. Creighton, T. E. Up the Kinetic Pathway. *Nature* **356**, 194 (1992).
23. David, D. C. *et al.* Widespread Protein Aggregation as an Inherent Part of Aging in C. elegans. *PLoS Biol* **8**, (2010).
24. Athamneh, A. I. & Barone, J. R. Enzyme-mediated self-assembly of highly ordered structures from disordered proteins. *Smart Materials and Structures* **18**, 104024 (2009).
25. Ridgley, D. M., Ebanks, K. C. & Barone, J. R. Peptide Mixtures Can Self-Assemble into Large Amyloid Fibers of Varying Size and Morphology. *Biomacromolecules* **12**, 3770-3779 (2011).
26. Ridgley, D. M., Claunch, E. C. & Barone, J. R. The effect of processing on large, self-assembled amyloid fibers. *Soft Matter* (2012).doi:10.1039/c2sm26496j
27. Fukuma, T., Mostaert, A. S. & Jarvis, S. P. Explanation for the mechanical strength of amyloid fibrils. *Tribol Lett* **22**, 233-237 (2006).
28. Stefani, M. & Dobson, C. M. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J Mol Med* **81**, 678-699 (2003).
29. Fändrich, M., Fletcher, M. A. & Dobson, C. M. Amyloid fibrils from muscle myoglobin. *Nature* **410**, 165-166 (2001).
30. Kendrew, J. C. Myoglobin and the Structure of Proteins. *Science* **139**, 1259-1266 (1963).
31. Kamino, K., Odo, S. & Maruyama, T. Cement Proteins of the Acorn Barnacle, Megabalanus rosa. *Biological Bulletin* **190**, 403-409 (1996).
32. McGuffin, L. J., Bryson, K. & Jones, D. T. The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404-405 (2000).
33. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157**, 105-132 (1982).

34. Gasteiger, E., Hoogland, C., Gattiker, A. & Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook* 571-607 (2005).

35. Walters, R. H. & Murphy, R. M. Aggregation Kinetics of Interrupted Polyglutamine Peptides. *Journal of Molecular Biology* **412**, 505-519 (2011).

36. Lee, C. C., Walters, R. H. & Murphy, R. M. Reconsidering the Mechanism of Polyglutamine Peptide Aggregation†. *Biochemistry* **46**, 12810-12820 (2007).

37. López de la Paz, M. *et al.* De novo designed peptide-based amyloid fibrils. *Proc Natl Acad Sci U S A* **99**, 16052-16057 (2002).

38. Scherzinger, E. *et al.* Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: Implications for Huntington's disease pathology. *Proc Natl Acad Sci U S A* **96**, 4604-4609 (1999).

39. Chiti, F., Stefani, M., Taddei, N., Ramponi, G. & Dobson, C. M. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* **424**, 805-808 (2003).

40. Hartgerink, J. D., Beniash, E. & Stupp, S. I. Peptide-amphiphile nanofibers: A versatile scaffold for the preparation of self-assembling materials. *Proc Natl Acad Sci U S A* **99**, 5133-5138 (2002).

41. Rauscher, S., Baud, S., Miao, M., Keeley, F. W. & Pomès, R. Proline and Glycine Control Protein Self-Organization into Elastomeric or Amyloid Fibrils. *Structure* **14**, 1667-1676 (2006).

42. Aggeli, A. *et al.* Engineering of peptide β-sheetnanotapes. *J. Mater. Chem.* **7**, 1135-1145 (1997).

43. Nowick, J. S. Exploring β-Sheet Structure and Interactions with Chemical Model Systems. *Acc. Chem. Res.* **41**, 1319-1330 (2008).

44. Waters, M. L. Aromatic interactions in peptides: Impact on structure and function. *Peptide Science* **76**, 435–445 (2004).

45. Perutz, M. F. Glutamine repeats and neurodegenerative diseases: molecular aspects. *Trends in Biochemical Sciences* **24**, 58-63 (1999).

46. Dueholm, M. S. *et al.* Functional amyloid in Pseudomonas. *Molecular Microbiology* **77**, 1009-1020 (2010).

47. Perutz, M. F., Finch, J. T., Berriman, J. & Lesk, A. Amyloid fibers are water-filled nanotubes. *Proc Natl Acad Sci U S A* **99**, 5591-5595 (2002).

48. Chen, S. & Wetzel, R. Solubilization and disaggregation of polyglutamine peptides. *Protein Sci* **10**, 887-891 (2001).

49. Sikorski, P. & Atkins, E. New Model for Crystalline Polyglutamine Assemblies and Their Connection with Amyloid Fibrils. *Biomacromolecules* **6**, 425-432 (2004).

50. DePace, A. H., Santoso, A., Hillner, P. & Weissman, J. S. A Critical Role for Amino-Terminal Glutamine/Asparagine Repeats in the Formation and Propagation of a Yeast Prion. *Cell* **93**, 1241-1252 (1998).

51. Bulone, D., Masino, L., Thomas, D. J., San Biagio, P. L. & Pastore, A. The Interplay between PolyQ and Protein Context Delays Aggregation by Forming a Reservoir of Protofibrils. *PLoS ONE* **1**, (2006).

52. Dyson, H. J., Wright, P. E. & Scheraga, H. A. The role of hydrophobic interactions in initiation and propagation of protein folding. *Proc Natl Acad Sci U S A* **103**, 13057-13061 (2006).

53. Chou, P. Y. & Fasman, G. D. Conformational parameters for amino acids in helical, β-sheet, and random coil regions calculated from proteins. *Biochemistry* **13**, 211-222 (1974).

54. Sinha, S. *et al.* Lysine-Specific Molecular Tweezers Are Broad-Spectrum Inhibitors of Assembly and Toxicity of Amyloid Proteins. *J. Am. Chem. Soc.* **133**, 16958-16969 (2011).

55. Ladiwala, A. R. A. *et al.* Rational design of potent domain antibody inhibitors of amyloid fibril assembly. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 19965-19970 (2012).

56. Ventura, S. *et al.* Short amino acid stretches can mediate amyloid formation in globular proteins: The Src homology 3 (SH3) case. *Proc Natl Acad Sci U S A* **101**, 7258-7263 (2004).

57. Chothia, C. & Janin, J. Principles of protein-protein recognition. *Nature* **256**, 705-708 (1975).

58. Sitia, R. & Braakman, I. Quality control in the endoplasmic reticulum protein factory. *Nature* **426**, 891-4 (2003).

59. Pifer, P. M., Yates, E. A. & Legleiter, J. Point Mutations in Aβ Result in the Formation of Distinct Polymorphic Aggregates in the Presence of Lipid Bilayers. *PLoS ONE* **6**, e16248 (2011).

60. Sinha, S., Lopes, D. H. J. & Bitan, G. A Key Role for Lysine Residues in Amyloid β-Protein Folding, Assembly, and Toxicity. *ACS Chem. Neurosci.* **3**, 473-481 (2012).

61. Lührs, T. *et al.* 3D structure of Alzheimer's amyloid-β(1–42) fibrils. *Proc Natl Acad Sci U S A* **102**, 17342-17347 (2005).

62. Yu, Y.-C., Berndt, P., Tirrell, M. & Fields, G. B. Self-Assembling Amphiphiles for Construction of Protein Molecular Architecture. *J. Am. Chem. Soc.* **118**, 12515-12520 (1996).

63. Luo, Y. *et al.* Mice deficient in BACE1, the Alzheimer's β-secretase, have normal phenotype and abolished β-amyloid generation. *Nature Neuroscience* **4**, 231-232 (2001).

64. McGlinchey, R. P. *et al.* The repeat domain of the melanosome fibril protein Pmel17 forms

the amyloid core promoting melanin synthesis. *PNAS* **106**, 13731-13736 (2009).

65. Berson, J. F. *et al.* Proprotein convertase cleavage liberates a fibrillogenic fragment of a resident glycoprotein to initiate melanosome biogenesis. *J. Cell Biol.* **161**, 521-533 (2003).
66. Zhang, S., Holmes, T., Lockshin, C. & Rich, A. Spontaneous assembly of a self-complementary oligopeptide to form a stable macroscopic membrane. *Proc Natl Acad Sci U S A* **90**, 3334-3338 (1993).
67. Vrhovski, B. & Weiss, A. S. Biochemistry of tropoelastin. *European Journal of Biochemistry* **258**, 1–18 (1998).
68. Baer, M., Schreiner, E., Kohlmeyer, A., Rousseau, R. & Marx, D. Inverse Temperature Transition of a Biomimetic Elastin Model: Reactive Flux Analysis of Folding/Unfolding and Its Coupling to Solvent Dielectric Relaxation†. *J. Phys. Chem. B* **110**, 3576-3587 (2006).
69. Bernacki, J. P. & Murphy, R. M. Model Discrimination and Mechanistic Interpretation of Kinetic Data in Protein Aggregation Studies. *Biophys J* **96**, 2871-2887 (2009).
70. Morris, A. M., Watzky, M. A. & Finke, R. G. Protein aggregation kinetics, mechanism, and curve-fitting: A review of the literature. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1794**, 375-397 (2009).
71. Ferrone, F. Analysis of protein aggregation kinetics. *Methods in Enzymology* **Volume 309**, 256-274 (1999).
72. Griffith, J. S. Nature of the Scrapie Agent: Self-replication and Scrapie. *, Published online: 02 September 1967; | doi:10.1038/2151043a0* **215**, 1043-1044 (1967).
73. Committee on Biomolecular Materials and Processes Inspired by Biology: From Molecules to Materials to Machines. 6 (2008).
74. Baker, D., Sohl, J. L. & Agard, D. A. A Protein-Folding Reaction Under Kinetic Control.

Nature **356**, 263-5 (1992).

75. Tang, T.-S. *et al.* HAP1 facilitates effects of mutant huntingtin on inositol 1,4,5-trisphosphate-induced Ca2+ release in primary culture of striatal medium spiny neurons. *European Journal of Neuroscience* **20**, 1779–1787 (2004).
76. Kenchington, W. The larval silk of Hypera spp. (Coleoptera: Curculionidae). A new example of the cross-β protein conformation in an insect silk. *Journal of Insect Physiology* **29**, 355-361 (1983).
77. Nova, A., Ketan, S., Pugno, N. M., Redaelli, A. & Buehler, M. J. Molecular and Nanostructural Mechanisms of Deformation, Strength and Toughness of Spider Silk Fibrils. *Nano Letters* **10**, 2626-2634 (2010).
78. Knowles, T. P. J. & Buehler, M. J. Nanomechanics of functional and pathological amyloid materials. *Nature Nanotechnology* **6**, 469-479 (2011).
79. Jordal, P. B. *et al.* Widespread Abundance of Functional Bacterial Amyloid in Mycolata and Other Gram-Positive Bacteria. *Appl Environ Microbiol* **75**, 4101-4110 (2009).

CHAPTER 2: Effect of Hydrophobicity and Glutamine Content on Cooperative Protein Aggregation

2.1 FTIR Analysis of Aggregation-Induced Structural and Chemical Changes in Templated Amyloid Fibers

2.1.1 Introduction

Protein secondary structure changes were observed using time-resolved Attenuated Total Reflectance Fourier Transform Infrared spectroscopy (ATR FTIR) in both liquid and solid phase samples. This method uses a radiant energy source to excite the bonds of analyte molecules with energy at a range of wavelengths in the infrared light spectrum, from 2.5 to 25 μm or 4000 to 400 wavenumbers (cm^{-1}). Wavenumbers are directly proportional to the radiation frequency, which is calibrated with a laser. The remaining energy, after molecular absorption, is measured by a thermal detector designed for high throughput. The FTIR software reports the energy detected as transmittance as a function of wavenumber. Molecular absorbance is related to transmittance by the following equation (1), where T is transmittance, I is the energy intensity reaching the detector after interacting with the sample, I_0 is the intensity of radiation from the source, and A is absorbance.[1] Both I and I_0 must be nearly monochromatic, or representing a narrow spread of wavelengths.

$$T = {}^{I}/_{I_0} = 10^{-A} \quad (1)$$

When molecular absorbance occurs, quantized amounts of energy are used for vibrational and rotational energy transitions resulting in a change in dipole moment at a frequency that is

also quantized, corresponding to the exact energy difference of the transition. Transition energy is observed at fundamental frequencies, overtones, combination bands, or as resonance splitting bands, depending on the vibrational interactions occurring. FTIR has the advantages of flexibility, high sensitivity, and quick response time, but it lacks the atomic resolution provided by X-ray or NMR techniques, so information regarding position and noncovalent interactions of individual amino acid residues must be inferred.[2] In addition, FTIR provides more information than circular dichroism, since changes in molecular absorbance may be observed in the other regions of the spectrum, such as the fingerprint region. At room temperature, FTIR is sensitive to liquid concentrations of 10 mg/mL or more, which could be pertinent to protein processing applications.

2.1.1.1 Infrared Absorbances of Interest

Several regions of the infrared spectrum were studied because they characterize the chemical and structural changes occurring in the aggregating protein. The primary absorbances studied were Amide I, Amide II, hydrophobic packing, and the fingerprint region. The Amide I region has been best correlated with crystalline structures of model proteins.[3] Between 1700 and 1600 cm^{-1}, Amide I appears as a very strong vibration due to the asymmetric carbonyl stretch of the peptide bond, coupled to in-phase N-H bending and C-N stretching.[4] These vibrations are sensitive to secondary structure because of the large mass contributed by this particular vibration, i.e. this stretch occurs at least once per amino acid, which combines with many simultaneous interactions such as dipole hydrogen bonding or coupling between transition dipoles.[5] The result is an IR absorbance that has varied position, compound shape, and intensity.[6] As a part of Amide I, α helix absorbance is typically observed near 1650 cm^{-1}, while β sheet absorbances are typically

split between a low frequency vibration, 1610-1630 cm^{-1}, and a weaker, high frequency vibration circa 1685 cm^{-1} that is difficult to distinguish from random coil and β turn conformations.[4,6–9]

Many authors have sought to correlate IR findings to other methods to make time-resolved detection of protein structure a straightforward, reliable technique. However, there are many complications that arise from the nature of complex protein macromolecule vibrations that make this particularly challenging. Absorption frequencies and intensities can vastly differ at different solution pH or ionic strength levels due to charge-induced dipoles and protein folding changes. Proteins that are assembled into configurations involving multiple associated units have more complex spectra. Amyloid fibers are an interesting example of this phenomenon.[10]

Changes observed in the Amide I vibration reflect changes in the protein structure. In the past, a number of experiments have sought to link Amide I changes to intramolecular interactions important to secondary structure, such as hydrogen bonding. A high degree of hydrogen bonding is expected to result in a lower Amide I frequency.[11] Other bond vibrations may absorb in the same wavenumber range as Amide I. Dried samples are expected to have lower frequency than liquids by 20-40 cm^{-1}, with additional downshifts accounted for by hydrogen bonding.[12] Primary amines are expected to vibrate at 1615 cm^{-1}, which could affect all of the templating peptides, since they have high Q, N, and K content, with expected intensity in order of P7 > P4 > P4 analog.[11] Strong amino acid side chain vibrations, such as glutamine, may contribute a fraction of the total Amide I intensity, so there is potential for interference when using Amide I to quantify secondary structure during self-assembly.[13] Aromatic amino acids also have strong absorbances near the β sheet peak. Folding of some amyloids and β hairpins has been shown to depend on aromatic interactions for energetics and directionality, but such

contributions could not be determined with the methods used in this study.[14,15]

2.1.1.2 Other Absorbances

Some regions not traditionally studied in protein applications show interesting correlations with fibrillar aggregation. The Amide II vibration band is an out-of-phase combination of in-plane N-H bend deformations and the C-N stretch, with small contributions by C=O in-plane bending, C-C stretch, N-C stretch.[4,16] It is informative in protein secondary structure changes, especially in amyloid formation and liquid to solid transitions.[12,17,18] Both Amide I and Amide II bands were used together in the present experiments for structural analysis. This combined approach was optimal when used in the past for protein structure modeling that included myoglobin as a calibration protein.[19] The peak width of Amide II was expected to increase with increasing hydrogen bonding, which was tracked during time-resolved FTIR and after drying.[11] A shifting of the Amide II peak to higher frequency by 5-20 cm^{-1} was expected, with additional upshifting due to hydrogen bonding differences.[12]

Weaker vibrations such as the Amide III vibrations are also sensitive to protein secondary structure and relatively free of interference from water vibrations, but the weak intensity limits the use of this region for structural prediction.[20] This vibration is generally found at or near 1270-1290 cm^{-1} and corresponds to the in-phase combination of the C-N stretch, the in-plane N-H bending vibration, the CH_3-C stretch, and the C_α-C stretch.[9,21,22] Other vibrations at frequencies less than 1500 cm^{-1}, in the so-called fingerprint region, can differ greatly for similar molecules. This difference occurs because of the influence of subtle steric or electronic effects, especially on single bonds.[11] Coupling of vibrations is specific to each carbon skeleton. Although proteins have a repetitive backbone in the amide linkage, many amino acid side chains contain

short carbon skeletons that can exert an aggregate effect.

Other IR vibrations analyzed include those corresponding to alkyl groups, especially vibrations of methyl groups and the methylene skeleton of side chains. These absorbances can reflect subtle changes in the organization of hydrophobic amino acids during folding and aggregation. Packing of methyl groups associated with the hydrophobic protein core was observed by our group in the past to be a critical metric for assembly as fibers stabilize and elongate.[23] The 1360 cm^{-1} absorbance can be assigned to the CH_3 symmetric deformation, $\delta_s(CH_3)$, and the 1410 cm^{-1} absorbance to the asymmetric CH_3 deformation, $\delta_{as}(CH_3)$, on the side groups of A, I, L, and V.[11,24,25] Considering $\delta_s(CH_3)$ as a measure of up and down movement and $\delta_{as}(CH_3)$ as a measure of side to side movement of the CH_3 amino acid side groups, the increase in the ratio $\delta_s(CH_3)/\delta_{as}(CH_3)$ indicated that A, I, L, and V side groups were laterally constrained in the amyloid structure because they were more able to vibrate up and down relative to side to side, which was termed "hydrophobic packing". Previous reports supported this conclusion.[23,26]

Protein FTIR spectra can sometimes show peaks at unexpected frequencies due to a unique quaternary structure, such as the coiled coil conformation seen in keratin, tropomyosin, and some spider silks.[27] Careful analysis must be conducted to avoid false positives due to assumptions about an unknown protein system. Although structurally built from α helices, the coiled coil Amide I vibration contains two additional absorbances in lower wavenumbers that overlap the region that is usually characteristic of β sheet structures.[28] This happens because of fundamental differences between hydrogen bonding in single α helices compared to multiple α helices, such as coiled coils. In coiled coils, backbone C=O and N-H hydrogen bonds are

mechanically distorted as a requirement to form the helix pitch necessary for coil association.[29] Coiled coil structures may also show FTIR developments in the hydrophobic packing region, since they contain hydrophobic interactions that form a tightly-packed core between helices.[29]

2.1.2 Experimental

2.1.2.1 Sample Preparation

Pure proteins and 1:1 molar mixtures of the synthetic proteins (P4 (2.3 mg/mL), P7 (2.4 mg/mL), and P4-AN (2.3 mg/mL) with myoglobin (My) (17.6 mg/mL) were dissolved in water, adjusted to pH 8 with small volumes of NaOH, and incubated at 37°C in centrifuge tubes and monitored for 35 days (P4 and P7) and 20 days (P4-AN). Pure Gd20KK (0.96 mg/mL), CB4 (1.02 mg/mL), and a 0.36:0.64 mol/mol mixture of each with My (12.5 mg/mL) were prepared in a similar way. Molar ratio has been shown to have a small influence on kinetics, with large excess of adder protein slowing kinetics. However, between template:adder molar ratios of 0.7:0.3 and 0.3:0.7, insignificant differences in kinetics were observed with no change in the size or properties of the final structures. Mixtures volumes were 10 mL and 5 mL. At the end of 35 days, liquid mixtures were dried at room temperature on Teflon®-coated aluminum foil in a fume hood.

2.1.2.2 Fourier Transform-Infrared (FTIR) Spectroscopy

Attenuated total reflectance (ATR) FTIR spectra of the incubating aqueous solutions were recorded daily on a Thermo Nicolet 6700 FTIR spectrometer (Thermo Fisher Scientific Inc., Waltham, MA) with a 45^{o} ZnSe crystal trough and a deuterated triglycine sulfate (DTGS) detector at a resolution of 4 cm^{-1}. Samples were mixed on a vortex mixer just prior to FTIR

scanning. The aqueous solvent was subtracted from the acquired spectrum to reveal the protein spectrum.[20,30] Three spectra were acquired at each point in time and averages and standard deviations reported for each condition. Dried mixtures were scraped with a spatula and analyzed. Three spectra of the dried mixtures were recorded at each point in time with a Smart Orbit diamond ATR accessory using a flat tip and averages and standard deviations reported. The spectra were collected and analyzed with OMNIC v8.1 software.

2.1.2.3 FTIR Amide I Analysis

Most proteins are not purely of one secondary structure and thus the Amide I absorbance is formed through contributions from all of the secondary structures present. It is possible to "deconvolute" the Amide I absorbance into its secondary structure components by representing it as a sum of peaks assigned to each structure.[3,5,8,31,32] The area of each peak is representative of the molar concentration of that particular secondary structure. The spectral region from 1720−1580 cm^{-1} was isolated and manually smoothed with the Savitzky-Golay algorithm using 7-11 points.[33] Next, the second derivative of the Amide I spectral region was taken without filtering to identify the individual Amide I components. Fourier self-deconvolution (FSD) was performed with Happ-Genzel apodization by setting enhancement to 2-3 and adjusting the bandwidth until absorbance maxima matched the second derivative minima.[3,34,35] Using the "Peak Resolve" feature of OMNIC v8.1, the Amide I absorbance was fit to a series of 6-7 Gaussian/Lorentzian peaks matching the FSD.[36] The Peak Resolve feature uses the Fletcher-Powell-McCormick algorithm to fit the peaks.[37] Amide I fitting was performed with a constant baseline, a target noise of 10.0, and an initial full width at half-height (FWHH) of 3.857. There was some overlap of absorbance between Amide I and Amide II, due to amino acid COO^- asymmetric stretch and/or NH_2^+ deformation vibrations,[9] which was accounted for as an edge-

effect peak in the deconvolution fitting procedure. This overlapping area was excluded from secondary structure calculations. Model fitting was conducted with a null hypothesis of no significant difference between the spectra and the peak fit, corresponding to an F value of 0. All fits conformed to an F value of less than 0.006.

2.1.2.4 Amide II Vibrations

The peak width at half height (FWHH) of the Amide II absorbance was calculated using the built-in peak tool in OMNIC, measuring the width at half of the uncorrected peak absorbance. This procedure assumed that baseline absorbance was 0, which was confirmed by a flat line at 0.0 significantly beyond the Amide II region on either side.

2.1.2.5 Hydrophobic Packing

Peak intensities were recorded manually from OMNIC spectra at peaks representing symmetric and asymmetric methyl deformations ca. 1365 cm^{-1} and 1410 cm^{-1}, respectively. The hydrophobic packing ratio $\delta_s(CH_3)/\delta_{as}(CH_3)$ was calculated from peak intensity readings. The averages and standard deviations were reported for each condition.

2.1.3 Results and Discussion

2.1.3.1 Amide I Solution FTIR

Amide I vibrations showed a progression in β sheet development over the course of 35 day incubations for CB4 peptide mixtures. CB4 was predicted to be 74% α-helical, but a significant β sheet peak was present in the FTIR spectra right after mixing (Figure 2.1.1a). This showed high instability of CB4 in water due to monomer hydrophobicity. The α helix was lost by 25

hours due to inter- and intra-molecular rearrangement of the many hydrophobic side chains away from solvent as β sheets formed. CB4:My had high initial β sheet and α helix absorbance,

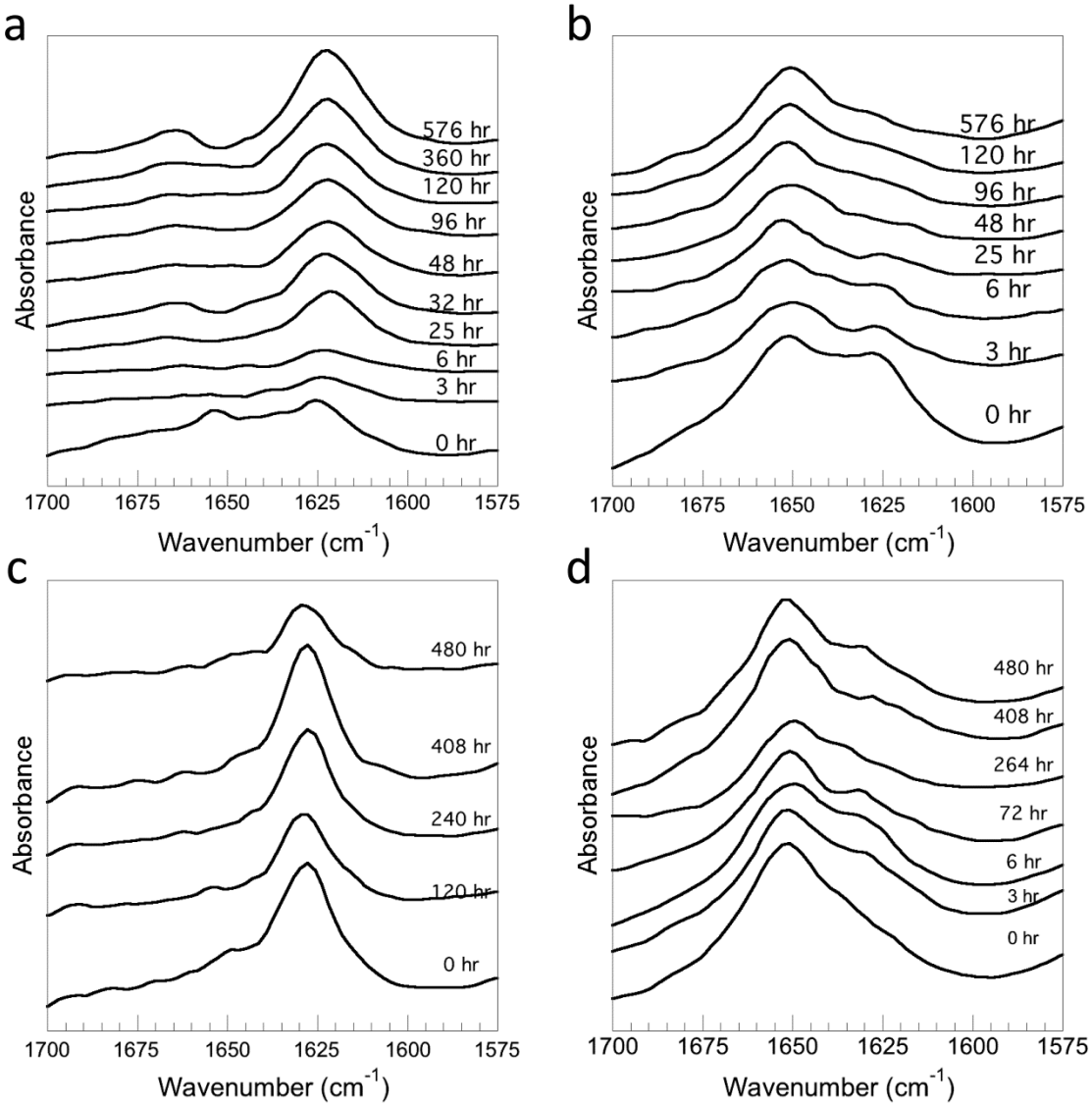

Figure 2.1.1. Amide I spectra of CB4(a), CB4:My (b), Gd20KK (c), and Gd20KK:My (d) during the incubation. CB4 (a) and Gd20KK:My (d) showed a clear decrease in α helix relative to β sheet peak area. The β sheet formed by template Gd20KK showed less development than CB4, which correlated with a lack of fibrils in AFM and SEM. CB4 showed development into a highly organized β sheet in water (a). The presence of the large adder protein seemed to stabilize the cooperative system in the case of Gd20KK:My (d), though structural developments seemed to stop after 6 hours. CB4:My (b) and Gd20KK (c) showed an initial β sheet, which decreased in intensity over time. CB4 may have been more effective in the self-templating configuration than with My, while Gd20 seemed to rely on the adder protein for stability.

corresponding to the CB4 and My components, respectively. After 3 hours, the β sheet absorbance broadened, especially in the lower wavenumber region, which would indicate the formation of the high strand density β sheets characteristic of amyloids. Some self-templating may have occurred, but a lack of significant α helix rearrangement indicated that myoglobin most likely remained in a near-native conformation. Similar CB4 and CB4:My SEM morphologies confirm the lack of rearrangement (Figure 2.2.4).

Gd20KK was predicted to be 90% α helix. Pure Gd20KK showed significant β sheet formation just after entering solution (Figure 2.1.1c), which developed more slowly over longer times as evinced by the loss of the 1650 cm^{-1} shoulder (α helix) relative to the 1625 cm^{-1} absorbance and 1610 cm^{-1} shoulder (β sheet). Gd20KK had the highest GRAVY score, with high GRAVY indicative of higher hydrophobicity, so it was expected to be highly unstable in water, even with the addition of two lysine residues. β sheet formation is typical of the formation of a template peptide that has the intrinsic capability to aggregate together with an adder protein. No pure Gd20KK fibers were observed in AFM or SEM. Micelles of pure Gd20KK may lack the water stability to assemble to larger, anisotropic β structures. In the presence of Myoglobin, Gd20KK acted as an effective template for cooperative β sheet assembly (Figure 2.1.1d). My provided enough water stability for further assembly, while interacting cooperatively with the template to transition from α to β structure. Both Gd20KK and CB4 contain aromatic residues that enable π-π stacking, whether pure or mixed with My. However, the wavenumber of phenylalanine and tryptophan residues overlaps with Amide I,[24] so those particular absorbances could not be distinguished. That particular class of noncovalent interactions was therefore

excluded.

Even though clear β sheet development occurred, at the end of a 20 day incubation, the Amide I peak of Gd20KK:My still showed considerable absorbance in the α helical range, near 1650 cm^{-1} (Figure 2.1.1d). Unaggregated My content may have been present in the mixture, which would cause α helical absorbance in Amide I. In the past, experiments with hydrolyzed gliadin and My showed that the 0.36:0.64 template to adder molar ratio had the largest conformation change. The base sequence of Gd20KK was the most hydrophobic peptide fragment resulting from tryptic hydrolysis of gliadin. It was also the most hydrophobic of all template peptides tested. These results indicate that other tryptic hydrolysis products of gliadin are important in conformation change and fiber formation when mixed with My.[23]

Amide I spectra for synthetic peptides P4, P7, and P4-AN and mixtures with My showed developments during the incubations that correlated with other findings. Pure P7 (Figure 2.1.2e) showed the most distinct β sheet development. P7 was predicted to be largely α-helical and upon initial solubilization P7 had a partially α-helical IR spectrum. Within just 5 days, the large β sheet absorbance began to split into two distinct peaks, representing low strand density β sheets (1622 cm^{-1}) and high density β sheets (1608 cm^{-1}).[2,10,38] The shift to lower frequency in the higher strand density β sheets was due to the effect of strong hydrogen bonding on the C=O stretch, which is typical of amyloid materials.[10] A lowered restoring force reduced the dipole energy, especially in the case of close packing.[24] The bimodal form of the β peak could have indicated a partitioning of backbone hydrogen bonds into stronger bonds that were interior to the growing fibril, and less dense bonds on the exterior that interacted with solvent and associating

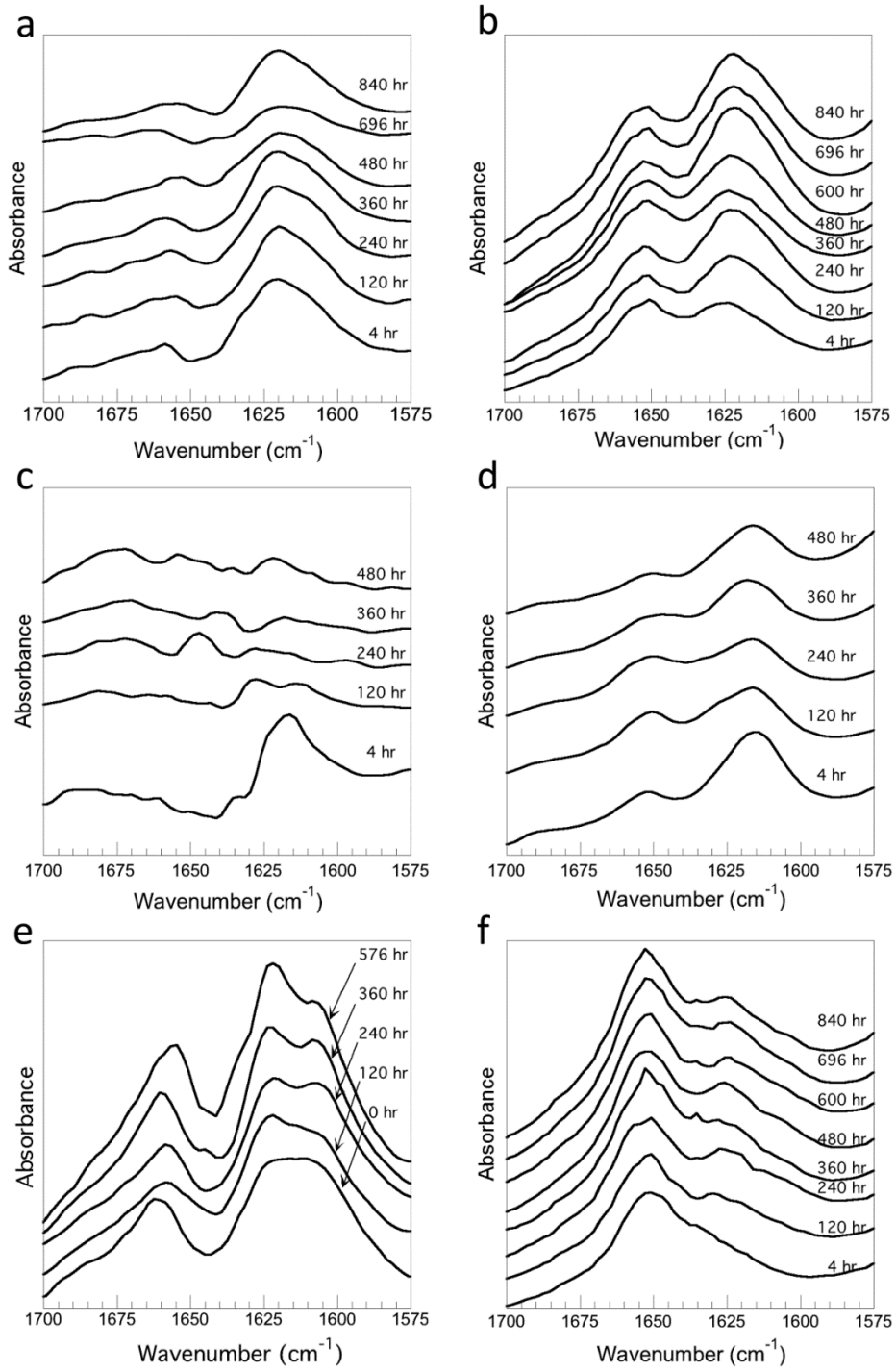

Figure 2.1.2. Amide I spectra of P4 (a), P4:My (b), P4-AN (c), P4-AN:My (d), P7 (e), P7:My(f) during the incubation. Of the pure peptides, P7 (e) showed the most distinct development of absorbance in the b sheet region, 1626 cm^{-1}. Templated peptides with My (b), (d), and (f) all showed loss of α helix absorbance at 1650 cm^{-1} relative to lower frequency vibrations, with the most dramatic changes in in P4:My (b) between 4 hr and 240 hr and. P7:My (f) also shows clear β sheet development between 4 hr and 240 hr.

peptides as fiber elongation proceeded. Hydrogen bonding was due to an attractive force between oppositely charged dipoles, and was stronger at small distances, as in Coulomb's law. Stronger intermolecular attractions led to a stronger, densely packed core. Dipole-dipole attraction only occurs down to a certain distance until nuclear repulsion dominates. It seems that Q-Q hydrogen bonds in P7 probably stayed in the attractive regime since a crystalline morphology was observed (Figure 2.1.1).

Pure P4 showed an initial β sheet peak which decreased in intensity somewhat over time. However, no distinct fibrils were observed for P4 in AFM or SEM. P4 peptides most likely formed globules stabilized by Q-Q bonds that were unable to self-assemble to a larger structure. The absorbance of the Q side chain may have contributed to a shoulder of the large β sheet peak observed for pure P4 and P7 at all times. However, the My mixture data is more consistent with changes in the peptide bond C=O stretch. P4-AN peptides showed a high initial β sheet absorbance when mixed in water, as predicted in PSIPRED. However, the P4AN spectra quickly became indistinct due to a lack of strong, regular hydrophilic interactions such as hydrogen bonding or salt bridging. Hydrophobicity was high in P4-AN, but the other hydrophilic residues were irregular in structure and in dipole strength, so side chain-side chain bonds and side chain-backbone bonds were unable to outcompete solvent-amino acid interactions, i.e. dissolution occurred. Some assembly progressed for P4AN peptides upon drying, but tapes were thin and required drying for intramolecular interactions to dominate other interactions (Figure 2.2.8a).

All My mixtures (Figure 2.1.2b, d, and f) showed increase in β sheet formation, which was observed as a loss of α helix absorbance at 1650 cm^{-1} relative to lower frequency β sheet vibrations. The most dramatic changes occurred between 4 hours and 240 hours for P4:My

(Figure 2.1.2b) and P7:My (Figure 2.1.2f). P4:My developed from equal amounts of α and β structures to mostly β sheet structure. This is most likely due to a slow, progressive unfolding and cooperative aggregation of the larger, more stable My adder protein. In the dried fibers, very little α helix seems to remain (Figure 2.1.6 and Figure 2.2.7). The Amide I absorbance showed a progressive increase in the 1628 cm^{-1} absorbance and the appearance of a shoulder at 1606 $cm^{-1.}$

Thus, β sheet aggregation occurred but there was more low strand density β sheet than high strand density β sheet. However, P7:My retained mostly α helical structure. This tendency seemed due to dominance by My, which was in stark contrast to the high initial β sheet content of the template spectra. P7 appeared to exclude bonding with My in preference for template-template interactions. By self-assembly in solution, P7:My developed a degree of β sheet structure, which was retained upon drying. P4-AN:My was relatively unchanged throughout incubation, except for fluctuations which might have indicated local rearrangements or intermediate structures.

2.1.3.2 Amide II

The peak width of the Amide II vibration, centered near 1550 cm^{-1}, was found to increase significantly for certain liquid mixtures as aggregation proceeded (Figure 2.1.3). The mixtures were strong fibril formers (P4:My, P4-AN:My, P7:My, Gd20KK:My, and CB4:My) and showed a strong linear increase in the Amide II FWHH over the entire time of observation (up to 35 days). The only pure peptide to display similar behavior was P7 (Figure 2.1.3b). Good linear fits were obtained for those six species for long time incubations, as shown below. Additional species are shown in A.3. The best fits were obtained for My mixtures with amphiphilic

sequences, while much lower R^2 values were obtained for CB4 and Gd20KK. This indicated that FWHH increase may require some degree of amphiphilic patterning. P4 and P4-AN had similar GRAVY scores to CB4, so hydrophobicity could not be implicated in FWHH increase. The linear fits of FWHH vs. time may be used for prediction of aggregation progress.

P4-AN:My had sufficient template hydrophobicity to assemble with My via the cooperative mechanism. These mixtures formed fibrils with a surprisingly high elastic modulus (Figure 2.2.9). Nanoindentation results confirmed that Q blocks were not necessary for strong intermolecular interactions and cooperative fibril formation. Higher FWHH may be a simple consequence of high β sheet character, or it could be due to increased conformational

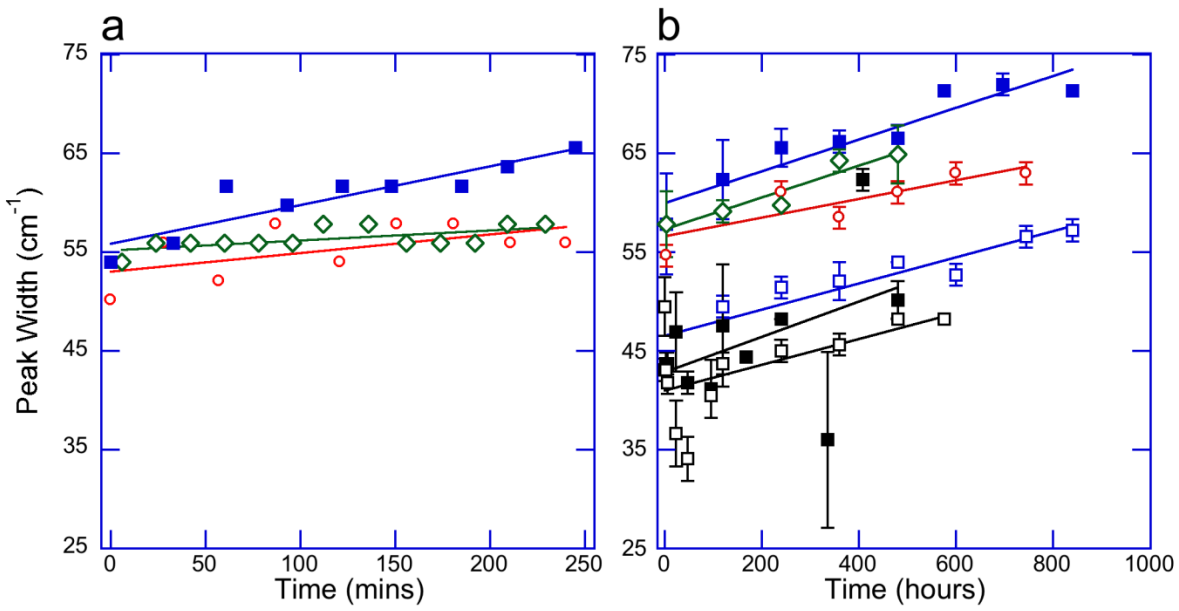

Figure 2.1.3. Amide II FWHH. This peak showed a strong increase especially in aggregation over long times. Samples are ■ P7, □ P7:My, ○ P4:My, ◇ P4-AN:My, □ CB4:My, ■ Gd20KK:My. In long time trials (b), relatively good linear fits allow prediction of FWHH based on aggregation progress. For P7, *FWHH* = 60.0 + 0.016 *t* (r^2=0.89); for P7:My, *FWHH* = 46.5 + 0.013 *t* (r^2=0.86); for P4:My, *FWHH* = 56.6 + 0.0095 *t* (r^2 = 0.73); for P4-AN:My, *FWHH* = 57.3 + 0.016 *t* (r^2=0.91); for CB4:My, *FWHH* = 41.0 + 0.013 *t* (r^2=0.31); and for Gd20KK:My, *FWHH* = 42.9 + 0.018 *t* (r^2=0.22). In short time trials (a), a smaller slope is apparent compared to long time trials (b). For short time line fits, P7 *FWHH* = 55.9 + 0.039 *t* (r^2 = 0.81); P4:My *FWHH* = 53.0 + 0.019 *t* (r^2 = 0.33); and P4-AN:My *FWHH* = 55.2 + 0.01 *t* (r^2 = 0.40).

flexibility.[4,19] The assembly process is a flux of molecules toward a thermodynamically favorable structure, which requires breaking less favorable interactions with solvent molecules and between amino acid side chains and sequestration of hydrophobic groups in order to overcome the unfavorable entropy of the ordered structure. The assembly process may be characterized by vibrational heterogeneity, at least while fibers are not yet in equilibrium.

2.1.3.3 Hydrophobic Packing

CB4:My mixtures showed hydrophobic packing, which appeared as an increase in the ratio of symmetric to asymmetric methyl deformation vibrations, $\delta\ (CH_3)_{,sym}/\delta\ (CH_3)_{,asym}$ (Figure 2.1.4). Gd20KK:My showed a very strong increase in hydrophobic packing during incubation, which was evidence of an increase in favorable enthalpy of β sheets. Conformational changes were not as pronounced as hydrophobic packing of these systems. The final product of hydrophobic packing interactions was shown in the dried fibers.

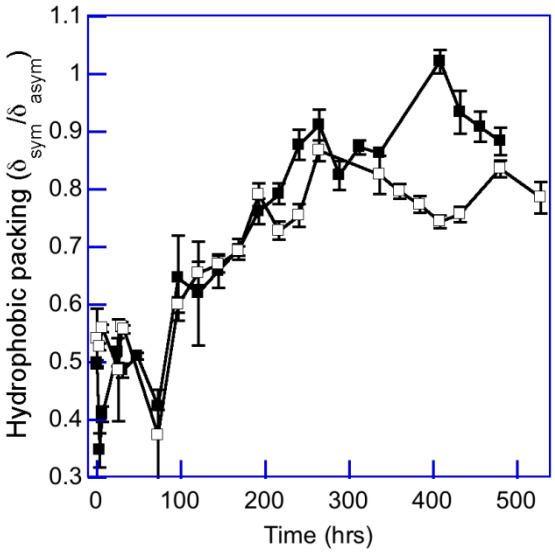

Figure 2.1.4. CB4:My □ and Gd20KK:My ■ showed strong hydrophobic packing as assembly proceeded. Both templates had a high degree of aliphatic amino acid side chains in their sequence, which were distributed randomly, which meant that there was a higher likelihood of packing between each beta strand, compared to the amphiphilic patterned sequences of the synthetic peptides.

Synthetic protein mixtures did not show nearly as much gain in hydrophobic packing as Gd20KK and CB4 systems, except for P4-AN (Figure 2.1.5, right ordinate). Although P4 and P4-AN had identical hydrophobic amino acid content and arrangement, P4-AN showed 7-fold hydrophobic packing increase while P4 showed almost no change. This indicated that the hydrophilic region had a large effect on folding in P4-AN. The Q-Q hydrogen bonds of P4 interacted mainly with each other, while the hydrophilic amino acids of P4-AN seemed to assist in hydrophobic sequestering and hence packing, perhaps by interacting with solvent molecules to exclude them from hydrophobic blocks. The difference in hydrophobic packing could indicate an important role for polar hydrogen bonds in stabilizing the outer structure while hydrophobic interactions stabilized the core. However, without the regular structure of Q-Q hydrogen bonding, P4-AN was greatly affected by the variability of the amino acids in its hydrophilic region, which led to a large variability in the hydrophobic packing peak intensities.

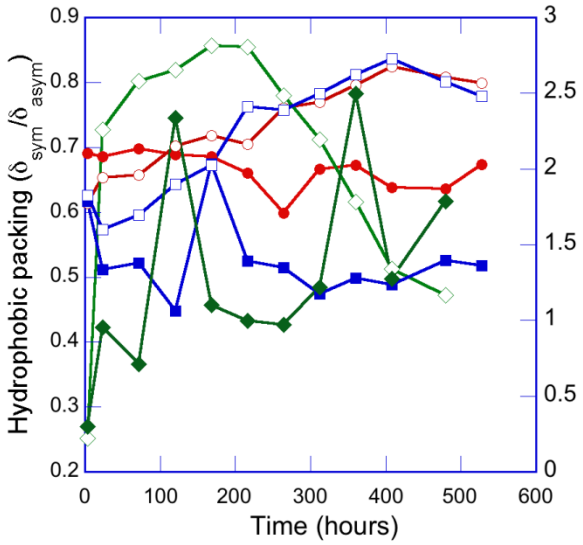

Figure 2.1.5. Hydrophobic packing of P4 ●, P4:My ○, P7 ■, P7:My □, and P4-AN:My ◇ (plotted on the left ordinate) and hydrophobic packing of P4-AN ◆ (plotted on the right ordinate). Although most mixtures showed developments between 0.3 and 1, P4-AN showed much higher values, but with much higher variability, perhaps due to high tendency of charged stabilizing side chains to rearrange. P4-AN:My shows asymmetric packing dominance after 200 hours, showing the impact of non-Q amphiphilic residues on packing.

The hydrophobic packing behavior of P4-AN:My showed a large initial jump in packing, presumably as β sheets formed to exclude water from hydrophobic regions. Interestingly, P4-AN:My showed a decrease in hydrophobic packing after 200 hours of assembly progress, which meant that self-assembly of cooperatively templated structures proceeded differently in systems without Q. Asymmetric bending, ie. the side-to-side deformation, may have been more important in packing rearrangements that occurred as P4AN:My fibers assembled.

2.1.3.4 Fingerprint Region

The region containing vibrations associated with single bonds occurs at lower frequency due to the lower energy of their dipole transitions.[11] In both P7 and P7:My, aggregation progress occurred simultaneously with an increase in the intensity of a band at 1315 cm^{-1} (Figure 2.1.2e and f). That vibration was probably due to skeletal wagging or rocking of glutamine-associated methylene groups, $\gamma_w(CH_2)$,[11] perhaps as a counter to movement of carbon-carbon stretching during rearrangement of Q residues to achieve optimal hydrogen bonding. However, P7:My had additional absorbances that correlated with self-assembly progress at 1417 cm^{-1} corresponding to methylene deformations, $\delta(CH_2)$.[24] That was likely another indicator of hydrophobic packing of neighboring carbon skeletons, which could have involved all amino acids in P7 and My. A vibration at 1250 cm^{-1} decreased with aggregation, which might correspond to $\delta(CH)$,[24] which indicates decreased tertiary carbon movement as beta sheets form and stabilize. This was probably due primarily to aggregation in the main chain. In P4-AN, vibrations at 1339 cm^{-1} increased with time. P4-AN:My showed a similar vibration at 1349 cm^{-1}. This vibration may have been due to histidine (H) ring stretching due to packing or interactions.[24] P4-AN:My also showed β-sheet aggregation at all times with a progressive shifting of the 1400 cm^{-1} absorbance

to 1410 cm^{-1} with time, and a concurrent loss of the 1365 cm^{-1} absorbance. The 1400 cm^{-1} absorbance was assigned to $\nu_s(COO^-)$ and the behavior would indicate ionic bonding through glutamic acid (E) on P4-AN.[24,39] Absorbances from and NH^{3+} deformation vibrations were also expected for salt bridging with charged lysine residues, but their loci at 1525 cm^{-1} (asymmetric) and 1625 cm^{-1} (symmetric) overlapped with the stronger Amide I and Amide II vibrations.[24]

2.1.3.5 Dried Aggregate Spectra

The greatest changes in absorbance occurred between the incubating solution and dried fiber. Dried P7 showed greatly reduced CH and CH_3 absorbances at 1400-1300 cm^{-1} relative to 1200 cm^{-1} and 1130 cm^{-1} absorbances, and a new absorbance appeared at 1180 cm^{-1} (Figure 2.1.6d). Dried P7:My had reduced absorbance at 1315 cm^{-1} and 1250 cm^{-1} relative to those at 1200 cm^{-1} and 1145 cm^{-1} (Figure 2.1.6k). For both P4-AN and P4-AN:My, all absorbances were greatly reduced relative to those at 1200 cm^{-1} and 1140 cm^{-1} and a new absorbance appeared at 1180 cm^{-1} (Figure 2.1.6f and n). P4-AN did not contain Q therefore the new absorbance at 1180 cm^{-1} observed in P4-AN, P7, and P4 was assigned to CH deformations. All dried, My-containing mixtures showed a dominant, new absorbance at 1200 cm^{-1}, which was most likely the C_α-C stretch, shifted down in frequency due to strong hydrogen bonding which lowered the restoring force of the stretch.[9,24] Although this stretching vibration occurred in all dried mixtures, it was only developed in the My mixtures upon drying. This vibration probably indicated a degree of tension between the backbone and side chains, which may be a result of packing to accommodate the aggregated amyloid structure. Varying degrees of torsion are shown in SEM images of twisted dried fibers (Figure 2.2.4 - Figure 2.2.8). For My mixtures, drying appeared to be an important part of the cooperative aggregation process. The templating mechanism was supported

by the presence of the fingerprint region vibrations in the synthetic peptides before the drying process, at the end of their self-assembly in solution. Templates were self-assembled by this time in the twisted, aggregated conformation. However, in solution, My was recruited by these templates but could not fully assemble to them without an additional thermodynamic input,

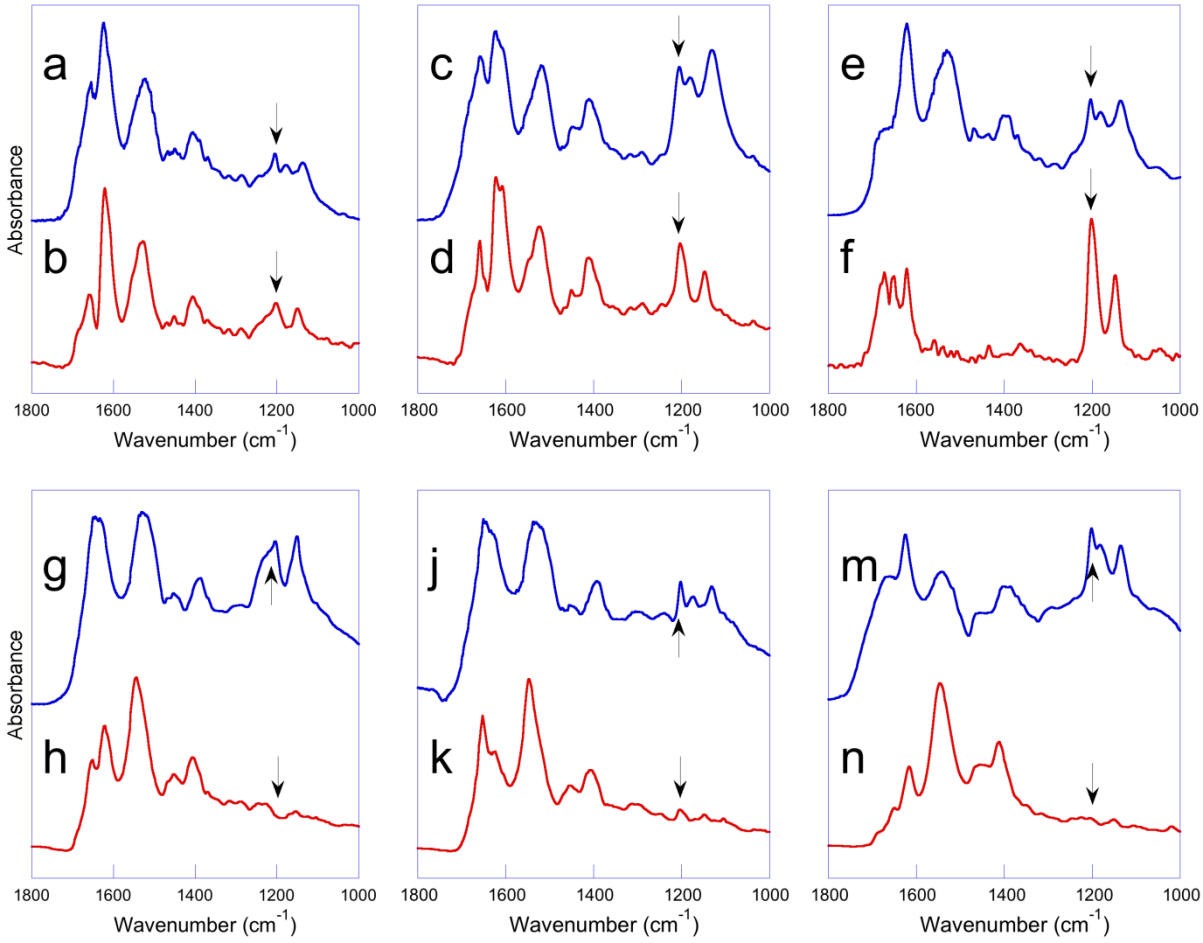

Figure 2.1.6. Dried (blue) and liquid (red) FTIR spectra show large changes occurred during drying of aggregates. Shown are P4 day 35 (a), P4 dried (b), P7 day 35 (c), P7 dried (d), P4-AN day 20 (e), P4-AN dried (f), P4:My day 35 (g), P4:My dried (h), P7:My day 35 (j), P7:My dried (k), P4-AN:My day 20(m), and P4-AN:My dried (n). Upon drying, the intensity of absorbances increased greatly in the fingerprint region. This was a new development for My-containing mixtures, which indicated that drying was necessary to complete self-assembly with the cooperative mechanism.

which occurred during drying. Gd20KK:My and CB4:My showed less distinct changes in drying (Figure 2.1.7). Fingerprint region vibrations were reduced relative to the vibration near 1400 cm-

[1]. This could indicate dominance of the asymmetric methyl vibration, which might imply that the geometry of inter-sheet hydrophobic interactions were in a head-to-head configuration once solvent was excluded.

Changes in Amide I and Amide II in drying were also distinct. P7 retained a distinct high density β peak showing Q-Q hydrogen bonding at 1603 cm^{-1}, which was a few wavenumbers lower than the solution spectra (Figure 2.1.6c and d). Both Q-containing templates P4 and P7 had a distinct peak at 1660 cm^{-1} that most likely corresponded to an organized structure, such as β turn. Typically, β turns are located at higher wavenumber but are likely involved in a number of

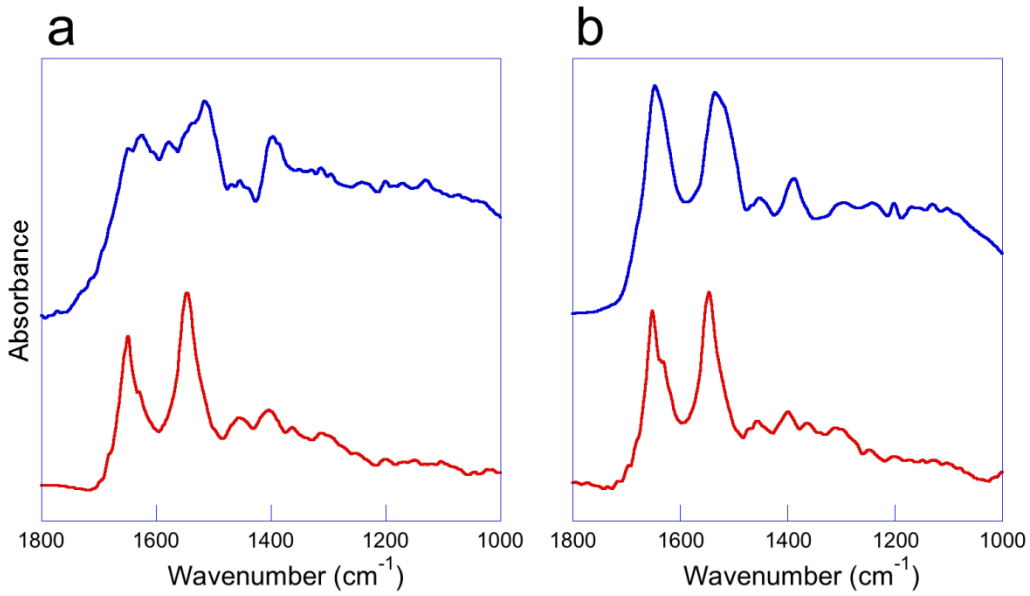

Figure 2.1.7 Dried (blue) and solution (red) FTIR spectra show less distinct changes upon drying of (a) CB4:My and (b) Gd20KK:My. Most likely, solvent exclusion occurred in these systems prior to drying because they lacked hydrophilic amino acid blocks. Hydrophobic packing changes showed the dominance of the peak near 1400 cm^{-1} upon drying relative to the peak at 1360 cm^{-1}.

noncovalent interactions that decrease the C=O stretching frequency. Turns located at β sheet edges were likely to be an important area for sheet-sheet packing and solvent exclusion. P4-AN instead showed a flatter shoulder, but had an identical β sheet peak to P4 at 1623 cm^{-1} (Figure

2.1.6e). In P4:My and P7:My, the Amide I spectra showed a broader, intense vibration with a center spanning from 1625 to 1650 cm^{-1} (Figure 2.1.6h and k). Since the same absorption appeared for assembled solution phase and dried phase, and there was no visible separation of secondary structural vibrations, it was inferred that a single conformation dominated, such as the twisted β sheet. As in the case of the coiled coil quaternary structure described previously, the effect of torsion and packing could have caused some atypical bands to appear. Amide II also showed a nearly identical intensity, shape, and position in liquid and dry samples. This indicated that the influence of Q-related hydrogen bonding on the peptide bond conformation was similar in P4 and P7. This was reflected in nearly identical nanoindentation results (Figure 2.2.9). This result also implied that the FTIR spectra in Amide I did not adequately convey the differences in length of Q block that were seen in the dried morphology. P4:My and P7:My may have assembled fibers that had similar chemical properties, but their differing morphologies points to factors other than secondary structural arrangements, such as the geometry of the final template-adder aggregate.

2.1.4 Conclusions

2.1.4.1 Influence of Glutamine

The highest Q-containing templates (P7) had very distinct β sheets which split into low and high strand density absorbances. Higher strand density absorbances were coincident with P7 crystallization from strong Q-Q hydrogen bonding, which caused the peaks to shift to lower wavenumber. When myoglobin was present with the P7 template, little α to β transition occurred, so the high glutamine peptide showed a preference for self-binding rather than with myoglobin. P7 was thus not effective as a template. A low level of glutamine in some unordered

templates was insufficient for amyloid development. CB4 and Gd20KK peptides with low levels of Q formed β sheets, but this is common among many proteins in water. At a higher level of Q in P4 (40%), templates did not form amyloids, losing β sheet structure over time, presumably due to hydrogen bonding with water.

2.1.4.2 Influence of Myoglobin

Myoglobin lent unstable templates Gd20KK, P4, and P4-AN a degree of water-stability that led to self-assembly via cooperative binding, which was shown in the α → β transition. In the absence of myoglobin, this transition occurred only for templates that had a high potential for Q-Q hydrogen bonds. Some mixtures showed evidence for unaggregated myoglobin in their FTIR spectra, which appeared as a residual peak at 1650 cm^{-1}. The presence of myoglobin in association with templates caused torsion in final dried structures, shown in a new C_{α}-C stretch in the dried aggregate.

2.1.4.3 Influence of Hydrophobicity

Template sequence affected changes in the hydrophobic packing ratio the most in the myoglobin mixtures. The least hydrophobic template P7 had the smallest hydrophobic packing increase, while all others showed that symmetric vibrations were favored over asymmetric. That meant that hydrophobic packing was a good indicator of aggregate core packing in proteins that were less solution-stable.

In pure synthetic templates, hydrophobicity caused unfolding in myoglobin at quantities which correlated with hydrophobic amino acid content in the template. Loss of α helix was instant in P4:My and P4-AN:My, which had templates of high hydrophobicity, while in P7:My,

myoglobin unfolding was somewhat slower. Myoglobin unfolding was an important prerequisite for cooperative assembly. However, in Gd20KK and CB4 mixtures with randomly patterned templates, myoglobin was less well-recruited to the aggregate, which was shown in the lack of strong $\alpha \rightarrow \beta$ transitions. The importance of hydrophobic patterning was seen in P4-AN:My, which had a distinct hydrophobic packing decrease after 200 hours but strong fiber morphology. In P4-AN, the hydrophilic portions of the pattern varied in size and charge, so hydrophobic packing favored asymmetric vibrations, perhaps due to a larger inter-β sheet distance in water.

2.2 Characterization of Morphology and Elastic Strength of Cast and Dried Templated Amyloid Fibers using Atomic Force Microscopy, Scanning Electron Microscopy and Nanoindentation

2.2.1 Introduction

2.2.1.1 Atomic Force Microscopy

Atomic force microscopy is used in characterization of protein morphology. This instrument generates a three-dimensional image of the protein using cantilever interactions with a biological surface. Thin films of experimental protein aggregate solutions were observed during fiber assembly using two scanning modes: contact mode and tapping mode. Contact mode is based on adjustment of the force resulting from the interaction between the cantilever and surface to a set point at which the tip of the cantilever stays pressed to the surface. A map of height information is derived from deflection of the cantilever as the force changes, which records the topology of an area limited to 100μm x 100μm. Such changes in the cantilever are sensed through a force

transducer, results in feedback adjustment of the cantilever position through a piezoelectric element. The measured force from cantilever displacement can also be used in material property determination.[40] Lateral twisting due to specific material sensitivity to deformation or texture may also be detected using piezoelectric elements in the x and y directions.[41] When operated in tapping mode, the cantilever tip oscillates just above the surface in contact with the sample force field, i.e. attraction due to a contamination layer, condensation, or charge. Tapping mode applies a smaller force to the sample surface, which is critical for soft samples that may be damaged, compressed or detached.[41] When used with the protein systems studied, tapping mode had excellent resolution while minimizing sample damage and artifacts, which were observed in contact mode. Probe-sample force may also be reduced by operating in a hydrated environment.

We have previously observed the self-assembly of large amyloid fibers, i.e., on the micron scale, to be a hierarchical process beginning at the protein molecule level. AFM was used to observe the hierarchy as a function of self-assembly time. The hierarchical process was observed in four steps where 1) protein molecules formed 10-30 nm protofibrils, 2) protofibrils aggregated into 100-200 nm fibrils, 3) fibrils aggregated into 400-800 nm large fibrils, and 4) large fibrils aggregated into 10-20 μm fibers.[42] AFM makes it possible to measure structures in steps 1-3. Fiber morphology and dimensions such as length, width, and pitch (if twisted) can be measured in AFM and used to quantitatively describe fiber structure. For instance, protofibril, fibril, and large fibril width are especially important in characterizing how each fits into the hierarchy observed in large amyloid fibers.[43,44] Other measures such as persistence length can indicate material stiffness and short-range axial interactions of the protein polymer, effectively expressing the energy needed to bend the fibril.[40,45] Persistence length correlates with assembly

progress and high beta content.[46]

2.2.1.2 Scanning Electron Microscopy

Scanning electron microscopy (SEM) is a technique used to image assembled structures on a larger scale, i.e. μm-mm. In this technique, samples are placed in a vacuum where a 1-20 kilo-electron volt accelerated electron beam sweeps in a rectangular raster pattern, interacting with the fibers to a depth of ~10nm. This energy beam generates a wide variety of detectable radiation.[47] In Field Emission SEM (FESEM), secondary electrons are detected and interpreted by an energy-dispersive X-ray detector and software. FESEM provides a large depth of field, which is more ideal for three-dimensional, low-moisture structures like fibers. Organic degradation is a common issue in electron microscopy of organic samples, so fibrous samples are exposed to lower energy electron beams and are coated with conductive, metallic coating to prevent charging.[47] Final, micron-sized fibers achieved in step 4 were imaged in FESEM. Similar measurements of fiber features can be made using software, as in AFM.

2.2.1.3 Nanoindentation

Nanoindentation is used to measure the physical properties of very small protein fibers that cannot be tested using macroscopic tensile or shear equipment. In nanoindentation, small samples are indented with a conical or pyramidal diamond tip where the tip dimension is on the order of 100 nm. The tip is pressed into the sample and the total force (F) on the tip is measured as a function of tip displacement (δ). The force on the indenter includes repulsive force and any collapse that occurs, on loading and unloading. Many indentation measurements of the background substance and locations in the material are made as a screening process so that the

sample can be identified by force-distance curve alone. Distinguishing the surface topography using indentation is important because indentation trials of the background or even the edge of a fiber can be misleading or can wrongly characterize the fiber. The elastic stiffness, or Young's Modulus, can be found from the F-δ curve. Stiffness of amyloids is important because their high modulus may explain the persistence of aggregates observed in disease states.[48,49]

The equation used to calculate the fiber elastic modulus from nanoindentation trials was derived by Oliver and Pharr,[50] based on Hertz's studies of contact deformation mechanics. The reduced Young's modulus E_r is expressed as a function of the measured stiffness S and tip-sample contact area A (Equation 2). Contact area differs by tip shape, but a 90° conical tip was used for the present studies. For this tip shape, the area is a function of δ (Equation 3). Reduced modulus is related to the modulus of the tip E_t and modulus of the sample E_s by the Poisson's ratios, ν_t and ν_s, respectively (Equation 4). Poisson's ratio indicates the change in volume on extension of a material body. The tip used was rated at a modulus of 1140 GPa and a Poisson's ratio of 0.07, which means there was almost no lateral contraction exhibited upon elongation. An assumed Poisson's ratio of 0.3 was used for ν_s, based on use in other amyloid studies and in typical polymers.[51]

$$E_r = \frac{\sqrt{\pi}}{2}\frac{S}{\sqrt{A}} \tag{2}$$

$$A = \sqrt{2}\pi\delta^2 \tag{3}$$

$$\frac{1}{E_r} = \left(\frac{1-\nu_t^2}{E_t} + \frac{1-\nu_s^2}{E_s}\right) \tag{4}$$

Young's modulus may also be measured by peak force loading using atomic force microscopy, which has the advantage of measuring aggregates that have not gone through complete assembly by drying.[48] Atomic force microscopy was used in the present study for surface imaging, but future work using the cantilever for force measurements will enable more accurate linking of surface topography to surface hardness and elasticity. Though fibers were measured at their most assembled and presumably most suitable state for industrial material applications, materials property progression from oligomer to protofibril and the dried product are of interest and may be simulated.[38] Such simulation was outside the scope of this work.

2.2.2 Experimental

2.2.2.1 Sample Preparation

Every 10 days, 50 μL of vortexed solution was spin coated onto freshly cleaved mica discs for 1 minute at 4000 rpm and AFM imaging performed. At the end of 20-35 days incubation, liquid mixtures were dried at room temperature on Teflon®-coated aluminum foil in a fume hood. Dried fibers were extracted using a stereoscope, fixed to aluminum specimen mounts with double-sided tape, and sputter-coated with gold-palladium for 40 seconds prior to SEM imaging.

2.2.2.2 Atomic Force Microscopy

Images of the deflection, lateral, and topographical surface signals were obtained using an Innova AFM (Bruker, Santa Barbara, CA) in contact mode with a 1-10 Ω-cm Phosphorus doped Silicon tip, or in tapping mode with a 0.01-0.025 Ω -cm Antimony (n) doped Silicon tip (Veeco Instruments, Inc., Plainview, NY). Final scans were performed at a rate of 0.1 Hz and a scan density of 512 scans per image. Image processing was performed using NanoScope Analysis

software V1.40 (Bruker). Images were fit to a 2nd order plane prior to surface profile image capture using the contact mode deflection signal or tapping mode phase and amplitude signals. 3-dimensional images and tangential surface slice measurements were generated using topographical signal images.

2.2.2.3 SEM Imaging

Scanning electron micrographs were generated using a LEO 1550 field-emission scanning electron microscope (Zeiss, Peabody, MA). An accelerating voltage of 5 kV with a working distance of 12 mm and In-Lens SE Detector were used. Images were evaluated and analysis performed by measuring fiber dimensions to scale using ImageJ v1.46r software (National Institutes of Health, Bethesda, MD). JMP statistical software was used for data analysis (SAS Institute Inc., Cary, NC).

2.2.2.4 Nanoindentation

Indentation experiments were performed on final, dried micron-sized fibers at room temperature and humidity using a Hysitron Triboindenter (Minneapolis, MN) with a 90° conical diamond tip. Experiments were in displacement-controlled (DC) mode with a maximum displacement of 1000 nm at a rate of 100 nm/s. Loading direction was normal to the fiber axis. Fiber elastic (Young's) modulus, *E*, was determined as previously reported from an average of 19, 14, 18, 18, 13, and 13 indents (for P4, P4:My, P7, P7:My, P4AN, and P4AN:My, respectively) that fit the expected F-δ curve.[50] All raw measurements are presented in Table D.2 - Table D.7 for reference.

2.2.3 Results and Discussion

2.2.3.1 AFM Images

Although it has been used successfully with other systems, AFM captured solution phase fibers for only some mixtures in the present experiments. This was probably due to the reported dependence on drying for full fiber assembly. P7 showed anisotropic structures to a much larger extent than P4 (Figure 2.2.1). No other protein systems showed these structures so they were ascribed to extensive crystallization through Q-Q bonding. These structures also showed a high degree of twist in SEM (Figure 2.2.6). Crystalline P7 structures showed a resistance to breakage evident in the very long length of SEM fibers. At the end of a 20 day incubation, Q-free P4-AN showed some oligomers, but no fibrils (Figure 2.2.2). The lack of amyloid fibrillation seen in AFM was also evident in a lack of P4-AN Amide I development in FTIR spectra (Figure 2.1.2).

Fibril-like structures were observed in My-containing mixtures only that were the result

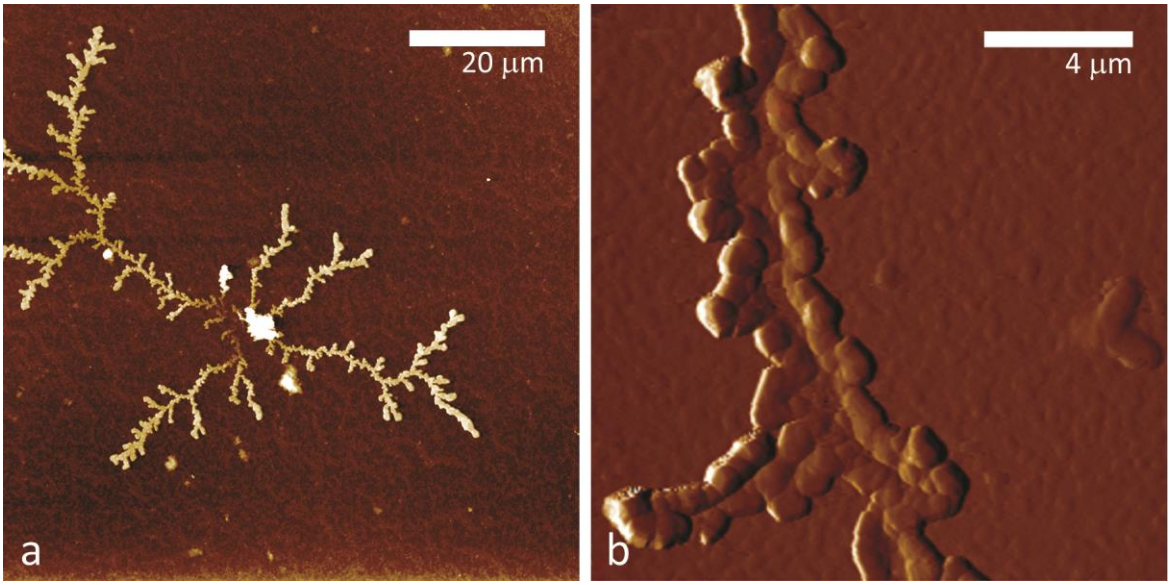

Figure 2.2.1. Tapping height images of P7 showed crystallization after 20 days incubation due to extreme self-association of Q blocks. The Q blocks of P7 had the potential for noncovalent interactions in both lateral and vertical directions due to propionamide side chain hydrogen bonding with the protein backbone and other side chains. P4 shows similar structures but to a lesser extent.

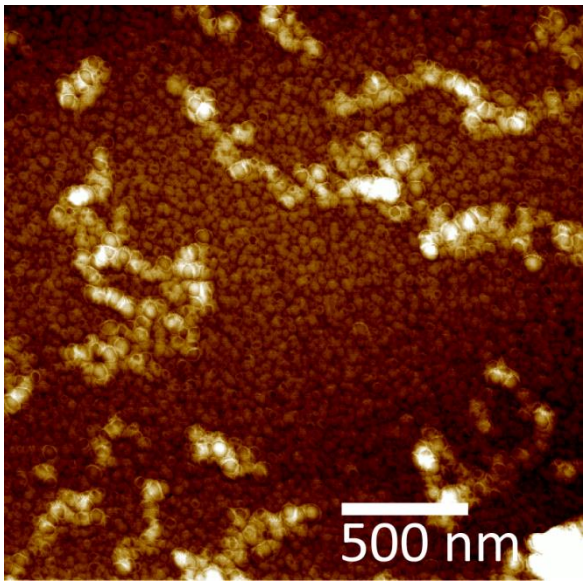

Figure 2.2.2. Tapping mode height image of P4-AN after 20 days incubation. Only oligomers formed from self-templating, which confirmed FTIR results that showed that the water-stability of My was needed for further P4-AN assembly.

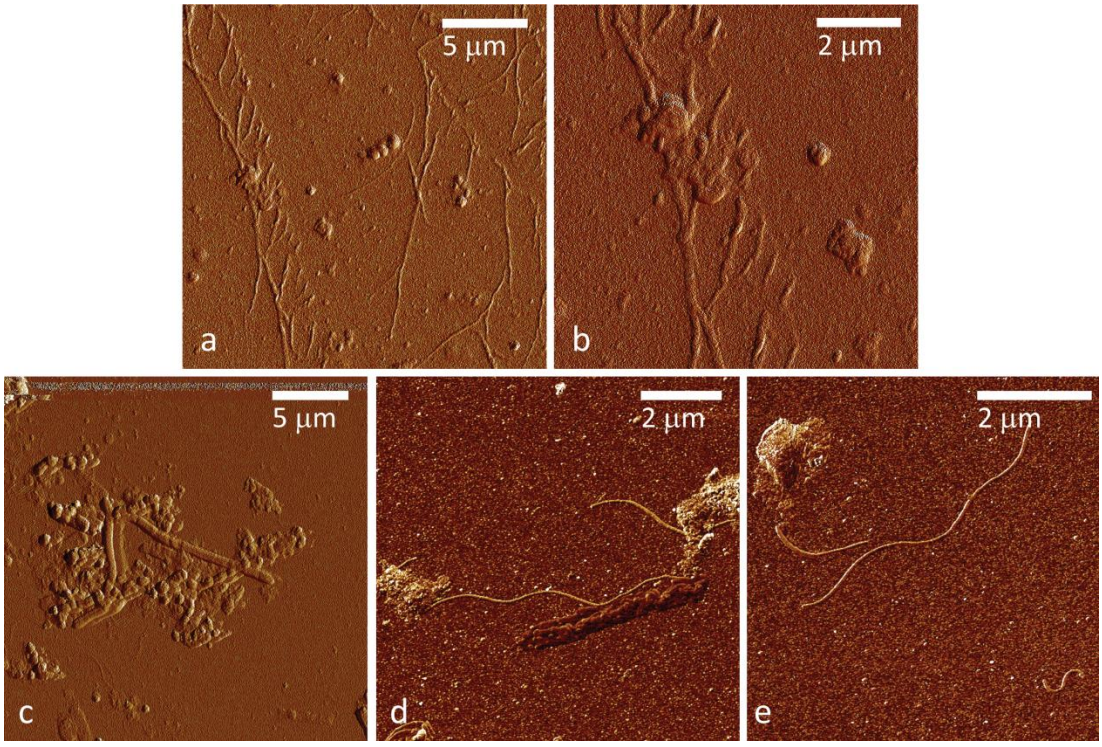

Figure 2.2.3 AFM scans in tapping mode. These images show that two contamination-related morphologies found in the same P4-AN:My sample that are easily distinguishable from the straight fibrillar aggregates typical of rigid amyloids. (a) and (b) show rigid fibers on a larger and smaller scale that are most likely amyloid. Also present were rod-shaped microbes dividing (c) and singular (d), both in the presence of smaller, helical flagella. These images show a similar structure to (e), which is most likely a separated flagellum.

of microbial contamination (Figure 2.2.3). These structures showed two morphologies: the first resembled the short, thick cell bodies which were attached end to end, as in reproducing E. Coli (Figure 2.2.3c),[52] and the second resembled thinner, helical flagella (Figure 2.2.3d and e).[53] Flagella may have been present without the cell body due to shear from vortexing of samples prior to loading in FTIR. These fibril-like structures represent a false positive because dimensions measured in AFM scans were similar to literature values for E. Coli and flagella.[54,55] Such structures were thus excluded from analysis of amyloid fibril properties and dimensions. However, the presence of microbes highlights an important factor in such protein aggregation experiments.

An excess of proteins can support the growth of biological organisms.[56] The organisms that grow in these settings may produce anisotropic structures as a part of their normal growth processes, but since the solution conditions provide very low nutrients, some of their normal processes may be altered.[57] The proteins used in the present experiments show a degree of digestibility by yeast in vivo, mammalian reticulocytes in vitro, and E. Coli in vivo, as determined by convention using the N-terminal amino acid, which is a part of the ExPASy server

Table 2.2.1 Summary of protein half-life indicated contamination was likely. However, it was observed primarily in My-containing systems.

Protein	Half life (hours)		
	Yeast	Mammalian Reticulocytes	E. Coli
Myoglobin	>20	30	>10
P4	0.05	5.5	0.03
P7	0.17	0.8	10
P4-AN	0.05	5.5	0.03
CB4	>20	1.9	>10
Gd20KK	>20	30	>10

analysis tools.[58] Although the half-life of My was longer than most of the peptides, solutions

contained a much larger mass of My due to its large molecular weight, and so mixtures were therefore more likely to provide nutrition to microbes.

Although microbial growth made definitive detection of fibrils more challenging in the present template-adder systems, microbes can coexist without issue in fiber-producing systems. Microbes are used in recombinant bacterial expression systems to produce amyloid-forming proteins in vivo. In related preliminary research, we have observed fiber formation from proteins expressed in recombinant bacteria. Bacteria were seen to attach to these fibers without disruption of structure (in preparation). Though bacteria are integral to such expressed protein systems, the fibers that can form are very large, dwarfing the microbes and their nanostructures. Fiber formation is hence distinguishable by size of the structure. Microbes were not observed to interfere with fiber formation, either imaged in SEM and AFM, or in FTIR.

2.2.3.2 SEM

We did not observe a hierarchical self-assembly in our two component protein mixtures as we did with hydrolyzed gliadin (multi-protein) template and pure protein adder systems. However, intermolecular interactions did occur in all systems, albeit different ones depending on protein amino acid composition. Thus, solvent loss seemed important in final, large-scale aggregation. The drying process involves diffusion of solvent through inner fiber layers toward the surface, where it evaporates. As solvent is excluded from the center of the structure, tighter packing of interior hydrophobic groups can occur at the same time that all hydrogen bond donor and acceptors groups have an enthalpic impetus to find a suitable acceptor or donor in the protein material remaining. After the liquid is gone, polymer rearrangement in air is minimal, but during evaporation, recruitment of free template and adder myoglobin to the elongating polymer is

favored because it provides beneficial enthalpic linkages to replace any lost interactions with water. Entropic barriers to aggregation depend on the specific sequences and their structural preferences.

The present work was motivated by initial observations of micron-sized fiber formation in tryptic hydrolysates of wheat gluten.[59] Hydrolyzed WG peptides consistently displayed round, micron-sized fibers with a high strand density β-sheet secondary structure. This was interesting because other studies on β-sheet formation in proteins and peptides displayed only nanometer-sized fibril formation. Initially, it was hypothesized that a component of the hydrolyzed protein mixture was responsible for fiber formation because unhydrolyzed WG at pH 8 and 37^{o}C would not spontaneously form β-sheet fibrils or fibers. However, WG has been shown to form fibrils under severe denaturing conditions, consistent with many other studies on amyloid fibril formation in the literature.[60–62] The current understanding of amyloid fibril formation is that fibrils form through main chain interactions that are independent of the amino acid sequence.[63] Accumulating evidence suggests that virtually any protein can fibrillate when provided the right conditions.[64,65] These are usually extreme denaturing conditions of low pH, high temperature, high stress, or high ionic strength that disrupt the protein's ability to fold based on side chain interactions, straightening the protein and allowing main chain interactions to dominate. Therefore, the same trypsin hydrolysis experiments were carried out on the gliadin and glutenin components of WG. Gliadin displayed strong β-sheet formation as evidenced by XRD and FT-IR while glutenin did not.[59] Hydrolyzed Gd also formed micron-sized fibers that were smaller in diameter and length than WG fibers. Examination of tryptic fragments of WG revealed the 3-22 peptide of gliadin (Gd20), the 162-207 peptide of gliadin (Gd46), and the 203-277 peptide of the

low molecular weight glutenin subunit (GtL75) to have a high propensity for β-sheet formation as determined by multiple prediction algorithms.[66] In hydrolyzed WG, the mole fraction of Gd:GtL was 0.51:0.49. All 3 proteins had high α-helix content but Gd20 was considered hydrophobic (positive GRAVY) compared to Gd46 and GtL75, which were considered hydrophilic (negative GRAVY, Table 2.2.2). α to β measured by FTIR spectroscopy were observed in hydrolyzed WG solutions that formed fibers.[23,59]

Table 2.2.2 Properties of peptides and proteins.

Protein	α(%)	# aa	MW	AI	GRAVY	pI	- (%)	+ (%)	TANGO
Gd20	85	20	2060	171.0	1.820	9.4	0.0	5.0	1477
Gd46	87	46	5422	108.0	-0.328	6.4	4.3	2.2	207
GtL75	47	75	8465	89.7	-0.668	5.4	2.7	1.3	771
My	76	154	17083	88.8	-0.381	7.2	13.6	13.6	624
Am	25	512	58549	69.6	-0.607	6.3	12.1	10.6	N/A

α is % α-helix; # aa is number of amino acids; MW is molecular weight in g/mol; AI is aliphatic index; GRAVY is grand average of hydropathicity; pI is isoelectric point; -, + is percentage of negative and positive amino acids, respectively; TANGO is a measure of β-sheet potential.

Replacing glutenin with My or amylase (Am) affected the α to β transition, hydrophobic packing absorbances $\delta_s(CH_3)$ and $\delta_{as}(CH_3)$, and fiber morphology. The observed changes offered insight into the multiple protein self-assembly process and how large fibers formed. Self-assembly based on complementary charge was eliminated as a possible mechanism because of the low level of charge at pH 8, and it has been shown that proteins require a large percentage of charged amino acids to self-assemble based on this mechanism.[67,68] In addition, not all protein mixtures displayed a charge mismatch. β-sheet formation was approximately equal to α-helix loss in the protein secondary structure indicating an α to β transition was important in forming

elementary units prior to fiber formation. Indeed, α-helical peptides have been shown to undergo α to β transitions more readily on hydrophobic surfaces.[69,70] Pure My had no conformation change and displayed no aggregation on any scale at the experimental conditions although it can at more severe denaturing conditions. From this previous work, Gd20KK was hypothesized as the primary template and two lysines were added to increase solubility. My was maintained as a strong adder protein that resisted conformation change at pH 8 and 37°C unless it was in the presence of a template. Here, the effect of template hydrophobicity and glutamine content is studied to complement the previous work on changing the adder protein properties.

In this previous work, correlation between adder protein glutamine content and twist in micron-sized fibers was also observed. For instance, hydrolyzed Gt proteins contained large segments of Q-blocks. Hydrolyzed Gt would not undergo conformation change or self-assemble into large fibers by itself.[59] But WG (or a combination of hydrolyzed Gd, GtL, and GtH proteins) would undergo large conformation change and self-assemble into perfectly cylindrical fibers. Replacing the Gt components with My again showed conformation change. However, My contained no Q-blocks and the resulting micron-sized self-assembled structure had a rectangular cross-sectional area referred to as a "tape". It was shown that twisting of the self-assembling structure was achieved through enhanced Q-Q bonding in Gt proteins to produce a cylinder. In the absence of Q-Q bonding, Gd:My systems self-assembled solely through hydrophobic interactions at scales higher than the hydrogen-bonded β-sheet to produce the untwisted, flat tape. In the present work, CB4 produced tapes that curled at the edges (Figure 2.2.4a) and

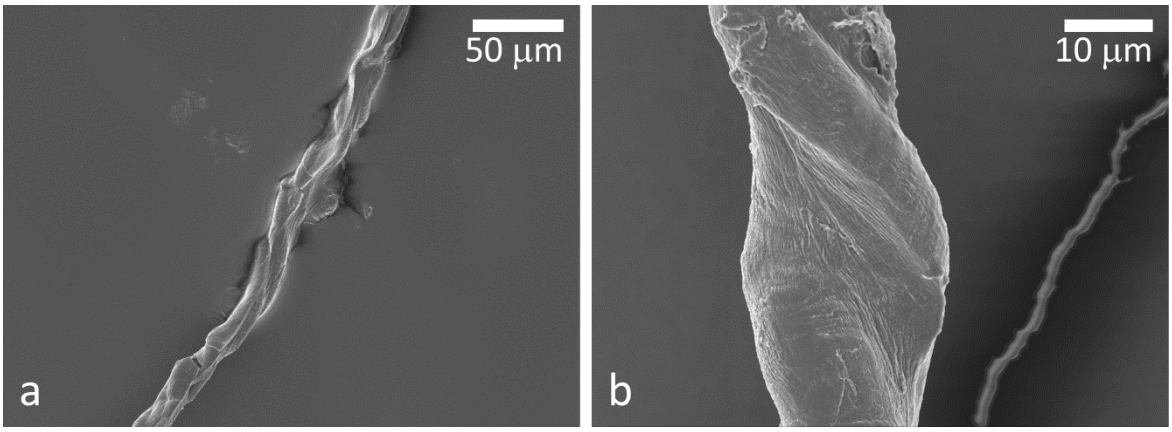

Figure 2.2.4. SEM images of CB4 at 500x (a) and CB4:My at 1000x (b). These mixtures formed irregular, twisted fibrils upon drying.

CB4:My produced flat tapes with some twist (Figure 2.2.4b). CB4 had 13% Q content but it did not seem to be enough to twist the structure but simply curl it. Myoglobin contained 4% Q without any Q-blocks. The twist in the CB4:My tapes may have been due to Q-Q bonding between the peptide and protein. An irregular CB4 and CB4:My fiber structure in SEM most likely resulted from the irregular distribution of Q and lack of repeats in the template or adder, which limited the Q-Q hydrogen bonding and close packing that was possible with larger Q blocks. Twisting was obviously still necessary for fiber stability, though fiber twisting appears to be both more regular and of a higher pitch in the mixture with My (Figure 2.2.4b). In one-way

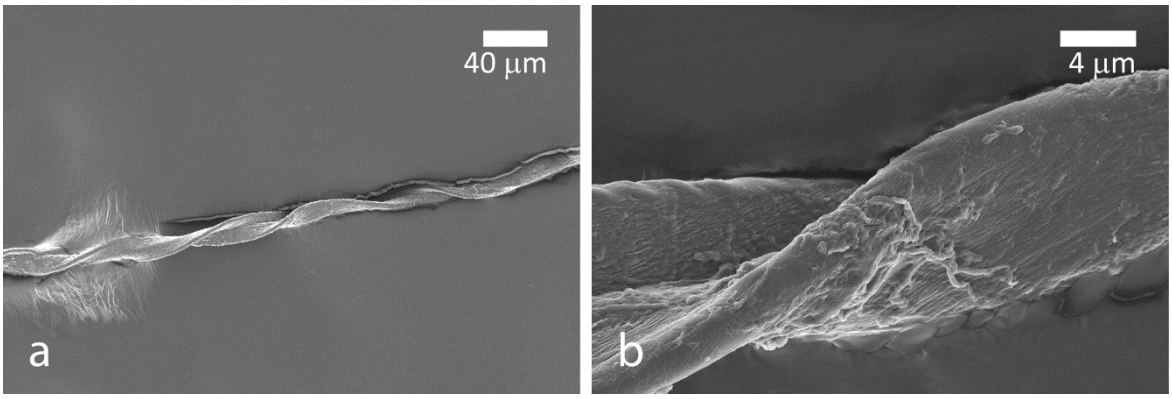

Figure 2.2.5. SEM images of Gd20KK:My at 1000x (a) and at 10,000x (b). These mixtures formed twisted tapes on drying.

analysis of variance of the measured dimensions, differences in pitch between all protein solutions studied were found to be significant (p=0.004), which provides additional evidence of the importance of the twisting morphology to Q-containing proteins (Table C.1).

Gd20KK formed tapes only in the presence of My (Figure 2.2.5). This peptide lacked Q and was reliant upon the action of hydrophobicity in the cooperative mechanism to self-assemble into larger structures. Gd20KK lacked both strong amino acid side chain hydrogen bonding and a regular repeating structure and could not form highly organized aggregates alone (Figure 2.2.5). This was also seen in a lack of Amide I development (Figure 2.1.1c).

P7 peptides with Q blocks 7 amino acids long self-assembled into an extremely twisted tape (Figure 2.2.6). In an amyloid tape, the flat part of the tape consists of protein strands

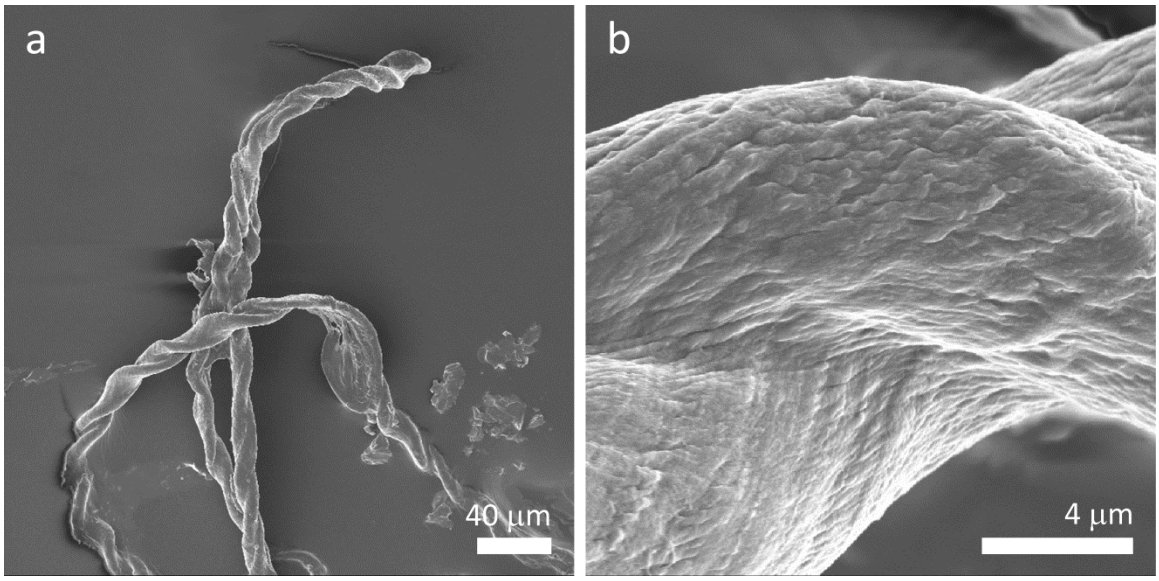

Figure 2.2.6. SEM images of P7 tapes at 762x (a) and 15,430x (b). This template showed a high degree of twisting due to the influence of Q-Q hydrogen bonding. The extremely coiled structure, long length, and high modulus (Figure 2.2.9) support the conclusion of a high degree of intermolecular interaction.

connected by hairpin turns or assembled as stacked monomers. The glutamine regime of the template protein may correspond directly to one strand in a β hairpin-like motif constituting the

tape.[71] Since hydrophobic residues are also important to fiber assembly, the configuration of synthetic peptides in a plausible tape would be Q block side chains facing solvent and hydrophobic amino acids on the interior, as in the model proposed by Aggeli *et al*.[43,72,73] In the P4 and P4-AN templates, each peptide would fold into segments of four, while in the P7 template, each peptide would fold into two segments of 7 Q residues and one segment of 6 hydrophobic residues. Interestingly, tape thickness followed the same pattern only for the My mixtures. That indicated that the number of layers of assembled β sheet depended on the presence of the adder protein to stack and overcome twisting. However, a smaller thickness was associated with a higher Q content, so hydrogen bonding was important to providing a counterbalance to torsion forces. P7:My fibers also had some degree of twist, but were amorphous, showing less core packing in their FTIR spectra and in the elastic modulus. A multiple comparisons test between mean pitch of synthetic peptides and mean pitch of My-

Table 2.2.3. Summary of morphological properties measured from FESEM pictures. Dimensions are reported as measured previously.[26] No Gd20KK fibers formed, so no measurements were reported.

Mixture	Fiber Thickness (μm)	Fiber Width W (μm)	Fiber Pitch h'(μm)	Fibril diam d (nm)	Fibril pitch h_f (nm)	Fibril angle γ (degrees)
Gd20KK:My	10.1 ± 2.9	18.7 ± 4.6	52.9 ± 16.7	479.9 ± 121.0	578.1 ± 198.5	16.1 ± 2.6
CB4	16.9 ± 1.3	16.9 ± 1.3	44.4 ± 44.4	272.3 ± 40.6	364.4 ± 51.1	13.2 ± 2.0
CB4:My	11.9 ± 7.5	17.3 ± 2.1	50.1 ± 7.4	331.2 ± 54.2	448.2 ± 98.1	28.9 ± 19.8
P4	7.8 ± 0.4	31.1 ± 12.4	132.7 ± 39.8	544.5 ± 305.5	1247.5 ± 552.5	11.6 ± 1.7
P4:My	8.4 ± 0.3	18.0 ± 3.0	178.5 ± 5.5	328.3 ± 81.8	1618.3 ± 431.7	16.4 ± 2.2
P7	5.9 ± 0.8	15.4 ± 2.0	112.1 ± 52.0	603.3 ± 136.8	833.3 ± 95.6	21.2 ± 5.5
P7:My	8.0 ± 4.8	19.4 ± 5.6	57.2 ± 21.5	468.9 ± 91.8	1771.7 ± 776.9	16.3 ± 4.1
P4AN	6.2 ± 1.6	12.1 ± 2.2	78.3 ± 32.3	372.5 ± 128.6	439.2 ± 126.3	14.4 ± 0.9
P4AN:My	13.1 ± 5.0	31.5 ± 12.6	92.2 ± 27.9	195.8 ± 62.8	235.7 ± 90.4	22.9 ± 6.6

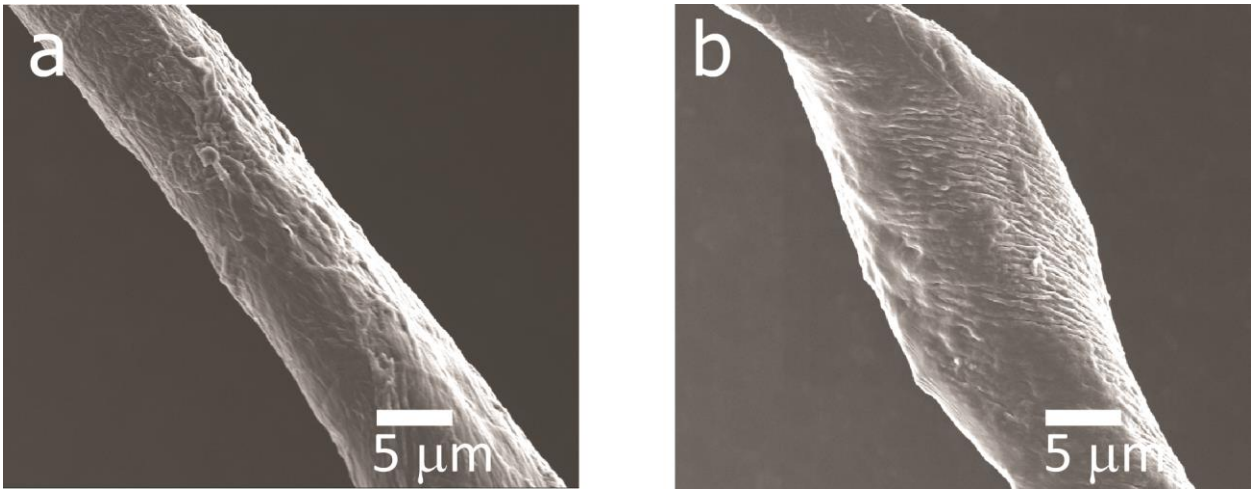

Figure 2.2.7. SEM images of P4:My at 7,570x (a) and P7:My at 6,800x (b). These mixtures formed fibers of similar diameter, but that differed in twist due to higher Q content of the P7 template.

containing peptides showed that the P7 template had equal twist to its mixture with P7:My and the P4:My mixture, but it differed from the P4 template, which had a larger pitch (Table C.2).

P4 peptides with Q blocks that were 4 amino acids long produced irregular tapes curled at the edges. P4:My morphology was similar to P7:My in dimensions and elastic modulus, but with less twist and a much longer pitch, which would be consistent with smaller Q-blocks

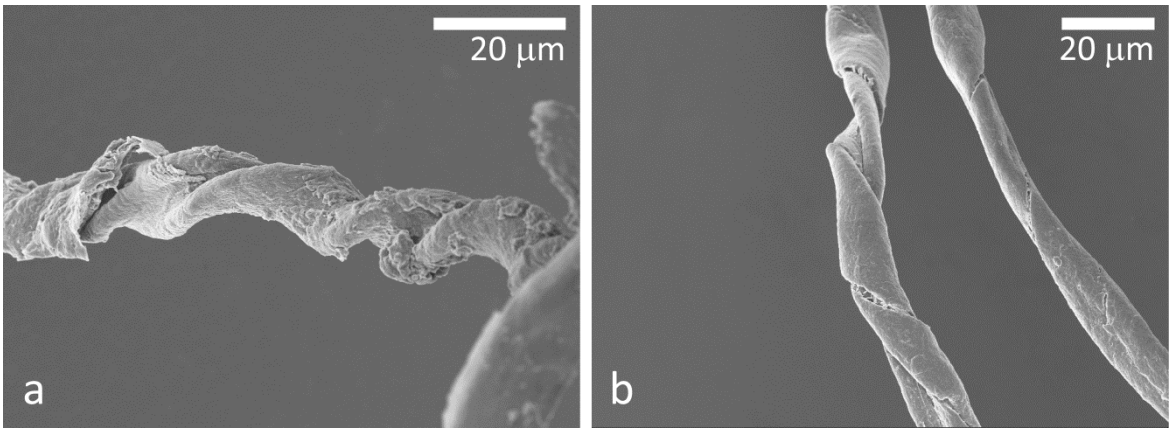

Figure 2.2.8. SEM images of P4-AN at 3,120x (a) and P4-AN:My at 2,120x (b). These mixtures showed high twisting. However, the templated mixture showed a closing of the tape due to development of complementary bonding between charged side chains localized to the interior face of the tape.

(Table 2.2.3).[74] Twisting could also occur from electrostatic repulsion, as was found in experiments on β-lactoglobulin fibers at low and high ionic strength.[75] It has been shown that Q-Q interactions in Q-blocks dominate over electrostatic repulsion to produce twist.[26] Q-free P4-AN had a neutral pI of ~7, giving it significantly less negative charge, which was expected to lead to less repulsion and less twisting. However, P4-AN had a twisted morphology regardless of no Q content, perhaps due to its amphiphilic sequence and water stability. Q-free P4-AN also assembled cooperatively with My due to high hydrophobicity combined with amphiphilicity, which lent it a degree of solvent stability and charge complementarity with the My sequence.

2.2.3.3 Nanoindentation

Pure peptides containing Q blocks had high elastic modulus, which was most likely a result of their crystallinity (Figure 2.2.9). The magnitude of the modulus was much higher than that of

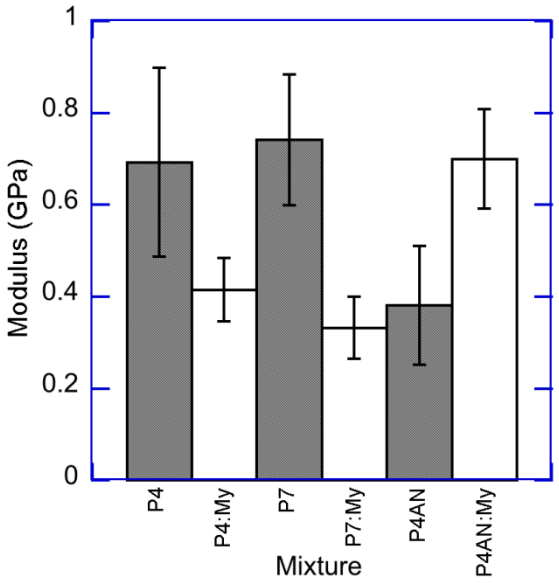

Figure 2.2.9. Elastic moduli of fibers of pure peptides and mixtures. This shows that, in the case of P4 and P7, the pure peptide was more rigid than the fibers from the mixture with My, almost by a factor of 2. However, for P4-AN:My, the mixed fibers were more rigid than the pure peptide. Although crystallinity due to a regular structure contributed to the high modulus of pure P4 and P7, the modulus of P4AN:My was due to more electrostatic interactions between the proteins as seen in FTIR and fiber morphology.

some crystalline amyloid fibrils and that of protofibrils observed during assembly progress, but one to two orders of magnitude smaller than predicted values for fully assembled amyloid fibers.[76,77] In P4:My and P7:My, the modulus was lower, correlating to diminished developments modulus may be a result of amyloid stabilization due to specific side chain groups, as was found during incubation. Although P4 and P7 were very effective at catalyzing the α to β transition in My (Figure 2.1.2), without hydrophobic packing, amyloid fibrils could not form the intermolecular interactions required for rigid mechanical properties. The magnitude of the modulus for Q-containing peptide mixtures was lower than that of silks, but comparable to actin, α-lactalbumin, αB-crystallin, some low-density plastics, and bovine hoof composites.[46,78,79] P4-AN:My mixtures, on the other hand, showed a gain in function reflected in an almost doubled modulus, similar to crystalline P4 and P7. Cucumber collagen had a similar modulus.[80] A high in other amyloid formers.[46] In this case, a higher modulus results from enhanced ionic interactions between the proteins.

2.2.4 Conclusions

Similarities were observed in AFM and SEM regarding the extent of self-assembly, especially in crystalline P7 but also in the tapes that formed for all myoglobin mixtures and for some pure peptides. This was observed despite morphology and dimension differences in fibers in AFM and SEM. Those discrepancies were thought to be due to the thermodynamic acceleration that the drying process provided to fiber assembly. In P7, the cast solution and the dry fiber had in common a very long length and high degree of torsion. This appeared to be a direct result of high Q content, but was most likely also influenced by hydrophobic solvent exclusion. However, there was a balance that was observed between the two types of side chain

interactions. Crystallization tendency was higher for a larger block of glutamine repeats. The structures formed by mixtures of myoglobin with glutamine-containing templates P4 and P7 had smaller size and weaker intermolecular bonding compared to the structures from pure templates. Myoglobin most likely provided many more amorphous regions in the fiber, which gave it increased elasticity. It is also thought that template-template bonding likely had an excluding effect on other proteins, which would prevent cooperative integration of myoglobin in the assembling fiber. Templating was much more important in P4-AN:My fibers, which assembled cooperatively to show a surprisingly high elastic modulus. Nanoindentation results confirmed that Q blocks were not necessary for strong intermolecular interactions and cooperative fibril formation.

Hydrophobic interactions were important for fiber growth via a cooperative mechanism involving short peptides and My at near-physiological conditions. Hydrophobic packing increased for most systems as aggregation progressed, but it did not always indicate fibrillar structure in solution, though it was always associated with fibrillar structure upon drying.

The hydrophobic nature of myoglobin's heme pocket makes it an ideal target for the highly hydrophobic template peptides P4, P4-AN, CB4, and Gd20KK. An accession point for templates is present in a hydrophilic edge in the heme pocket. Propionic groups from glutamic acid are present at this locus, which makes it attractive to salt-bridging template side chains such as in K.[81] In this way, template-My interactions may replace native interactions of two critical My propionic acid groups, allowing protein insertion, disruption of folding, and cooperative assembly as an amyloid β structure. Highly conserved hydrophobic residues were identified as possible nucleation points of My folding,[82] so disruption of their structure could be responsible

for whole-protein disruption and refolding. Many proteins share a similar structure, in the globin family and in others, so this self-assembly process may be applied for a number of two-protein mixtures with the potential to create an array of self-assembling fibers of differing geometries and strengths.

References

1. Gunzler, H. & Gremlich, H.-U. *IR Spectroscopy*. (Wiley-VCH, 2002).
2. Kubelka, J. & Keiderling, T. A. Differentiation of β-Sheet-Forming Structures: Ab Initio-Based Simulations of IR Absorption and Vibrational CD for Model Peptide and Protein β-Sheets. *J. Am. Chem. Soc.* **123,** 12048–12058 (2001).
3. Byler, D. M. & Susi, H. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers* **25,** 469–487 (1986).
4. Barth, A. & Zscherp, C. What Vibrations Tell About Proteins. *Quarterly Reviews of Biophysics* **35,** 369–430 (2002).
5. Surewicz, W. K., Mantsch, H. H. & Chapman, D. Determination of protein secondary structure by Fourier transform infrared spectroscopy: A critical assessment. *Biochemistry* **32,** 389–394 (1993).
6. Barth, A. Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1767,** 1073–1101 (2007).
7. Nevskaya, N. A. & Chirgadze, Y. N. Infrared spectra and resonance interactions of amide-I and II vibrations of α-helix. *Biopolymers* **15,** 637–648 (1976).
8. Arrondo, J. L. R., Muga, A., Castresana, J. & Goñi, F. M. Quantitative studies of the structure of proteins in solution by fourier-transform infrared spectroscopy. *Progress in Biophysics and Molecular Biology* **59,** 23–56 (1993).
9. Parker, F. S. *Applications of infrared spectroscopy in biochemistry, biology, and medicine*. (Plenum Press, 1971).
10. Zandomeneghi, G., Krebs, M. R. H., McCammon, M. G. & Fandrich, M. FTIR reveals structural differences between native beta sheet proteins and amyloid fibrils. *Protein Sci* **13,**

3314–3321 (2004).

11. Gunzler, H. & Gremlich, H.-U. in *IR Spectroscopy: An Introduction* 171–278 (Wiley-VCH, 2002).
12. Kauzmann, W. The Physical Chemistry of Proteins. *Annual Review of Physical Chemistry* **8,** 413–438 (1957).
13. Venyaminov, S. Y. & Kalnin, N. N. Quantitative IR spectrophotometry of peptide compounds in water (H2O) solutions. I. Spectral parameters of amino acid residue absorption bands. *Biopolymers* **30,** 1243–1257 (1990).
14. Gazit, E. A possible role for π-stacking in the self-assembly of amyloid fibrils. *FASEB J* **16,** 77–83 (2002).
15. Honda, S., Kobayashi, N. & Munekata, E. Thermodynamics of a β-hairpin structure: evidence for cooperative formation of folding nucleus. *Journal of Molecular Biology* **295,** 269–278 (2000).
16. Bandekar, J. Amide modes and protein conformation. *Biochimica et Biophysica Acta* **1120,** 123–143 (1992).
17. Oberg, K. A., Ruysschaert, J.-M. & Goormaghtigh, E. The optimization of protein secondary structure determination with infrared and circular dichroism spectra. *European Journal of Biochemistry* **271,** 2937–2948 (2004).
18. Sarroukh, R. *et al.* Transformation of amyloid β(1–40) oligomers into fibrils is characterized by a major change in secondary structure. *Cell. Mol. Life Sci.* **68,** 1429–1438 (2011).
19. Dousseau, F. & Pezolet, M. Determination of the secondary structure content of proteins in aqueous solutions from their amide I and amide II infrared bands. Comparison between classical and partial least-squares methods. *Biochemistry* **29,** 8771–8779 (1990).

20. Vedantham, G., Sparks, H. G., Sane, S. U., Tzannis, S. & Przybycien, T. M. A Holistic Approach for Protein Secondary Structure Estimation from Infrared Spectra in H2O Solutions. *Analytical Biochemistry* **285,** 33–49 (2000).
21. Fraser, R. D. B. & Price, W. C. Infra-Red Dichroism and Protein Structure. *, Published online: 20 September 1952; | doi:10.1038/170490a0* **170,** 490–491 (1952).
22. Miyazawa, T., Shimanouchi, T. & Mizushima, S. Normal Vibrations of N-Methylacetamide. *The Journal of Chemical Physics* **29,** 611–616 (1958).
23. Ridgley, D. M., Ebanks, K. C. & Barone, J. R. Peptide Mixtures Can Self-Assemble into Large Amyloid Fibers of Varying Size and Morphology. *Biomacromolecules* **12,** 3770–3779 (2011).
24. Barth, A. The infrared absorption of amino acid side chains. *Progress in Biophysics and Molecular Biology* **74,** 141–173 (2000).
25. Miyazawa, T., Shimanouchi, T. & Mizushima, S. Characteristic Infrared Bands of Monosubstituted Amides. *The Journal of Chemical Physics* **24,** 408–418 (1956).
26. Ridgley, D. M., Claunch, E. C. & Barone, J. R. The effect of processing on large, self-assembled amyloid fibers. *Soft Matter* (2012). doi:10.1039/c2sm26496j
27. Sutherland, T. D., Young, J. H., Weisman, S., Hayashi, C. Y. & Merritt, D. J. Insect Silk: One Name, Many Materials. *Annual Review of Entomology* **55,** 171–188 (2010).
28. Reisdorf, W. C. & Krimm, S. Infrared Amide I' Band of the Coiled Coil. *Biochemistry* **35,** 1383–1386 (1996).
29. Heimburg, T., Schunemann, J., Weber, K. & Geisler, N. FTIR-Spectroscopy of Multistranded Coiled Coil Proteins. *Biochemistry* **38,** 12727–12734 (1999).
30. Goormaghtigh, E., Cabiaux, V. & Ruysschaert, J.-M. Secondary structure and dosage of

soluble and membrane proteins by attenuated total reflection Fourier-transform infrared spectroscopy on hydrated films. *European Journal of Biochemistry* **193,** 409–420 (1990).

31. Vecchio, G., Bossi, A., Pasta, P. & Carrea, G. Fourier-transform infrared conformational study of bovine insulin in surfactant solutions. *International Journal of Peptide and Protein Research* **48,** 113–117 (1996).
32. Weisman, S. *et al.* An Unlikely Silk: The Composite Material of Green Lacewing Cocoons. *Biomacromolecules* **9,** 3065–3069 (2008).
33. Savitzky, A. & Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **36,** 1627–1639 (1964).
34. Barlow, D. E. *et al.* Characterization of the Adhesive Plaque of the Barnacle Balanus amphitrite: Amyloid-Like Nanofibrils Are a Major Component. *Langmuir* **26,** 6549–6556 (2010).
35. Surewicz, W. K. & Mantsch, H. H. New insight into protein secondary structure from resolution-enhanced infrared spectra. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **952,** 115–130 (1988).
36. Kauppinen, J. K., Moffatt, D. J., Mantsch, H. H. & Cameron, D. G. Fourier transforms in the computation of self-deconvoluted and first-order derivative spectra of overlapped band contours. *Anal. Chem.* **53,** 1454–1457 (1981).
37. Bradley, M. in *Thermo Fisher Scientific, I.* (Thermo Fisher Scientific, Inc.) **50733,** (2007).
38. Paparcone, R., Cranford, S. & Buehler, M. J. Compressive deformation of ultralong amyloid fibrils. *Acta Mech Sin* **26,** 977–986 (2010).
39. Böcker, U. *et al.* Revealing Covariance Structures in Fourier Transform Infrared and Raman Microspectroscopy Spectra: A Study on Pork Muscle Fiber Tissue Subjected to Different

Processing Parameters. *Appl. Spectrosc.* **61,** 1032–1039 (2007).

40. Adamcik, J., Berquand, A. & Mezzenga, R. Single-step direct measurement of amyloid fibrils stiffness by peak force quantitative nanomechanical atomic force microscopy. *Applied Physics Letters* **98,** 193701–193701–3 (2011).
41. Eaton, P. & West, P. *Atomic Force Microscopy*. (Oxford University Press, 2010).
42. Ridgley, D. M. & Barone, J. R. Evolution of the Amyloid Fiber over Multiple Length Scales. *ACS Nano* **7,** 1006–1015 (2013).
43. Aggeli, A. *et al.* Hierarchical self-assembly of chiral rod-like molecules as a model for peptide β-sheet tapes, ribbons, fibrils, and fibers. *PNAS* **98,** 11857–11862 (2001).
44. Knowles, T. P. J., Smith, J. F., Devlin, G. L., Dobson, C. M. & Welland, M. E. Analysis of structural order in amyloid fibrils. *Nanotechnology* **18,** 044031 (2007).
45. Sperling, L. H. in *Introduction to Physical Polymer Science* (Wiley-Interscience, 2006).
46. Knowles, T. P. *et al.* Role of Intermolecular Forces in Defining Material Properties of Protein Nanofibrils. *Science* **318,** 1900–1903 (2007).
47. Pawley, J. B. in *Biological Low-Voltage Scanning Electron Microscopy* (Schatten, H. & Pawley, J. B.) 27–106 (Springer New York, 2008). at <http://link.springer.com.ezproxy.lib.vt.edu:8080/chapter/10.1007/978-0-387-72972-5_2>
48. Smith, J. F., Knowles, T. P. J., Dobson, C. M., MacPhee, C. E. & Welland, M. E. Characterization of the nanoscale properties of individual amyloid fibrils. *Proc Natl Acad Sci U S A* **103,** 15806–15811 (2006).
49. vandenAkker, C. C., Engel, M. F. M., Velikov, K. P., Bonn, M. & Koenderink, G. H. Morphology and Persistence Length of Amyloid Fibrils Are Correlated to Peptide Molecular Structure. *J. Am. Chem. Soc.* **133,** 18030–18033 (2011).

50. Oliver, W. C. & Pharr, G. M. An Improved Technique for Determining Hardness and Elastic Modulus Using Load and Displacement Sensing Indentation Experiments. *Journal of Materials Research* **7,** 1564–1583 (1992).
51. Del Mercato, L. L. *et al.* Amyloid-like Fibrils in Elastin-Related Polypeptides: Structural Characterization and Elastic Properties. *Biomacromolecules* **9,** 796–803 (2008).
52. Margolin, W. FtsZ and the division of prokaryotic cells and organelles. *Nat Rev Mol Cell Biol* **6,** 862–871 (2005).
53. Macnab, R. M. How Bacteria Assemble Flagella. *Annual Review of Microbiology* **57,** 77–100 (2003).
54. Chao, Y. & Zhang, T. Optimization of fixation methods for observation of bacterial cell morphology and surface ultrastructures by atomic force microscopy. *Applied Microbiology and Biotechnology* **92,** 381–392 (2011).
55. Sloboda, R. D. Flagella and Cilia: The Long and the Short of It. *Current Biology* **19,** R1084–R1087 (2009).
56. Merchant, S. S. & Helmann, J. D. in *Advances in Microbial Physiology* (Robert K. Poole) **Volume 60,** 91–210 (Academic Press, 2012).
57. Tyler, K. M. & Engman, D. M. Flagellar elongation induced by glucose limitation is preadaptive for Trypanosoma cruzi differentiation. *Cell Motility and the Cytoskeleton* **46,** 269–278 (2000).
58. Gasteiger, E., Hoogland, C., Gattiker, A. & Bairoch, A. in *The Proteomics Protocols Handbook* 571–607 (Humana Press, 2005).
59. Athamneh, A. I. & Barone, J. R. Enzyme-mediated self-assembly of highly ordered structures from disordered proteins. *Smart Materials and Structures* **18,** 104024 (2009).

60. Bernardin, J. E. & Kasarda, D. D. Hydrated protein fibrils from wheat endosperm. *Cereal Chemistry* **50,** 529–536 (1973).

61. Bernardin, J. E. & Kasarda, D. D. The microstructure of wheat protein fibrils. *Cereal Chemistry* **50,** 735–744 (1973).

62. Kasarda, D. D., Bernardin, J. E. & Thomas, R. S. Reversible aggregation of α-gliadin to fibrils. *Science* **155,** 203–205 (1967).

63. Fandrich, M. & Dobson, C. M. The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *EMBO J.* **21,** 5682–5690 (2002).

64. Dobson, C. M. Protein misfolding, evolution and disease. *Trends in Bio* **24,** 329–332 (1999).

65. Uversky, V. N. & Fink, A. L. Conformational constraints for amyloid fibrillation: the importance of being unfolded. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* **1698,** 131–153 (2004).

66. Maurer-Stroh, S. *et al.* Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nature Methods* **7,** 237–242 (2010).

67. Gazit, E. Self-assembled peptide nanostructures: the design of molecular building blocks and their technological utilization. *Chemical Society Reviews* **32,** 1263–1269 (2007).

68. Dzwolak, W. & Marszalek, P. E. Zipper-like properties of [poly(L-lysine) + poly(L-glutamic acid)] b-pleated molecular self-assembly. *Chemical Communications* (2005).

69. Sethuraman, A. & Belfort, G. Protein structural perturbation and aggregation on homogeneous surfaces. *Biophysical Journal* **88,** 1322–1333 (2005).

70. Sethuraman, A., Vedantham, G., Imoto, T., Przybycien, T. & Belfort, G. Protein unfolding at interfaces: Slow dynamics of α-helix to β-sheet transition. *Proteins: Structure, Function, and Bioinformatics* **56,** 669–678 (2004).

71. Wei, G. & Shea, J.-E. Effects of Solvent on the Structure of the Alzheimer Amyloid-[beta](25-35) Peptide. *Biophysical Journal* **91,** 1638–47 (2006).

72. Perutz, M. F. Glutamine repeats and neurodegenerative diseases: molecular aspects. *Trends in Biochemical Sciences* **24,** 58–63 (1999).

73. Sikorski, P. & Atkins, E. New Model for Crystalline Polyglutamine Assemblies and Their Connection with Amyloid Fibrils. *Biomacromolecules* **6,** 425–432 (2004).

74. Perutz, M. F., Johnson, T., Suzuki, M. & Finch, J. T. Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *PNAS* **91,** 5355–5358 (1994).

75. Bolisetty, S., Adamcik, J. & Mezzenga, R. Snapshots of fibrillation and aggregation kinetics in multistranded amyloid β-lactoglobulin fibrils. *Soft Matter* **7,** 493–499 (2011).

76. Guo, S. & Akhremitchev, B. B. Investigation of Mechanical Properties of Insulin Crystals by Atomic Force Microscopy. *Langmuir* **24,** 880–887 (2008).

77. Xu, Z., Paparcone, R. & Buehler, M. J. Alzheimer's Aβ(1-40) Amyloid Fibrils Feature Size-Dependent Mechanical Properties. *Biophysical Journal* **98,** 2053–2062 (2010).

78. Fried, J. in *Polymer Science & Technology* 194 (Prentice Hall, 2003).

79. Tombolato, L., Novitskaya, E. E., Chen, P.-Y., Sheppard, F. A. & McKittrick, J. Microstructure, elastic properties and deformation mechanisms of horn keratin. *Acta Biomaterialia* **6,** 319–330 (2010).

80. Heim, A. J., Matthews, W. G. & Koob, T. J. Determination of the elastic modulus of native collagen fibrils via radial indentation. *Applied Physics Letters* **89,** 181902–181902–3 (2006).

81. Dickerson, R. E. & Geis, I. *The structure and action of proteins*. (Harper & Row, 1969).

82. Dyson, H. J. & Wright, P. E. Unfolded Proteins and Protein Folding Studied by NMR. *Chem. Rev.* **104,** 3607–3622 (2004).

CHAPTER 3: Conclusions

The objectives of the present research were completed. It was determined that the major template property affecting cooperative assembly of mixed protein systems was the hydrophobicity and specific patterning of the template. The present research may be used in further work to explore the factors governing recruitment of the large adder protein to amyloid structures. Many economically competitive bulk proteins are available for this work, including the trypsin-hydrolyzed fragments of wheat gluten. In addition, ionic interactions hypothesized to control Q-free template interactions may be identified in more detail by varying solution pH during assembly.

3.1 Major Impact of Hydrophobicity, Patterning

The presence of hydrophobic residues in the template sequence was important for controlled hierarchical fiber growth via a cooperative mechanism involving template peptides and My at near-physiological conditions. Template sequence affected changes in hydrophobic packing the most in the My mixtures. Hydrophobic packing changes increased for most systems as aggregation progressed, but it did not always indicate fibrillar structure in solution. However, hydrophobic packing was always associated with fibrillar structure upon drying. The least hydrophobic P7 had the smallest hydrophobic packing increase, while all others showed that symmetric vibrations are favored over asymmetric. That meant that hydrophobic packing was a good indicator of aggregate core packing in proteins that were less solution-stable.

In the pure synthetic templates, hydrophobicity caused unfolding in My at rates which correlated with the quantity of hydrophobic residues in the template. My unfolding was an

important prerequisite for cooperative assembly. However, in Gd20KK and CB4 mixtures with randomly patterned templates, My was not well recruited to the aggregate, which was shown in the lack of strong $\alpha \rightarrow \beta$ transitions. The importance of hydrophobic patterning was shown in P4-AN:My, which showed a distinct hydrophobic packing decrease after 200 hours but strong fiber morphology. In P4-AN, the hydrophilic portions of the pattern were varied in size and charge, so hydrophobic packing favored asymmetric packing, perhaps because of a larger inter-β sheet distance in water.

Similarities were observed in AFM and SEM regarding the extent of self-assembly, especially in crystalline P7 but also in the tapes that formed for all My mixtures and for some pure peptides. This was observed despite morphology and dimension differences in fibers in AFM and SEM due to the thermodynamic advantage that the drying process provided to assembly. In P7, the cast solution and the dry fiber had in common a very long length and high degree of torsion. This appeared to be a direct result of high Q content, but was likely influenced by hydrophobic solvent exclusion. However, a balance that was observed between the two types of side chain interactions. Crystallization tendency was higher for a larger block of glutamine repeats. The structures formed by glutamine-containing proteins P4 and P7 lacked the size and material properties found in the pure proteins, so template-template bonds likely excluded cooperative integration of My. Templating was more important in P4-AN:My fibers, which formed fibers with a surprisingly high elastic modulus. Nanoindentation results confirmed that Q blocks were not necessary for strong intermolecular interactions and cooperative fibril formation.

The hydrophobic nature of the heme pocket of myoglobin makes it an ideal target for the highly hydrophobic template peptides P4, P4-AN, CB4, and Gd20KK. An accession point for

templates is present in a hydrophilic edge in the otherwise hydrophobic heme pocket of My.[1] In this way, template-My interactions may replace native interactions of two critical My propionic acid groups, allowing protein insertion, disruption of folding, and cooperative assembly as an amyloid β structure. Highly conserved hydrophobic residues were identified as possible nucleation points of My folding,[2] so disruption of their structure could be responsible for whole-protein disruption and refolding. Many proteins share a similar structure, in the globin family and in others, so this self-assembly process may be applied for a number of two-protein mixtures with the potential to create an array of self-assembling fibers of differing geometries and strengths.

Myoglobin lent unstable templates Gd20KK, P4, and P4-AN a degree of water-stability that led to self-assembly via cooperative binding. In the absence of myoglobin, this transition occurred only for templates that had a high potential for Q-Q hydrogen bonds. Some mixtures with templates showed evidence for unaggregated My, shown in a residual FTIR peak at 1650 cm^{-1}. The presence of the large protein My assembled with templates caused torsion in final dried structures, shown in a new C_α-C stretch in the dried aggregate.

3.2 Influence of Glutamine

The highest Q-containing templates (P7) had very distinct β sheets which split into low and high strand density absorbances in FTIR spectra. Higher strand density absorbances were coincident with P7 crystallization from strong Q-Q hydrogen bonding, which caused these peaks to shift to lower wavenumber. When myoglobin was mixed with the P7 template, little α to β transition occurred, so the high glutamine peptide was ineffective as a template, since it preferred self-catalysis. A low level of glutamine in some unordered templates was insufficient for

amyloid development. Peptides with low levels of Q formed β sheets, but this is common among many proteins in water. At a higher level of Q, templates were still unable to form amyloids, losing β sheet structure over time, presumably due to hydrogen bonding with water. This information regarding the balance needed to achieve two-protein assembly in templates with Q will be useful in designing future self-assembling protein systems with customizable material properties.

References

1. Dickerson, R. E. & Geis, I. *The structure and action of proteins*. (Harper & Row, 1969).
2. Dyson, H. J. & Wright, P. E. Unfolded Proteins and Protein Folding Studied by NMR. *Chem. Rev.* **104,** 3607–3622 (2004).

Appendix A: Supporting Information for FTIR Analysis

This appendix presents supporting information for FTIR analysis, pertaining to the Amide I indicators α/β and hydrophobic packing, the Amide II peak width, and procedures used to derive information from the FTIR spectra. The purpose of these measures was to determine extent of fibril self-assembly during protein mixture incubations. As shown in Figure A.1.1-Figure A.1.10, it is possible to add information to the interpretation of Amide I peaks by limiting the structures considered to α helix and high density β sheet, then taking the ratio of the respective peak areas (Figure A.1.11). However, qualitative discussions of IR spectra were determined to be sufficient, so the graph of the ratio of α/β structures vs. time is included in the appendix rather than in the text. The deconvolution method used to determine the Amide I spectral information in A.1 is presented in A.2. Changes in the Amide II peak full width at half height (FWHH) were used in Section 2.1.3.2 as an indicator of self-assembly progress, but the peak widths for a number of mixtures were not reported in the results since they did not show any change over time. In general, long time incubations showed a trend for My-containing mixtures, but did not show a trend for synthetic peptides with low Q, P4 and P4AN. P4AN did not show a significant Amide II peak, which indicated that the peptide may have dissolved without forming a secondary structure. Short time incubations did not show any trend for P4 and P4AN spectra, and short time incubations were not performed for P7:My mixtures since they showed development during long-time incubations. The results indicated that there was little structural transition in the short time and long time incubations of P4 and P4AN. These two structures had sufficiently stable β sheet in the liquid, which resulted in an unchanged Amide II vibration.

A.1 Amide I Deconvolution Data – Peak Areas vs. Time

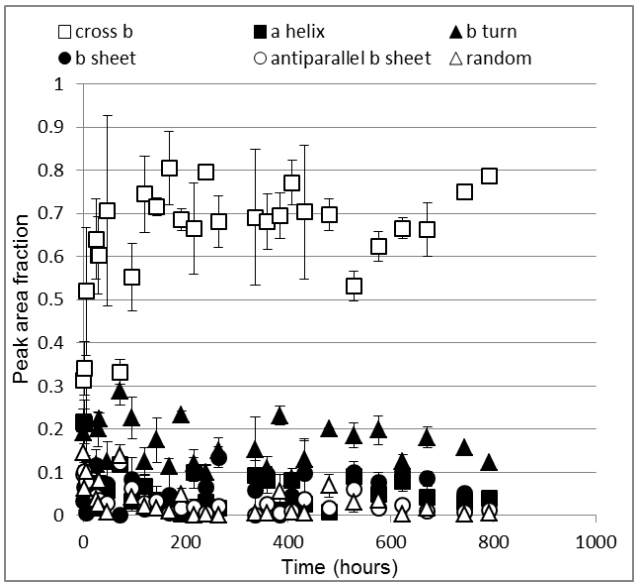

Figure A.1.1 CB4 Amide I peak areas over long-time incubation at pH 8, 37°C

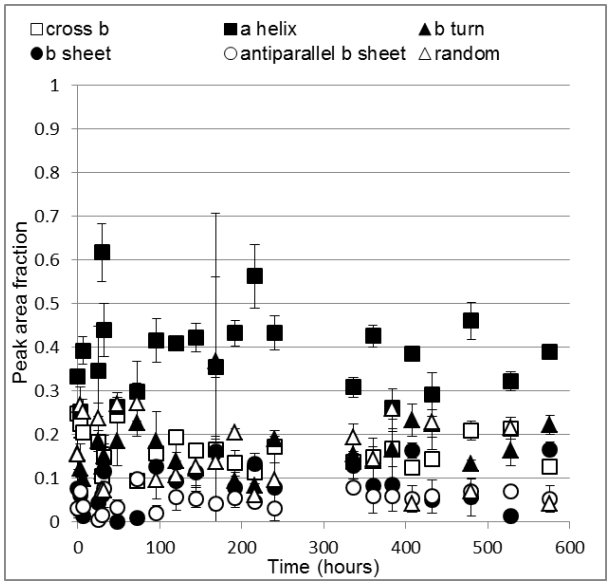

Figure A.1.2 CB4:My Amide I peak areas over long-time incubation at pH 8, 37°C

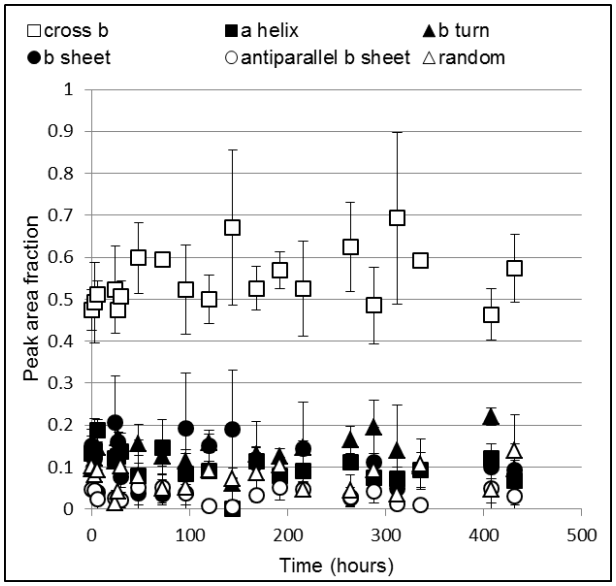

Figure A.1.3 Gd20KK Amide I peak areas over long-time incubation at pH 8, 37°C

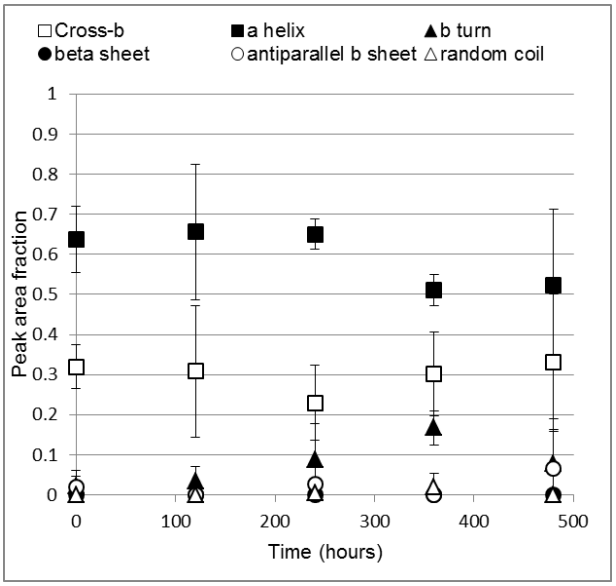

Figure A.1.4 Gd20:My Amide I peak areas over long-time incubation at pH 8, 37°C.

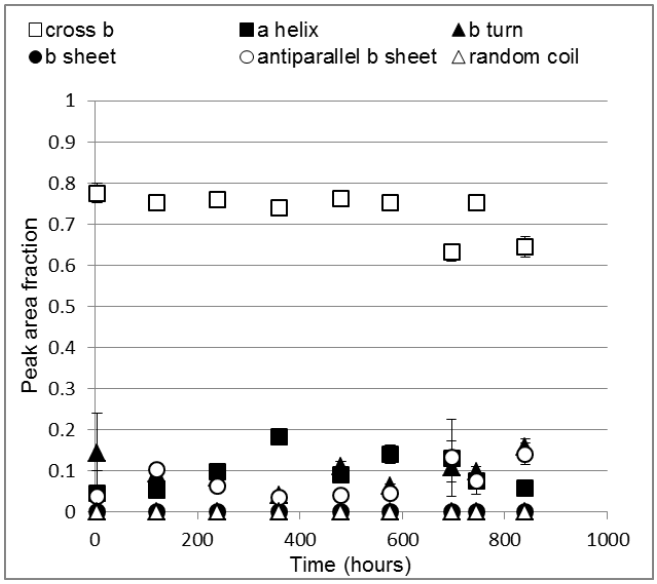

Figure A.1.5 P4 Amide I peak areas over long-time incubation at pH 8, 37°C.

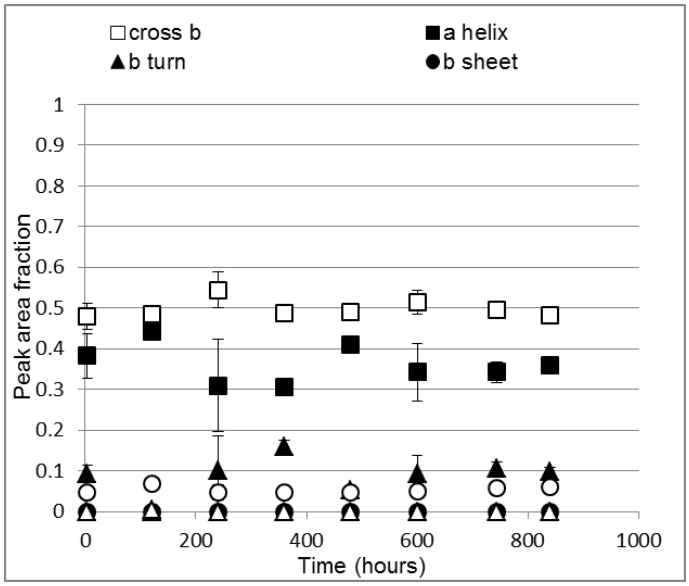

Figure A.1.6 P4-My Amide I peak areas over long-time incubation at pH 8, 37°C.

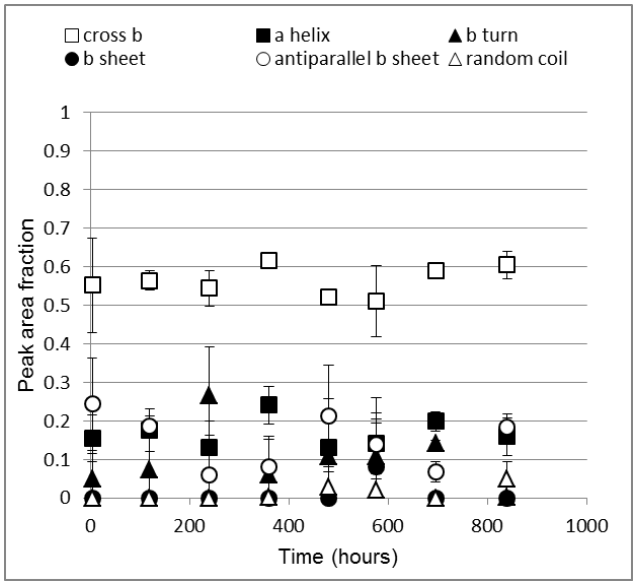

Figure A.1.7 Amide I peak areas over long-time incubation at pH 8, 37°C

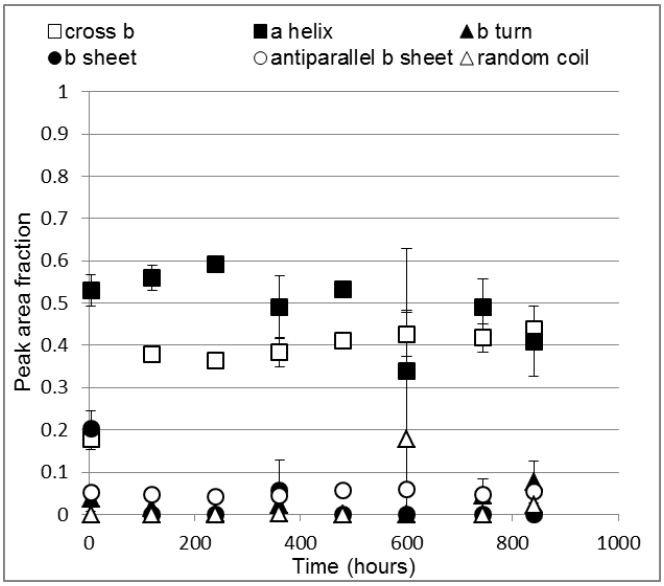

Figure A.1.8 P7-My Amide I peak areas over long-time incubation at pH 8, 37°C

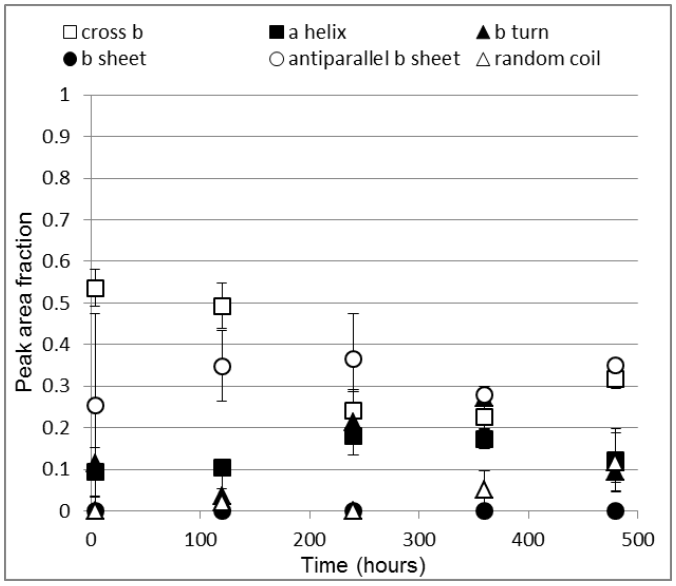

Figure A.1.9 P4AN Amide I peak areas over long-time incubation at pH 8, 37°C

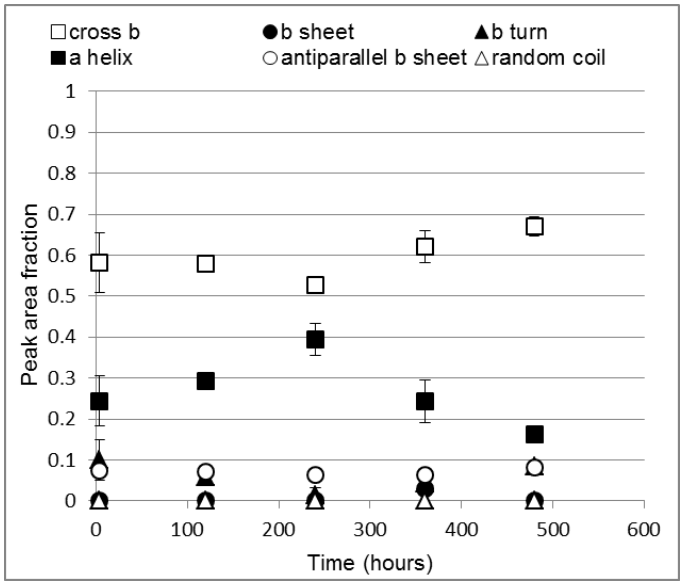

Figure A.1.10 P4AN:My Amide I peak areas over long-time incubation at pH 8, 37°C

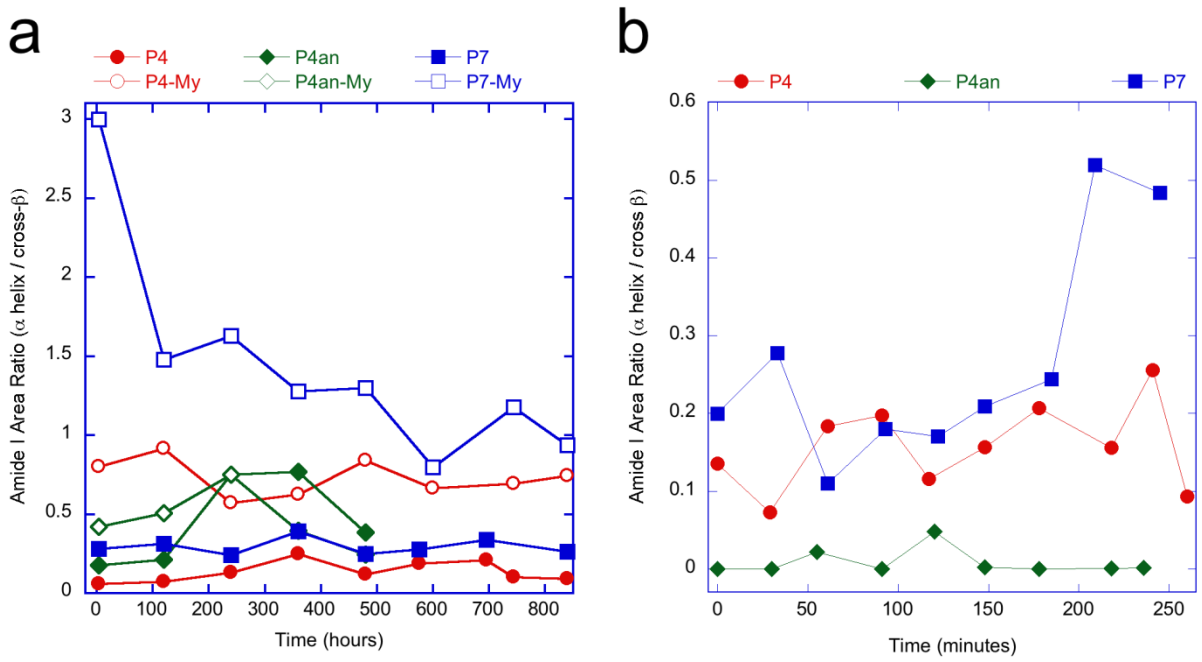

Figure A.1.11. A plot of nonzero α/β transitions over long time (a) and short time (b). Short time plots (b) are limited to species which showed a nonzero change.

A.2 Procedure for FTIR Spectral Deconvolution of Amide I Peak

1. Open the spectrum in OMNIC. Select the wavenumber range that bounds the feature of interest (eg. For amide I, look from 1770-1590 cm^{-1}).
2. Smoothing of the spectra: Smooth the spectra, keeping important features (eg. Those that indicate actual 2° structure components) and smoothing meaningless noise. This aspect requires some knowledge of the expected vibrational modes in the region of interest. Typically, the Amide I region contains six to nine components.[1]
    a. Process -> Smooth (manual)
    b. Select a number of points that satisfies the requirements. Fewer points works better to make sure important features are not smoothed out. Eg. 5-9 smooth points

3. The second derivative. Choose the important peaks of the spectra by those that have significant negative peaks in the second derivative. Count these – this will be used in Peak Analysis later.[2,3,4]
4. Fourier Self-Deconvolution. This uses Happ-Genzel apodization to smooth the spectrum. Apodization determines spectral resolution.[4] Use Enhancement values 2-3 and bandwidth that gives well-defined spectra with similar number of features as the second derivative step.
    a. Process -> Fourier Self-Deconvolution
5. Peak Analysis. Analyze the smoothed spectra to calculate area of peaks corresponding to secondary structure features. Choose constant baseline and Gaussian-Lorentzian curves. Sensitivity may be set to 'high' depending on the fit achieved. The most critical parameter to adhere to in this step is to the number of peaks. Change this using the Peak button the left-hand margin. There should be 6 or so peaks in the Amide I region, for example, based on second derivative analysis.[1] Make sure the list of peaks is trimmed to include the same # of entries as you counted in the second derivative analysis (you may add 1 peak for a 'fudge factor' to account for edge area effects, where Amide I is affected by aromatic residue absorbance or other nearby, overlapping peaks).
    a. Analyze -> Peak resolve
    b. Set baseline and peak type (Gaussian/Lorentzian)
    c. Find peaks
    d. Peaks button – Delete peaks exceeding the desired # of peaks counted. You may want to choose specific peaks to keep based on location – the fit that results may depend on initial guesses.

e. Fit peaks: repeat until a good fit is achieved with peaks at reasonable locations corresponding to secondary structure.

6. Peak Resolve Statistics button: The results will give an F statistic, standard error, points in region, number of constraints (degrees of freedom). These may be used to calculate a reduced χ-square. See USGS page 169.[5]

7. Area analysis of Secondary Structural Features:

   Feature identification should include all structural components that are expected in the region of interest. For Amide I vibrations, we expect a number of beta sheet bands, alpha helices, beta turns, random or unordered coil. Characteristic wavenumber for assignment varies from protein to protein. Average values are given for literature sources in Table A.1.

Table A.1 Vibrational frequencies for Amide I features

* molar distinction coefficients in parenthesis, ∫ standard deviation included from n=17 total proteins analyzed, [a] Calculated, [b] Theoretical, from literature, [c] Experimental, from literature

Feature	Byler and Susi 1986[1∫]	Vedantham et al. 2000[6]*	Ridgley 2011[7]
Beta chain	1624 ± 2.4 1637 ± 1.4, 1631 ± 2.5 1675 ± 2.6	1630[b], 1633[c], 1633 (0.82)[a] 1688[c] 1691 (0.13)[a], 1695[b]	1611-1630 Cross-β 1630-1637 β sheet 1679-1688 Antiparallel
Helix	1654 ± 1.5	1652 (1.00)[a], 1653[b], 1656[c]	1647-1662
Random coil	1645 ± 1.6	1620 (2.85)[a] 1638[c] 1646 (1.51)[a] 1650[c] 1662[c], 1663[b], 1665[b], 1667[c] 1676[a]	1637-1647
Turns and Bends	1663 ± 2.2 1670 ± 1.4 1683 ± 1.5 1688 ± 1.1 1694 ± 1.7	1613 (0.10)[a] 1629 (0.26)[a] 1652[b] 1666[b,c] 1668 (0.20)[a] 1672[c] 1680[c], 1686[b], 1688[c] 1690[b] 1720 (0.17)[a]	1662-1678 1689-1699

References

1. Byler, D. M. & Susi, H. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers* **25**, 469–487 (1986).
2. Maddams, W. F. The Scope and Limitations of Curve Fitting. *Appl. Spectrosc.* **34**, 245-267 (1980).
3. Susi, H. & Michael Byler, D. Protein structure by Fourier transform infrared spectroscopy: Second derivative spectra. *Biochemical and Biophysical Research Communications* **115**, 391-397 (1983).
4. Kauppinen, J. K., Moffatt, D. J., Mantsch, H. H. & Cameron, D. G. Fourier transforms in the computation of self-deconvoluted and first-order derivative spectra of overlapped band contours. *Anal. Chem.* **53**, 1454-1457 (1981).
5. Helsel, D. R. & Hirsch, R. M. A3. Statistical Methods in Water Resources. *Hydrologic Analysis and Interpretation* **4**, (2002).
6. Vedantham, G., Sparks, H. G., Sane, S. U., Tzannis, S. & Przybycien, T. M. A Holistic Approach for Protein Secondary Structure Estimation from Infrared Spectra in H2O Solutions. *Analytical Biochemistry* **285**, 33-49 (2000).
7. Ridgley, D. M., Ebanks, K. C. & Barone, J. R. Peptide Mixtures Can Self-Assemble into Large Amyloid Fibers of Varying Size and Morphology. *Biomacromolecules* **12**, 3770-3779 (2011).

A.3 Amide II Analysis Data

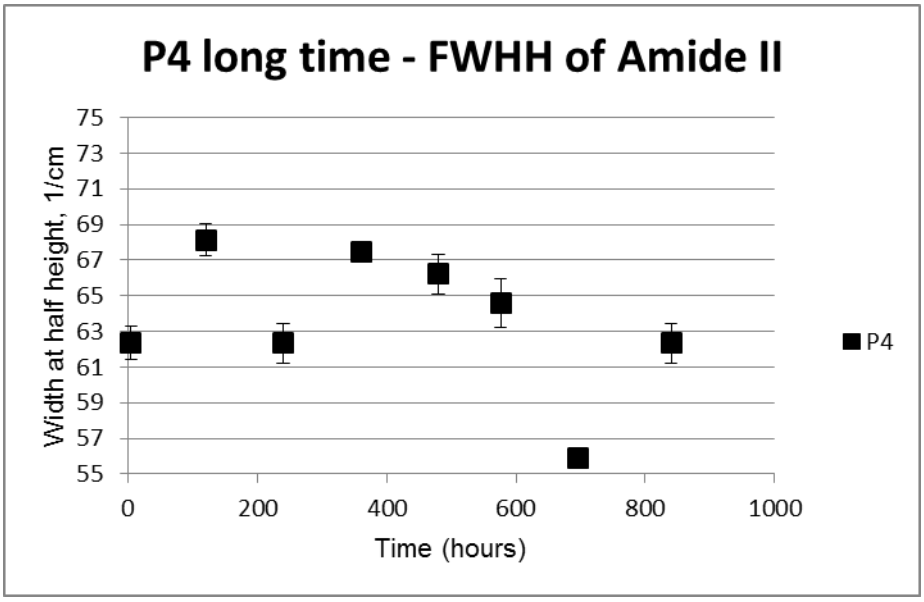

Figure A.3.1 FWHH of P4 long time incubation lacks a significant trend

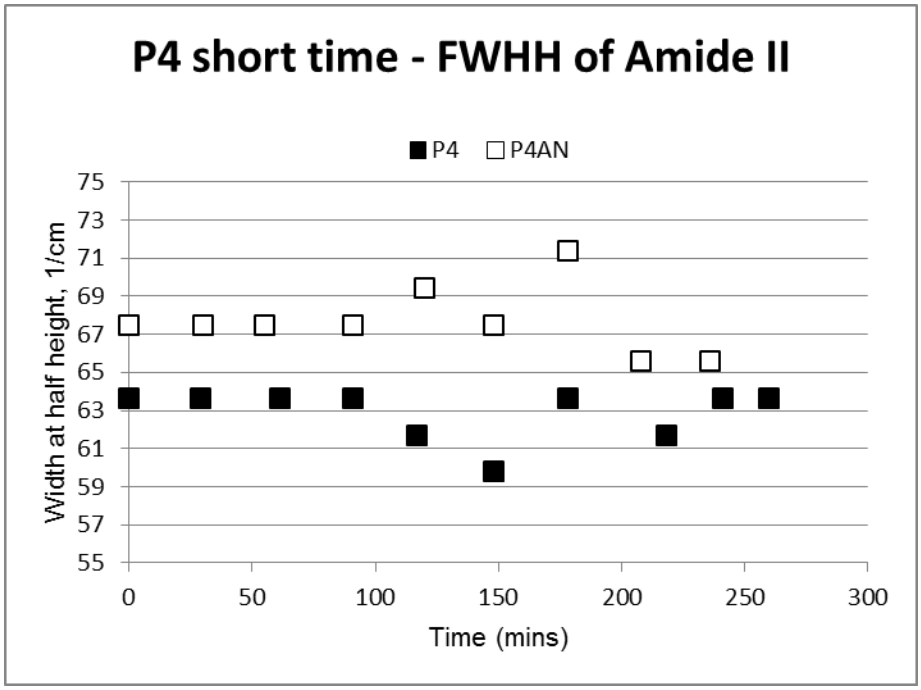

Figure A.3.2 FWHH of P4 and P4AN short time incubations lack a significant trend

A.4 Method for Amide II Analysis (Full Width at Half Height)

The full width at half height is an important parameter in spectral curve fitting that describes the bandwidth of vibrations. Mathematically, the full width at half maximum (Γ) affects the shape of Gaussian and Lorentzian curves.

From the integral of a Gaussian curve function $F_G(\nu)$, the peak area is:

$$\int F_G(\nu)d\nu = \sqrt[2]{\pi/ln2}\ A_{max}\ \Gamma/_2^1$$

The integral of a Lorentz curve function $F_L(\nu)$ gives the peak area:

$$\int F_L(\nu)d\nu = \pi\ A_{max}\ \Gamma/_2^1$$

A.4.1 Baseline

For secondary structure predictions using Amide I and Amide II regions of IR spectra, baseline subtraction and normalization is recommended in investigations of the statistical accuracy of such procedures.[2] According to Pitha and Jones,[3] a linear baseline parallel to the x-axis only requires two points well-removed from the peak area of interest. For peaks fitted to a Lorentzian profile, wide wings result in a very wide distribution of intensity, even up to five half-widths from the peak center.[4] Therefore, one can choose a linear baseline point above 1750 cm^{-1}, past the Amide I peak tails, and one baseline point between 1500 and other, smaller peaks occurring in the fingerprint region between 1000 cm^{-1} and the Amide I peak. For protein spectra this baseline may occur at an absorbance of zero if no negative peaks are present lying below the zero absorbance level.

A.4.2 Peak Measurement

In OMNIC, full width at half height may be measured using the peak height tool (see the Figure below).

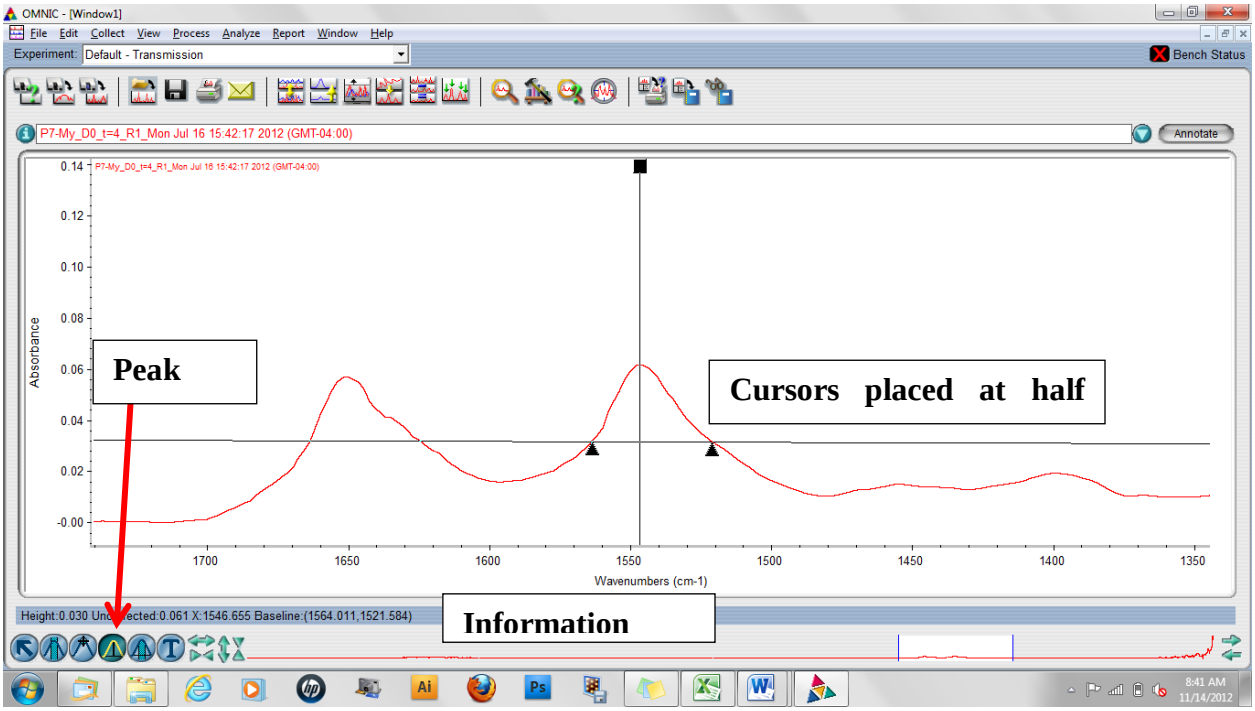

Figure A.4.1 Using the Peak Tool in OMNIC to determine Full width at Half Height

Narrow the display limits around the Amide II peak (or peak of interest) such that a relatively flat baseline may be viewed on either side – eg. 1750 to 1350 cm^{-1}. Click on the peak and move the vertical bar so that the center is in the center of the peak. The uncorrected height may be read at the bottom grey information panel. This gives the absolute value of the absorbance, from 0 to the intersection of the vertical line with the spectra. For example, if the uncorrected height is 0.10, the half-height will occur at 0.05. Position the horizontal line by moving the triangular cursors so that you have a Height equal to half your peak (0.05 in this example), which you can check in the information panel. The uncorrected height should still read from 0. Calculate the

width: the baseline points will give you the frequencies on either side of your curve at the half height. Subtract the smaller from the larger. For example, if you have Baseline (1565.0, 1518.0), the width is calculated as 1565.0 - 1518.0 = 47 cm^{-1}.

References

1. Gunzler, H. & Gremlich, H.-U. Spectrometers. *IR Spectroscopy: An Introduction* 82-84 (2002).
2. Oberg, K. A., Ruysschaert, J.-M. & Goormaghtigh, E. The optimization of protein secondary structure determination with infrared and circular dichroism spectra. *European Journal of Biochemistry* **271**, 2937–2948 (2004).
3. Pitha, J. & Jones, R. N. A Comparison of Optimization Methods for Fitting Cuves to Infrared Band Envelopes. *Canadian Journal of Chemistry* **44**, 3031-3050 (1966).
4. Maddams, W. F. The Scope and Limitations of Curve Fitting. *Appl. Spectrosc.* **34**, 245-267 (1980).

Appendix B: Supporting Information for AFM

This appendix presents information used in the analysis of AFM scans and a method used to do so. Although analyses were confounded by contaminated anisotropic structures, the same analysis may be applied if the identity of the fiber is assured prior to analysis.

Method for Measuring Fibrils using NanoScope Analysis (Bruker)

1. **Open file**: Open a topographical scan image (these are identified by a TOPO near the file extension. They also have a distinct appearance from phase amplitude or deflection scans.)
2. **Section Analysis Tool**: Select the topo image and click the section tool (icon: a box with a knife cutting it, or use the menu item Analysis>Section). The image appears in an inset with two plots. The top is a distance vs. height plot and the bottom is a frequency vs. measurement plot, which tells you information about the sections that you define.
3. **Draw the slice**: You can make up to three section measurements at once. Draw a line for each – you can differentiate between them by color. Each line you draw shows a surface section in the topmost plot.
4. **Define the slice to measure**: Move the dashed vertical boundary bars to choose where to begin and end your measurement. This will typically be at the edge of an object of interest, eg. A globule. You can check the location of the crosses (+) on the image at left (which correspond to the dashed boundary bars) to make sure the edges are in the right location.
5. **Measurements:** A list of measurements will be at the bottom in three colored rows corresponding to the line sections you draw. You can right click and choose which

measurements are tabulated here. For height of a feature, look at Rmax (distance between lowest and highest point in the measurement range). For width and length measurements, look at either Horizontal Distance or Surface Distance. Surface distance includes any rises or falls in the surface, but horizontal distance is a straight line distance.

6. **Recording data**: It is easy to record these measurements in an open Excel document separate from the NanoScope Analysis program. Although NanoScope Analysis will convert the colored row of measurements in the Section Tool into an Excel document for you, it may be better to organize it in a way that you have defined for your data. An example is shown below:

Table B.1 Example of Recorded AFM measurements

Fibril ID	Z height (nm)			Width (nm)			Length (nm)	
bright lower	131.46	164.46	129.74	1278	1293	1440	4774	
dark upper	70.49	96.16	61.71	895	903	903	4509	
cross-ways long	110.81	87.63	48.55	827	716	665	3603	4703
left vertical	82.20	96.34	82.63	1050	1072	869	9395	
right diagonal	51.39	48.73	41.62	649	623	602	6345	

Appendix C: Supporting Information for SEM

This appendix presents the results of statistical analysis of dimension measurements made using ImageJ software measurements of SEM images. One-way analysis of variance was conducted on dimension measurements by solution (P4, P7, P4:My, P7:My) in JMP using Tukey's Honestly Significant Difference for multiple comparisons. Fiber pitch was the only measurement to vary by solution ($p=0.004$, Table C.1). This confirmed the importance of twisting in differentiating solution morphologies. P4 was found significantly different than P7 and P7:My, with P4:My not significantly different than either of those groupings (See Table C.2). Two-factor analysis could not differentiate between low and high levels of glutamine content and low and high levels of hydrophobicity. In addition, a nested ANOVA analysis was performed to determine the amount of variation due to different samples. This analysis found that the different samples varied much more than the variation between treatments, which explains the lack of significant results. Adding more fiber samples to the analysis would make it much more robust.

Statistical Analysis of Measurements

Table C.1 1-way ANOVA summary for fiber dimensions to test significance of mixture on measurements. The only significant difference was observed in pitch, expressed as twisting. Twisting was a very important factor related to glutamine content.

Dimension	N	F Statistic	P-value
Fiber Width	17	0.25	0.86
Fiber Thickness	17	0.58	0.64
Fiber Pitch	20	6.73	0.004*
Fibril diameter	19	0.41	0.75
Fibril pitch	23	0.49	0.70
Angle to fiber, γ	31	0.36	0.78

Table C.2 ANOVA table for fiber pitch, the only dimension to show significance between the four mixtures tested singly. * Means followed by the same letter are not significantly different at the 0.05 level (experimentwise) using Tukey's HSD.

Mixture	N	Mean (μm)*	Standard Error (μm)
P4	5	93.0[a]	17.9
P4:My	4	178.5[ab]	17.0
P7	4	85.8[b]	21.7
P7:My	4	35.8[b]	3.5

Table C.3 Results of nested analysis. Random effects of sample on dimension measurements comprised a large portion of the variation, whereas variation from the treatment was a more minor component, except in the case of overall fiber pitch and fibril diameter.

Dimension	% Variation from Sample	% Variation from Solutions
Fiber Width	93	7
Fiber Thickness	85	15
Fiber Pitch	0	100
Fibril diameter	0	100
Fibril pitch	46	54
Angle to fiber, γ	27	73

Appendix D: Supporting Information for Nanoindentation

This information is presented to support the choice of dataset presented for nanoindentation results (2.2.3.3). A screening process was done at the time of testing to determine which of the elastic moduli calculated corresponded to an indentation that was in contact with the fibril surface or in contact with the background material or other debris. Outliers were identified through rating the hysteresis pattern of loading and unloading of the indenter on a scale of 0 to 10, with 10 representing a perfect fit and 0 representing erroneous data. Lower scores were excluded as outliers. The total number of well-fitting points that were subsequently used for modulus calculation is reported in the table below (Table D.1). Analytical screening was needed because the equipment used did not allow visual confirmation of the locus of indentation. The author gratefully acknowledges Devin Ridgley for performing these measurements and calculations.

Table D.1 Summary of well-fitting nanoindentation results and the calculated Young's moduli and standard error.

Mixture	N used for calculation	Modulus (GPa)
P4	20	0.69 ± 0.21
P4:My	13	0.41 ± 0.07
P7	18	0.74 ± 0.14
P7:My	18	0.33 ± 0.07
P4AN	13	0.38 ± 0.13
P4AN:My	13	0.70 ± 0.11

Table D.2 Nanoindentation results for P4. A subset of well-fitting points was chosen for calculation and reported.

Fit	Modulus (GPa)	Outlier (Y=Yes, N=No)
0	0.46	Y
0	0.59	Y
0	0.78	Y
0	1.33	Y
3	0.04	Y
0	0.68	Y
7	0.10	N
8	0.14	N
7	0.09	N
7	0.10	N
7	2.58	N
7	1.95	N
7	0.81	N
7	0.35	N
7	0.15	N
3	0.31	Y
7	1.16	N
7	0.29	N
7	0.45	N
5	0.08	Y
5	1.34	Y
7	0.10	N
2	0.96	Y
7	0.49	N
7	3.15	N
7	0.55	N
4	0.08	Y
3	0.03	Y
3	0.01	Y
3	0.03	Y
3	0.03	Y
0	0.04	Y
0	30.39	Y
2	0.08	Y
8	0.39	N
2	0.30	Y
2	0.03	Y
0	0.00	Y
0	4.72	Y
0	235.38	Y
0	227.09	Y
0	0.03	Y
0	0.03	Y
7	0.17	N
8	0.16	N

Table D.3 Nanoindentation results for P4:My. A subset of well-fitting points was chosen for calculation and reported.

Fit	Modulus (GPa)	Outlier? (Y=Yes, N=No)
1	1800.07	Y
1	1800.28	Y
1	103.63	Y
1	0.01	Y
1	0.01	Y
0	2.32	Y
0	3.32	Y
0	1800.13	Y
0	44.77	Y
0	13.69	Y
0	0.03	Y
6	0.07	Y
6	0.07	Y
8	0.35	N
7	0.39	N
7	0.13	N
5	0.03	Y
1	0.03	Y
9	0.18	N
1	0.92	Y
7	0.22	N
7	2.49	N
8	0.68	N
6	0.15	Y
6	0.08	Y
7	0.46	N
6	0.09	Y
7	0.13	N
1	0.12	Y
7	0.59	N
7	0.53	N
2	0.04	Y
1	0.04	Y
6	0.05	Y
1	0.02	Y
7	0.51	N
6	0.09	Y
7	0.45	N
7	1.06	N
6	0.66	Y

Table D.4 Nanoindentation results for P7. A subset of well-fitting points was chosen for calculation and reported.

Fit	Modulus (GPa)	Outlier? (Y=Yes, N=No)
0	22.47	Y
0	1.41	Y
1	0.03	Y
1	0.01	Y
1	0.01	Y
1	0.03	Y
7	1.50	N
7	0.83	N
8	1.90	N
8	0.22	N
7	1.82	N
7	0.15	N
7	0.08	N
1	0.02	Y
1	0.05	Y
1	163.93	Y
1	0.03	Y
8	0.26	N
7	0.55	N
1	0.24	Y
6	2.67	Y
7	0.48	N
1	0.14	Y
8	0.15	N
5	0.03	Y
1	241.06	Y
8	0.23	N
5	0.09	Y
7	1.14	N
7	0.39	N
7	1.25	N
9	0.24	N
1	0.18	Y
7	1.25	N
5	0.12	Y
7	0.90	N

Table D.5 Nanoindentation results for P7:My. A subset of well-fitting points was chosen for calculation and reported.

Fit	Modulus (GPa)	Outlier? (Y=Yes, N=No)
2	2.32	Y
0	4.06	Y
0	1.74	Y
1	48.41	Y
1	1799.22	Y
1	0.05	Y
7	0.07	N
1	0.03	Y
1	0.02	Y
7	0.26	N
7	0.56	N
7	0.33	N
8	0.18	N
6	0.04	Y
6	0.02	Y
4	2.36	Y
7	0.27	N
7	1.16	N
7	0.41	N
7	0.35	N
7	0.29	N
7	0.77	N
7	0.35	N
7	0.52	N
7	0.21	N
4	3.81	Y
7	0.12	N
4	0.10	Y
7	0.07	N

Table D.6 Nanoindentation results for P4AN. A subset of well-fitting points was chosen for calculation and reported.

Fit	Modulus (GPa)	Outlier? (Y=Yes, N=No)
0	2.83	Y
0	3.21	Y
9	0.18	N
3	0.05	Y
3	0.02	Y
0	0.01	Y
0	0.03	Y
0	0.02	Y
7	0.22	N
3	0.80	Y
8	0.17	N
8	0.13	N
8	0.24	N
8	0.41	N
2	0.07	Y
5	1.58	Y
5	2.94	Y
5	4.69	Y
5	4.02	Y
7	1.01	N
5	0.36	Y
8	0.29	N
1	0.17	Y
7	1.71	N
7	0.09	N
7	0.19	N
7	0.16	N
7	0.13	N

Table D.7 Nanoindentation results for P4AN:My. A subset of well-fitting points was chosen for calculation and reported.

Fit	Modulus (GPa)	Outlier? (Y=Yes, N=No)
7	1.12	N
7	0.42	N
7	0.32	N
7	0.11	N
7	0.71	N
7	0.14	N
7	1.24	N
4	1.77	Y
6	2.42	Y
1	0.14	Y
7	0.74	N
7	1.20	N
7	0.65	N
1	0.05	Y
1	0.11	Y
7	0.62	N
7	0.69	N
7	1.14	N
1	0.08	Y
0	0.19	N
0	0.21	N
0	0.25	N
0	0.20	N
0	0.25	N
0	0.24	N