

DESIGN, FABRICATION AND MOLECULAR MODELING OF PROTEIN
SUBUNITS FOR USE IN A NOVEL HYDROGEL

By

CHRISTOPHER LIAM GAUGHAN

A Dissertation submitted to the
Graduate School-New Brunswick
Rutgers, The State University of New Jersey
in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Chemical and Biochemical Engineering

written under the direction of

David Ira Shreiber

and approved by

New Brunswick, New Jersey

[October 2009]

ABSTRACT OF THE DISSERTATION

Design, Fabrication And Molecular Modeling Of Protein Subunits For Use In A Novel
Hydrogel

by CHRISTOPHER LIAM GAUGHAN

Dissertation Director:

David Ira Shreiber

Use of bioinspired, genetically engineered proteins in tissue engineering scaffolds represents a new opportunity for engineering these constructs. However, the production and rational modification of new, artificial proteins is hindered by significant gaps in knowledge regarding expression of artificial gene constructs in *E. coli* and their molecular modeling. This thesis focuses on the production of a novel hydrogel scaffold composed of four self-assembling protein modules and their rational modification using Molecular Dynamics (MD) simulations. Two of the modules are based on the ABA triblock copolymer design. In this triblock, a hydrophilic, random coiled region is flanked by 28 amino-acid α -helical endblocks. The purpose of these endblocks is to function as virtual crosslinkers and support network formation. The length of the endblocks can be changed by the addition of two unlinked, fiber-forming peptides and thus potentially alter the gelation and melting points of the hydrogel. We evaluate the efficacy of production of these endblocks by two separate expression strategies in *E. coli* and demonstrate their ability to form hydrogels. Furthermore, we analyze the Gibbs free energy of formation of oligomeric intermediates that arise early on during fibrillogenesis from the unlinked

peptides using the MM/PBSA module of Amber 9. Thermodynamic data demonstrates changes in the primary structure of these peptides affect the stability of the intermediate that seeds fiber formation. This analysis also suggests a shift in the fiber forming mechanism from monomer addition to protofibril addition. We offer how this data can be used to improve interhelical interactions between endblocks and unlinked peptides and how to develop coarse-grain models of fiber formation.

Table Of Contents

ABSTRACT OF THE DISSERTATION	i
Table Of Contents	iv
List Of Figures	vii
List of Tables	xii
Chapter 1- Introduction.....	1
Hydrogels as biomaterials.....	1
Aggregating proteins as materials for hydrogels	2
Genetically Engineered Hydrogels	4
The Saf Peptide System	7
Intermolecular Forces That Stabilize The Staggered Dimer Conformation	10
Molecular Dynamics Simulation of Protein Aggregation	11
Thesis Overview	13
References.....	14
Chapter 2- Hydrogel Fabrication: Expression and Purification of Water Associating Proteins	17
Abbreviations.....	17
Introduction.....	18
Hydrogel Design: Four Self-Assembling Protein Subunits	22
Unlinked Subunit Structure and Fiber Formation.....	23
Linked Subunit Structure	25
Materials and Methods.....	28
Linked Gene Design and Expression Vectors	28

Detailed Methods: Cloning Expression and Purification of Linked Proteins	32
Results	37
Linked Proteins as MBP Fusions	37
Expression and Purification of 6x His Tagged Linked Proteins	42
Hydrogel Formation: Determination of Gelation Temperature	44
Rheometry	46
Discussion	50
Expression of Linked Proteins	50
MBP-Fusion Proteins	51
6x His Tagged Proteins	52
Hydrogel Formation	54
Rheology	56
Conclusion	58
References	59
Chapter 3- Molecular Modeling and MD Simulation of Early Oligomeric Intermediates Formed During Fibrillogenesis of Saf Peptides	62
Introduction	62
Materials and Methods	65
Molecular Dynamics	71
Results	73
Dimer Stability	73
Stability of oligomers	75
Discussion	83

References.....	88
Chapter 4- Discussion	91
References.....	95
Chapter 5- Future Directions	97
Improved Design and Expression of Triblock Proteins	97
Improving Dimer Stability	99
Developing Coarse-Grain Models of Peptide Aggregation and Hydrogelation	102
References.....	103
Bibliography	105
Curriculum Vitae	113

List Of Figures

Figure 1- Helical wheel diagram of a typical coiled-coil. The 'a' and 'd' positions usually occupied by hydrophobic amino-acids while the 'e' and 'g' positions are occupied by oppositely charged ordered pairs	6
Figure 2- Triblock design of Tirrell et al. includes coiled-coil cross linkers separated by long hydrophilic amino-acid repeats known as the [(AG) ₃ PEG] ₁₀ region. Taken from [16].....	6
Figure 3- Formation of the offset, staggered dimer.	8
Figure 4- "Sticky End" fiber elongation mechanism as proposed for the Saf system of peptides. Progressive addition of monomer units to hydrophobic ends propagates the axial lengthening of fibers.	9
Figure 5- Long fibers aggregate laterally to produce thicker fibers.	10
Figure 6- Intermolecular forces promote staggered dimer formation (A) Four pairs of oppositely charged amino-acids (Red- Glu, Blue- Lys), two 'above' (A and C) and two 'below' (B and D). (B) Hydrogen bonding interactions arising from glutamine residues.	11
Figure 7- Helical wheel diagram of typical coiled-coil-complex. Hydrophobic amino acids at positions 'a' and 'd' drive coiled-coil formation while oppositely charged ordered pairs at the 'e' and 'g' positions confer oligomerization specificity.....	22
Figure 8- Amino-acid sequence of unlinked peptides. These 28 amino-acid subunits have alpha helical structures and form long, coiled-coil fibers when combined in solution	24

Figure 9- Elongation of unlinked peptide units into fibers is initiated by (A) formation of a staggered dimeric structure and propagated by (B) "Sticky End" fibrillogenesis by progressive addition of unlinked monomers.....	24
Figure 10- Schematic triblock structure and sequence of the linked protein SAF1L. The 6x histidine purification tag is highlighted in red.	25
Figure 11- Schematic triblock structure and sequence of the linked-protein SAF2L. The 6x histidine tag is highlighted in red.....	26
Figure 12- Due to the fibrillogenic nature of the Saf system of peptides, endblocks could be elongated by addition of unlinked subunits whose sequence is complimentary to that of the endblocks.	27
Figure 13- Cartoon of hydrogel network structure	27
Figure 14- Schematic representation of the pMal plasmid vector with multiple cloning site highlighted (taken from [54]).	29
Figure 15- Schematic representation of the pJ201 vector with Saf1L and Saf2L genes placed between Bgl2 and HindIII restriction sites. T7 viral promoter drives transcription of linked constructs.....	31
Figure 16- 10% SDS-PAGE of samples treated with and without Genenase 1 for four hours. Single fragments in the low molecular weight region can be seen in the Genease 1 treated lanes due to non-specific cleavage.	39
Figure 17- MALDI-TOF spectrum of Saf2L-MBP showing a peak at 58776 Da.....	40
Figure 18- MALDI-TOF spectrum of seven-hour cleavage reaction of Saf2L-MBP.	41
Figure 19- MALDI-TOF spectrum of the same mixture as in Figure 18 with diminished peak at 16.2 kDa and fragmentation peaks in the low molecular weight region.	42

Figure 20- Co-purified samples of his-tagged Saf1L and Saf2L to determine that the linked proteins could be resolved as separate peaks using MALDI-TOF.	43
Figure 21- MALDI-TOF spectrum demonstrating the presence of several impurities, including a peak at 10670 Da. This peak was observed in all samples tested.....	44
Figure 22- An inverted 500uL eppendorf tube of a hydrogel formed by mixing the four protein subunits.....	45
Figure 23- The distribution of gelation temperatures obtained by falling ball experiments was variable.	46
Figure 24- Parallel Plate Rheometry data with temperature sweep showing gelation behavior of (A) Generation 1 mixture and (B) Generation 2 mixture.....	47
Figure 25- Parallel Plate Rheometry data with frequency sweep showing viscoelastic behavior of gels formed from (A) Generation 1 Mixture and (B) Generation 2 mixture.....	49
Figure 26-Side chain locations of the G1, G2 and G3 monomers were threaded onto the structure of tropomyosin to match those in the helical wheel diagram specified by Woolfson et al.....	69
Figure 27- Glutamate-Lysine ordered pairs help stabilize the staggered dimeric structure: (A) and (B) E15-K34, E20-K29, E22-K41, E27-K36. (B) E15-K41, E20-K36, E22-K48, E27-K43, K29-E55, E35-K50.....	70
Figure 28- Eight monomer units were grafted onto tropomyosin to simulate a small fiber ('minifiber') comprised of many such repeating units. Glu-Lys ordered pairs are divide into 4 repeating blocks based on their periodicity along the coiled-coil structure.....	71

Figure 29- Distance between alpha carbon centers of mass (COM) amongst indicated ordered pairs in the dimer (left) and averaged ordered pairs in the minifiber (right) (A) Generation 1 (B) Generation 2.	77
Figure 30- Ordered pair distances in generation 3 dimer (left) vs. minifiber (right). Error bars indicate standard deviation of ordered pair distances with time.	78
Figure 31- Progressive engineering of Saf system monomers produced dimers with different stabilities in simulations.	79
Figure 32- Thermodynamic behavior of the three generations of oligomers is significantly different. G3 oligomerization proceeds with undulations in energetics whereas G2 displays a single energy minimum at the pentameric state.	80
Figure 33- A bias against addition of Type1 and Type2 monomers exists in G3 oligomerization evidenced by static energy changes in the formation of odd numbered oligomers (right) vs. even numbered oligomers (left).	81
Figure 34- G3 oligomers are structurally sound throughout the oligomerization process whereas G1 and G2 oligomers do not attain structural stability until the pentameric state.	82
Figure 35- Protofibril addition is a plausible mechanism as simulated by the joining of (A) 2 pentameric units to form the decamer (blue) vs monomer addition (black) in G2 oligomerization. B) 2 tetrameric units to form the octamer (blue) vs monomer addition (black) to form the octamer in G3 oligomerization.	83
Figure 36- The G2 fiber mixture displays superior coiled-coil forming ability as compared to the G1 mixture.	99

Figure 37- (A) Residues at the 'a' and 'd' positions in the helical wheel diagram represent the hydrophobic pocket region where phenylalanine residues will be placed. B) Staggered dimer with the overlap region highlighted.....	101
Figure 38- Example of placement of the phenylalanine residues in the hydrophobic, overlap region of the staggered dimer to promote pi-stacking in the dimer (A) Placement of 4 phenylalanine residues. B) Placement of two phenylalanine residues.	102

List of Tables

Table 1- MALDI-TOF Results of Expression of MBP-SafL Fusion Proteins	39
Table 2- Proposed identity of 10.67 kDa peaks found in mass spectra.	54
Table 3- Amino-Acid Sequences of Type1 and Type2 Monomers	66
Table 4- Side-Chain Placement on Crystal Structure of Tropomyosin (PDB Identifier 1C1G).....	68

Chapter 1- Introduction

Protein-based biomaterials are increasingly being used for biomedical applications. Though proteins have great potential for use in such materials, their main drawback is that engineering their properties remains elusive. The goal of this research is to describe the design and fabrication of a protein-based hydrogel that could be customized to support the viability and function of potentially any cell type. To enhance the potential to affect modifications, the hydrogel has a modular design. This allows for separate engineering of the functional aspects of each subunit. We illustrate how these various modules are fabricated in *E. coli* from synthetic genes that code for these subunits. At the same time we illustrate how atomistic molecular modeling and simulation can be used to rationally re-design the hydrogel to potentially alter some of its physical properties.

Hydrogels as biomaterials

Hydrogels are defined as materials that, upon hydration, absorb an amount of water no less than 90% of their dry weight. They have been employed as materials for controlled drug-release devices, biosensors, and contact lenses [1-4], and also have compelling use as tissue engineering scaffolds. Hydrogels are excellent candidates for scaffolds because they replicate the hydrated environment of cellular tissue. As a consequence, they allow for the diffusion of nutrients to and waste away from cells housed in the scaffold. Furthermore, their porous nature facilitates invasion by support cells, delivery of trophic molecules, and diffusion of O₂ from blood vessels within or in

proximity to the hydrogel implant. Thus, they mimic the natural cellular environment and match the water content of soft tissue[5].

A Tissue Engineering Scaffold (TES) is essentially an engineered extracellular matrix. As such, it is important that a scaffold can be rationally modified with affect on hydrogel physical properties to provide an optimal environment for the specific cellular or tissue application. Specific properties that can affect tissue and cellular function include the mechanical strength and stiffness; the presence, number and type of cell attachment ligands; pore size distribution; as well as aggregation properties.

The structure of most hydrogels is an aqueous phase “entrapped” in an insoluble mesh, the nature of which typically differentiates among types of hydrogels. This mesh can be the result of polymerization of chemically synthesized molecules into covalently bonded networks or from the aggregation of low molecular-mass organic gelators (LMOG's) into self-assembled fibrillar networks (SAFIN's) [6]. This latter group, known as molecular hydrogels, is the dominant focus of modern hydrogel research [6].

Aggregating proteins as materials for hydrogels

The constituents of the insoluble part of the gel network can be broadly separated into naturally and chemically synthesized monomers. The latter, synthesized monomers generally aggregate into covalently bonded networks whereas naturally occurring LMOG's often aggregate due to van der Waals forces. As described below, the difference in the degree to which each type of monomer can be engineered is significant.

The prime advantages to using a chemically synthesized scaffold is that these materials have been well characterized, are readily available, and are relatively inexpensive. Thus, their targeted formulation towards gel-specific physical parameters is relatively straightforward. However, their synthetic nature presents the disadvantage that they may degrade in situ into potentially cytotoxic by-products. For example, the polymer polyacrylamide is a non-toxic substance but the acrylamide monomers have been shown to be neurotoxic under conditions of chronic exposure [7]. Furthermore, they have limited inherent biofunctionality and thus offer no molecular motifs for cell attachment.

Thus, protein based materials are finding increasing use in tissue engineering [8]. Such materials imitate the natural extracellular surroundings accurately and are excellent candidates for cell compatibility. This is especially true of protein hydrogels [8-11]. With respect to their use as scaffolds for 3-D cell culture and tissue engineering, proteins offer the advantage of better replicating the in vivo cellular environment as most natural scaffolds (e.g. collagen) are protein based [11]. Hydrogels composed of protein subunits are particularly advantageous due to inherent potential for biofunctionality providing binding sites that govern cell adhesion. Furthermore, since proteins have the innate ability to alter their 3-D conformations in response to changes in temperature, pH, and ionic strength, many protein hydrogels can be structurally sensitive to stimuli, and can easily be developed to form a fibrillar network in the presence of cells, thereby trapping them directly in the 3D construct, rather than relying on cell migration for invasion into the construct

Genetically Engineered Hydrogels

Advances in the fields of gene synthesis and controlled expression of proteins from plasmid vectors in bacterial hosts have opened new avenues for the use of proteins as biomaterials. One of the major advantages to using genetically engineered proteins is that biopolymers can be produced that have uniform composition, molecular weight and amino-acid sequences [12]. Furthermore, genetically engineered protein polymers offer an astonishing degree of resolution with respect to engineering properties as a choice of 20 amino acids can be placed into the covalently bonded backbone of a peptide or protein. As a consequence, proteins can potentially be made with tailored features that exhibit different fiber forming and gelation properties [13]. The design of proteins for use as LMOG's in hydrogels reflects naturally occurring protein folding and aggregation secondary structures such as β -sheets and α -helices.

For example, the α -helical secondary structure governs aggregation of the e leucine zipper protein-folding motif, based on the residue sequence of the Jun oncogene product. This motif is characterized by repeating units of seven amino acids often written as abcdefg. The a and d residues are usually the hydrophobic amino-acids leucine and the e and g positions occupied by amino acids with opposite charge (e.g. glutamate-lysine pairs). The consequence of the placement of these key amino acids is the formation of a coiled-coil, where the a and d residues run along a hydrophobic interface between the two α -helices and thus foster their aggregation of the due to the hydrophobic effect. The e and g residues serve to stabilize interhelical interactions (**Figure 1**).

Genetically engineered hydrogels have been developed based on the leucine zipper motif in the lab of D.A. Tirrell [14-17]. The hydrogels have a triblock architecture consisting of terminal leucine zipper domains separated by a long hydrophilic linker region (**Figure 2**). This linker region consists of repeating units of $[(AG)_3PEG]_{10}$ and serves as a unit to allow gel swelling behavior. The leucine zipper domains serve as crosslinking regions that governed pH dependent reversible gelation due to ionization of the e and g residues at the core positions.

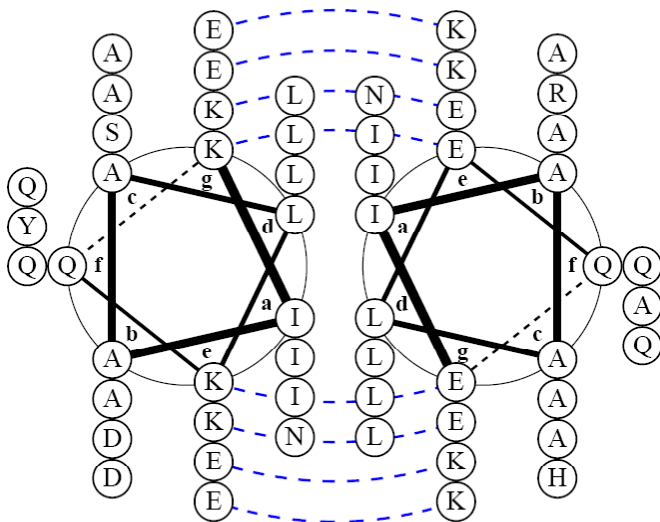


Figure 1- Helical wheel diagram of a typical coiled-coil. The 'a' and 'd' positions usually occupied by hydrophobic amino-acids while the 'e' and 'g' positions are occupied by oppositely charged ordered pairs

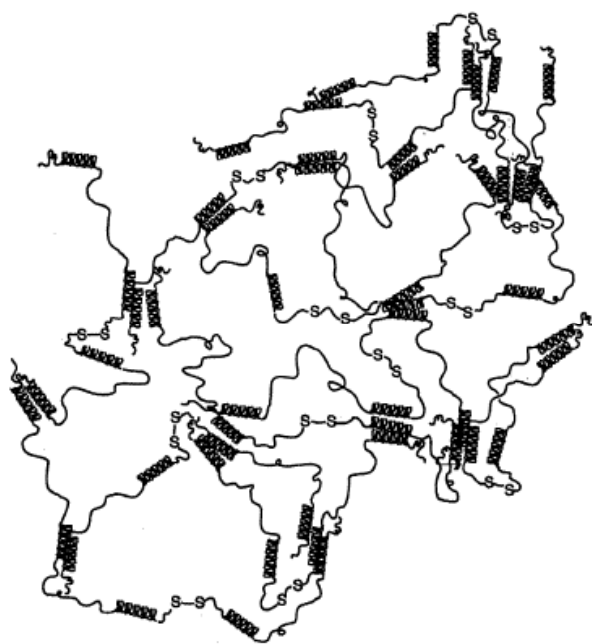


Figure 2- Triblock design of Tirrell et al. includes coiled-coil cross linkers separated by long hydrophilic amino-acid repeats known as the $[(AG)_3PEG]_{10}$ region. Taken from [16].

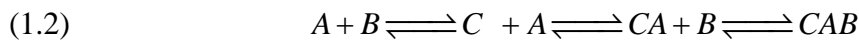
The Saf Peptide System

Several small peptide alpha-helices have been designed by Woolfson et al.[18, 19] (known as the Saf System of peptides) that have the potential to be incorporated into biomaterials. These peptides are unique in that they can, on mixing of complimentary pairs in solution, self assemble into protein fibers several hundred microns long. These fibers have been shown to be composed of long coiled-coils formed by the aggregation of the alpha-helical peptides. Furthermore, changing certain amino acids in the modules has resulted in fibers with different physical properties. The exact mechanism by which these peptides aggregate to form fibers has not yet been elucidated. However, the formation of a staggered-dimeric intermediate has been implicated as being critical for fiber formation and has been proposed as a seeding structure for subsequent fibrillogenesis [18].

The mechanism of aggregation of the Saf peptides has not yet been established by experimental means. However, a theoretical mechanism has been suggested based on peptide design parameters (outlined below). The first step in fiber forming pathway has been proposed to be formation of a staggered dimer:



where A and B are the 28 (or 35) amino-acid alpha-helices and C is the critical dimeric intermediate that seeds addition of progressive alpha-helical units as in:



and thus permit fiber elongation. We refer to the multimers that appear during the early, axial elongation steps (after dimerization) as oligomeric intermediates.

Dimerization is initiated when two monomeric random coils join together in solution as in **Figure 1**.

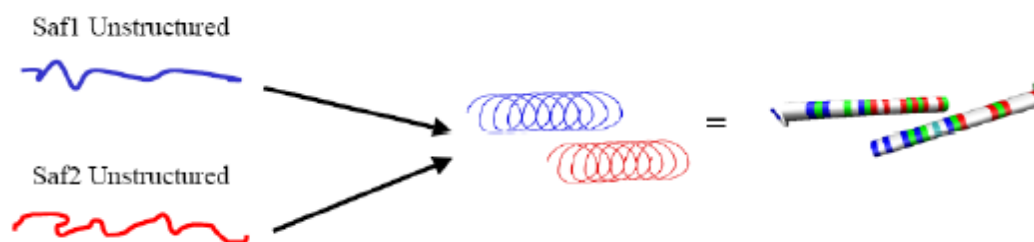


Figure 3- Formation of the offset, staggered dimer.

The main consequence of the dimer's staggered orientation is that hydrophobic amino acids are left exposed to the polar solvent. In the second step of fibrillogenesis, staggered dimers serve as seeds for axial elongation of fibers by virtue these exposed hydrophobic amino acids. Thus, fiber elongation is thought to proceed by progressive addition of Saf monomeric units to these exposed non-polar side chains. In so doing, however, the placement of each monomer produces a new exposed hydrophobic end that then leads to the addition of another monomer and a type of chain reaction is propagated (**Figure 4**).

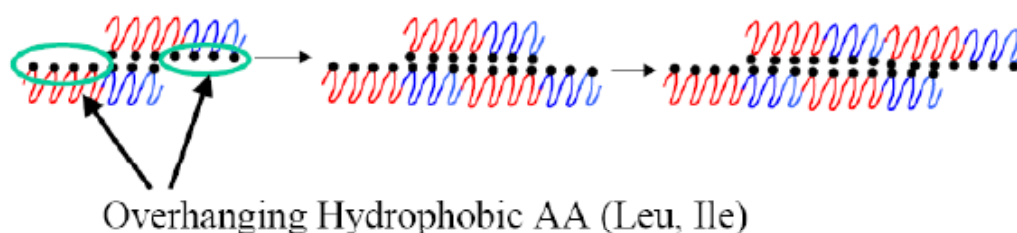


Figure 4- "Sticky End" fiber elongation mechanism as proposed for the Saf system of peptides. Progressive addition of monomer units to hydrophobic ends propagates the axial lengthening of fibers.

It is theorized that once fibers have achieved some critical length, they begin to aggregate laterally and thus become thicker (**Figure 5**) [18]. Fiber diameters of 45, 68 and 58 nm have been reported for three separate complimentary pairs of Saf peptides [20].

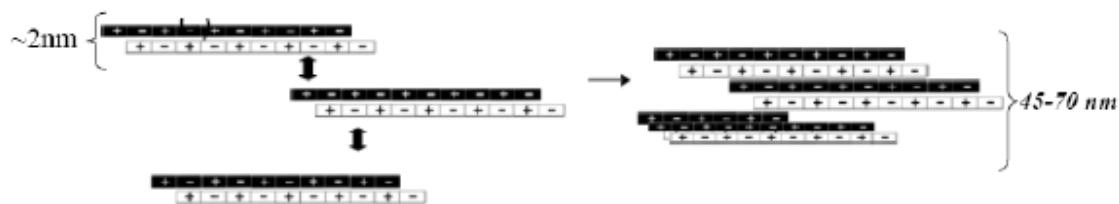


Figure 5- Long fibers aggregate laterally to produce thicker fibers.

Intermolecular Forces That Stabilize The Staggered Dimer Conformation

The Saf monomers were designed with two structural motifs that foster the staggered dimer as a significant conformational energy minima:

1. Charge-charge interactions as a result of strategic placement of oppositely charged amino-acid side chains.
2. Hydrogen bonding interactions resulting from placement of asparagine residues.

Charge-charge interactions are exerted through four pairs of oppositely charged lysine and glutamate residues. These residues are placed as co-planar pairs; two above and two below an imaginary plane bisecting the dimer (**Figure 6**).

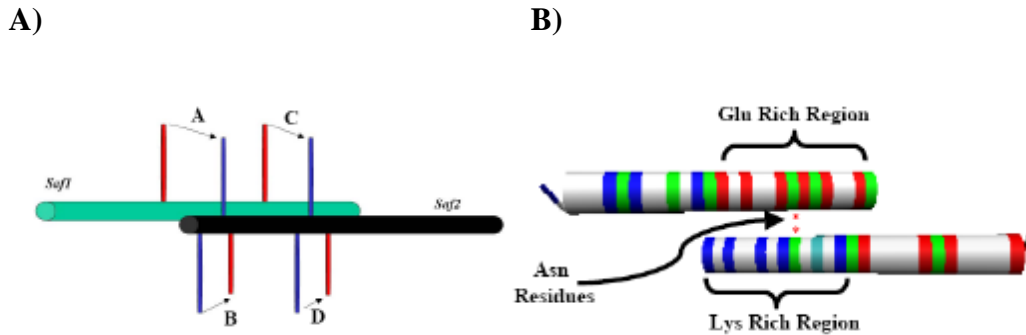


Figure 6- Intermolecular forces promote staggered dimer formation (A) Four pairs of oppositely charged amino-acids (Red- Glu, Blue- Lys), two 'above' (A and C) and two 'below' (B and D). (B) Hydrogen bonding interactions arising from glutamine residues.

Molecular Dynamics Simulation of Protein Aggregation

While proteins have tremendous potential for modification, it is still unclear how changes in amino acid sequences affect final biomaterial properties. This is primarily a consequence of the relatively limited knowledge that presently exists on protein-protein interactions and protein folding, which This is especially true in comparison to well characterized chemically synthesized monomers. Thus, efforts are underway to elucidate the mechanisms that underlie protein aggregation and folding pathways.

While experimental techniques such as NMR, CD Spectroscopy, X-ray crystallography and calorimetry have yielded a wealth of information on these aspects of

protein properties, they can be extremely expensive and time consuming. Furthermore, as in the case of X-ray crystallography, only a select number of proteins are amenable to investigation via these experimental techniques. These circumstances have spurred the development of computational techniques to analyze and predict protein behavior with changes in amino-acid sequence and under various experimental conditions.

The computational methods that address protein-protein interactions and folding can be broadly subdivided into two classes: 1) Course-grain models and 2) atomistic models. Both models have their advantages and disadvantages. In the case of course grain models, one can take advantage of approximations of protein structure to minimize the strain on the limitations of present computing power. For example, modeling proteins as “beads on a string” divides a protein sequence into broad categories, or beads, that reflects the groupings of amino acid side chains with respect to polarity and charge. While these models can enable molecular investigations over relatively long periods of time (μs to s), their main limitation is found in ignoring the role that the solvent (water). This limitation is significant because the change in entropy of the solvent ($\Delta S_{\text{solvent}}$) during protein folding and aggregation is often the driving force for minimization of the Gibbs free energy and thus governs the fraction of the predominant form of the protein(s) at equilibrium.

In contrast, computational methods such as Molecular Dynamics (MD) take into account every atom of the system (protein and solvent). In the case of MD, all atoms in the protein are assigned a set of internal coordinates with the solvent explicitly represented. Velocity vectors for each atom are then generated from the Maxwell-Boltzmann distribution at the specified temperature. Interactions between the atoms in the

protein/solvent system are described by a set of potential energy functions commonly known as the “force field”. The force field is parameterized using both quantum mechanical calculations and experimental data from small molecule systems [21]. Newton’s equations of motion are then solved for each atom in the system:

$$(1.3) \quad F_i = -\nabla_i V(r_1, \dots, r_n) = m_i \frac{d^2 r_i}{dt^2}$$

where m_i is the mass of the i^{th} particle.

The main drawback to the use of MD, especially when applied to protein systems, is that it is computationally demanding. Since these simulations represent each atom of both solute and solvent explicitly, a significant proportion of the computation time can be devoted to the trajectory of the solvent alone. Thus, MD studies on proteins must be confined to relatively small systems over short time periods (ns range).

Despite these limitations, MD studies on aggregating protein systems explore the phenomena governing the initial stages of aggregate formation and the properties of early intermediates. Studied aggregation systems include β -amyloid, prion proteins as well as designed α -helical peptides [22-24]. While the β -amyloid and prion protein systems have been extensively explored, there are very few studies regarding MD on biomaterial, self-assembling systems.

Thesis Overview

While several protein hydrogels have been fabricated [25-28], rational modification of their physical properties remains elusive. Control of physical properties is

essential for a biomedically relevant hydrogel scaffold. Our goal was to develop a hydrogel whose physical properties could be modulated in a controlled manner. This would be accomplished through fabrication of four separate protein modules. The advantage to such a design is that each module could be separately engineered and combined in different proportions so that the hydrogel would exhibit pre-defined gelation properties and could be tailored for use with different cell types. Chapter 2 of the thesis describes the design and fabrication of the four protein modules. A significant proportion of this chapter is dedicated to the describing expression of two large protein subunits in *E. coli*. The chapter ends by describing the gelation characteristics of the resulting hydrogel. Chapter 3 focuses on the use of MD to describe the early aggregation events during the formation of the fibers of the hydrogel. Finally, Chapter 4 summarizes the results, discusses the major conclusions and Chapter 5 outlines future work that could be done to improve the prospects of developing robust hydrogels amenable to physical modification.

References

1. Nakagawa, T. and J. Ito, Drug delivery systems for the treatment of sensorineural hearing loss. *Acta Otolaryngol Suppl*, 2007(557): p. 30-5.
2. Hamley, I.W., Peptide fibrillization. *Angew Chem Int Ed Engl*, 2007. **46**(43): p. 8128-47.
3. Dong, S. and X. Chen, Some new aspects in biosensors. *J Biotechnol*, 2002. **82**(4): p. 303-23.
4. Stapleton, F., et al., Silicone hydrogel contact lenses and the ocular surface. *Ocul Surf*, 2006. **4**(1): p. 24-43.
5. Martens, P.J., S.J. Bryant, and K.S. Anseth, Tailoring the degradation of hydrogels formed from multivinyl poly(ethylene glycol) and poly(vinyl alcohol) macromers for cartilage tissue engineering. *Biomacromolecules*, 2003. **4**(2): p. 283-92.
6. Weiss, R.G. and P.T. , *Molecular Gels: Materials with Self-assembled Fibrillar Networks*. 2005.

7. Lehning, E.J., et al., Biochemical and morphologic characterization of acrylamide peripheral neuropathy. *Toxicol Appl Pharmacol*, 1998. **151**(2): p. 211-21.
8. Stegemann, J.P., S.N. Kaszuba, and S.L. Rowe, Review: advances in vascular tissue engineering using protein-based biomaterials. *Tissue Eng*, 2007. **13**(11): p. 2601-13.
9. Daamen, W.F., et al., Elastin as a biomaterial for tissue engineering. *Biomaterials*, 2007. **28**(30): p. 4378-98.
10. Koide, T., Designed triple-helical peptides as tools for collagen biochemistry and matrix engineering. *Philos Trans R Soc Lond B Biol Sci*, 2007. **362**(1484): p. 1281-91.
11. Woolfson, D.N. and M.G. Ryadnov, Peptide-based fibrous biomaterials: Some things old, new and borrowed. *Curr Opin Chem Biol*, 2006. **10**(6): p. 559-67.
12. Dinerman, A.A., et al., Swelling behavior of a genetically engineered silk-elastinlike protein polymer hydrogel. *Biomaterials*, 2002. **23**(21): p. 4203-10.
13. Ehrick, J.D., et al., Genetically engineered protein in hydrogels tailors stimuli-responsive characteristics. *Nat Mater*, 2005. **4**(4): p. 298-302.
14. Shen, W., et al., Assembly of an Artificial Protein Hydrogel through Leucine Zipper Aggregation and Disulfide Bond Formation. *Macromolecules*, 2005. **38**(9): p. 3909-3916.
15. Shen, W., et al., Tuning the erosion rate of artificial protein hydrogels through control of network topology. *Nat Mater*, 2006. **5**(2): p. 153-8.
16. Petka, W.A., et al., Reversible Hydrogels from Self-Assembling Artificial Proteins. *Science*, 1998. **281**(5375): p. 389-392.
17. Wheeldon, I.R., S.C. Barton, and S. Banta, Bioactive proteinaceous hydrogels from designed bifunctional building blocks. *Biomacromolecules*, 2007. **8**(10): p. 2990-4.
18. Pandya, M.J., et al., Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis. *Biochemistry*, 2000. **39**(30): p. 8728-34.
19. Andrew M. Smith, E.F.B., Wayne R. Edwards, Maya J. Pandya, Derek N. Woolfson, Engineering Increased Stability into Self-Assembled Protein Fibers. *Advanced Functional Materials*, 2006. **0000**(00): p. 1-9.
20. Balgude, A.P., et al., Agarose gel stiffness determines rate of DRG neurite extension in 3D cultures. *Biomaterials*, 2001. **22**(10): p. 1077-84.
21. LeSar, A.R.a.R., Modeling and Simulation of Biomaterials. *Annu. Rev. Mater. Res.*, 2004. **34**: p. 279-314.
22. Rezaei-Ghaleh, N., M. Amininasab, and M. Nemat-Gorgani, Conformational changes of {alpha}-chymotrypsin in a fibrillation-promoting condition: A molecular dynamics study. *Biophys J*, 2008.
23. Melnik, T.N., et al., Shift of fibril-forming ability of the designed alpha-helical coiled-coil peptides into the physiological pH region. *Protein Eng*, 2003. **16**(12): p. 1125-30.
24. Anand, P., F.S. Nandel, and U.H. Hansmann, The Alzheimer's beta amyloid (Abeta1-39) monomer in an implicit solvent. *J Chem Phys*, 2008. **128**(16): p. 165102.

25. Stabenfeldt, S.E., A.J. Garcia, and M.C. LaPlaca, Thermoreversible laminin-functionalized hydrogel for neural tissue engineering. *J Biomed Mater Res A*, 2006. **77**(4): p. 718-25.
26. Yan, H., et al., Thermoreversible protein hydrogel as cell scaffold. *Biomacromolecules*, 2006. **7**(10): p. 2776-82.
27. Wang, C., R.J. Stewart, and J. Kopecek, Hybrid hydrogels assembled from synthetic polymers and coiled-coil protein domains. *Nature*, 1999. **397**(6718): p. 417-20.
28. Jung, J.P., et al., Modulating the mechanical properties of self-assembled peptide hydrogels via native chemical ligation. *Biomaterials*, 2008. **29**(13): p. 2143-51.

Chapter 2- Hydrogel Fabrication: Expression and Purification of Water Associating Proteins

Abbreviations

6x- Six repeating monomers (as in 6-Histidines in a row)

A- Alanine

aa- amino acid

Bp- base pairs of nucleotides

D- Aspartate

Da- Daltons (g/mol)

E- Glutamate

G- Glycine

G'- Storage Modulus

G''- Loss Modulus

H- Histidine

His- Histidine

I- Isioleuciene

IMAC- Immobilized Metal Affinity Chromatography

IPTG- Isopropylthiogalactoside

K- Lysine

L- Leucine

MALDI-TOF- Matrix-assisted laser desorption/ionization- Time of Flight (mass spectrometry)

MBP- Maltose Binding protein

MWCO- Molecular Weight Cut-Off

N- Asparagine

P- Proline

Pa- Pascal's (N/m^2)

PEG- Poly-ethylene-glycol

Q- Glutamine

R- Arginine

S- Serine

R- Arginine

W- Tryptophan

Y- Tyrosine

Introduction

Protein based, self-assembling hydrogels are now being investigated for use as tissue engineering scaffolds, drug delivery devices, and 3-D cell culture [1-6]. Use of proteins in these materials stems from the desire to more accurately replicate the extracellular matrix environment [7]. In contrast to hydrogels derived from synthetic monomers, protein hydrogels possess inherent biocompatibility and biodegradability via in situ protein turnover pathways [8]. Furthermore, protein based materials have tremendous potential as engineering-amenable substrates. Proteins serve as building blocks for a wide array of elaborate structures with distinct physical properties. Using macromolecular complexes in hydrogel design also enables modular separation of

functionally independent peptide domains that can be engineered individually. Thus, they are potentially suited for tailoring to any number of specific biomedical applications.

Several protein-based hydrogels have been produced from naturally occurring proteins such as hen-egg lysozyme, fish protein, and fibrin [9, 10]. These proteins are readily available and are inexpensive to produce. While use of these proteins produces robust hydrogels, thousands of years of evolution have already optimized their properties for their specific functions, and as such, they are limited with respect to the amount they can be tailored for different biomedical applications.

Advances in the fields of total gene synthesis and protein expression in bacterial hosts have now augmented the repertoire of proteins that can be used for hydrogel research and manufacture. Furthermore, recombinant production of hydrogel proteins in bacterial hosts ensures monodispersed protein production with defined molecular weights, compositions, sequences and stereochemistry [11, 12]. Thus, proteins can be rationally designed and produced with specific aggregation properties and environmental sensitivities.

Recently, genetically engineered hydrogels have been made as ABA triblock copolymers, in which a hydrophilic B block is flanked on either side by short hydrophobic A blocks [13-15]. At least three such B blocks have been used in hydrogels. The first is an Ala-Gly rich region (denoted $[(AG)_3PEG]_{10}$) marked by a random coil structure. This amino acid sequence makes the region behave as a polyelectrolyte and was developed in the lab of David Tirrell [13]. The second is an elastin like protein sequence composed of Val-Pro-Gly-Val-Gly repeats [16]. The third is the hydrophilic polymer poly-ethylene glycol (PEG) [17].

The A blocks, which function as virtual crosslinkers, are α -helices whose primary structure essentially induces the formation of coiled-coils [18]. This naturally occurring protein folding motif is characterized by a repeating, seven-residue (heptad) sequence of amino acids generally denoted as “a, b, c, d, e, f, g” based on their positions in the coiled-coil complex. The “a” and “d” amino acids are generally hydrophobic residues such as leucine and isoleucine and function as the driving force for coiled-coil formation via the hydrophobic effect. The “e” and “g” residues are generally oppositely charged pairs such as lysine-glutamate and confer specificity of association between helices (**Figure 7**). Use of the coiled-coil motif in the engineered endblocks is advantageous because small changes in the amino-acid sequences can drastically effect thermal stability and specificity of helical association [19], thereby allowing great versatility in design.

Several coiled-coil forming motifs have been used for A block design, including those with hydrophobic residue placement derived from the Jun oncogene product [13], elastin [20], the yeast transcriptional factor GCN4 [14, 21, 22] and fibrin [17]. However, very few of these A blocks have been systematically altered to effect gelation properties. Work by Kopecek et al demonstrated that minor, but tailored alterations to the coiled-coil forming A block (VSSLESK)₆ led to dramatic changes in gelation properties [22]. In these studies, the V and K residues in the “a” position destabilized hydrophobic interactions in the core of the coiled-coil. Replacing S residues for K in the “c” position of the third, fourth, and fifth heptads introduced electrostatic repulsive forces between “c” and “g” residues. Such amino-acid alterations decreased the melting temperature, changed the concentration of peptides needed for gel formation and changed the oligomerization state of the coiled-coil assembly. Other studies by the same group also

demonstrated that changing the length of the A blocks resulted in hydrogels with different physical characteristics [14, 22]. Thus, to fabricate hydrogels with targeted physical characteristics it is essential to use A blocks where changes in amino-acid composition will have a predictable effect on coiled-coil behavior.

To address this issue, we propose using the Saf system of peptides developed by Woolfson et al as endblocks in triblock hydrogels. This system consists of several α -helical peptides that form long, coiled-coil fibers when complementary pairs are combined in solution [23-28]. These peptides have been extensively engineered with respect to their amino-acid composition to create fibers of varying widths, kinetics of formation, melting temperatures, and morphology.

In this chapter we outline the design and fabrication of a novel, genetically engineered hydrogel composed of four peptide subunits or modules. Two of the subunits are based on the triblock ABA architecture and are denoted “linked-units” hereon. The A blocks are type 1 or type 2, 28 amino-acid α -helices which comprise Woolfson's generation 1 fibers [28] and function as coiled-coil crosslinkers (See Below). The B block is the $[(AG)_3PEG]_{10}$ hydrophilic random-coil-linker region [13] described above. The remaining two subunits are the isolated 28 amino-acid α -helical endblocks and are denoted “unlinked-units”. Incorporation of the unlinked units is an innovative feature of the hydrogel design that enables virtual elongation of the of the coiled-coil endblocks which could lead to tunable hydrogel properties such as described by Kopecek et al. above. Thus, there are two potential routes for adjusting hydrogel properties: 1) changing the amino-acid composition of the coiled-coil endblocks and 2) adjusting the length of the crosslinker via addition of the unlinked units.

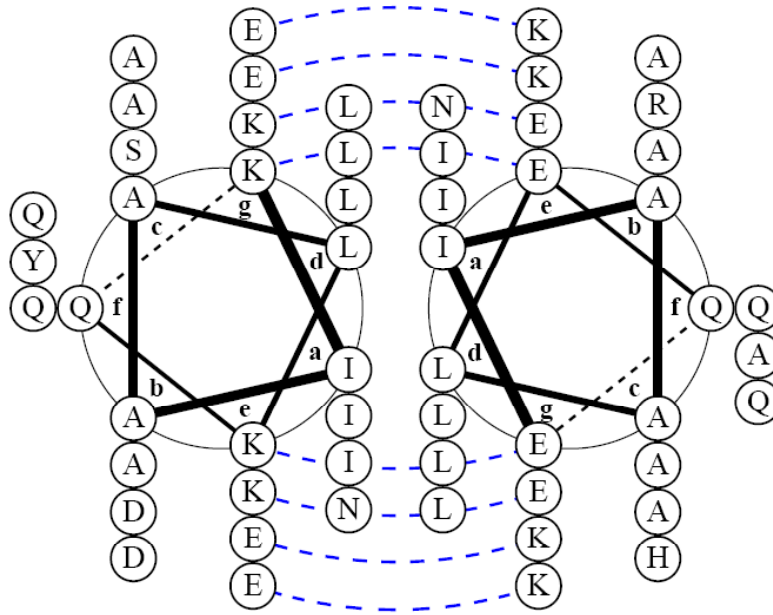


Figure 7- Helical wheel diagram of typical coiled-coil-complex. Hydrophobic amino acids at positions 'a' and 'd' drive coiled-coil formation while oppositely charged ordered pairs at the 'e' and 'g' positions confer oligomerization specificity

Hydrogel Design: Four Self-Assembling Protein Subunits

The hydrogel design incorporates four protein subunits that will allow for separate engineering of these modules. The two linked-units, which have triblock architecture, were expressed from genes that have been synthesized and placed into plasmid vectors for subsequent expression in E coli. The unlinked units are 28 amino-acid alpha helices that were chemically synthesized in a peptide synthesizer.

The peptide units used in the proposed hydrogel are taken from the work of two separate groups. A reversible hydrogel developed by Tirrell et al. consists of multidomain triblock peptides that self assemble at physiologic temperature and pH [13,

29]. Resulting hydrogels were found to be reversible when pH and temperature were varied. Modifications of amino acids in the helices were also found to alter the gel-sol transition point, thus suggesting that material properties can be tailored for use as a tissue engineering construct. Gelation was driven by aggregation of α - helices to form coiled-coil dimers whereas a separate block mediated association with water. We use this latter block in our own design for the water uptake function in the gel.

To augment the length of the coiled-coil crosslinks we have adopted the design of Woolfson et. al. to yield the unlinked protein subunits. These investigators demonstrated fiber formation from the use of two 28 amino acid peptides. Fiber formations occurred via formation of long coiled-coils and were observed to grow as long as 100 μm . In theory, unlinked peptides whose structures are complimentary to the endblock helices will elongate this endblock due to “sticky end” fibrillogenesis (described below). Thus, targeted engineering of these amino-acid sequences could lead to crosslinks with different properties and thus alter the physical characteristics of the resulting hydrogels as described by Kopecek et al. [22].

Unlinked Subunit Structure and Fiber Formation

The self-assembly of the fibers their constituent unlinked peptide is governed by hydrophobic effects. To form a coiled-coil, two α -helices join together to form a dimer (**Figure 9A**). In this dimer, the orientation of the helices with respect to one another is staggered due to the presence of charged side chains within the helix (**Figure 9A**). The consequence of this orientation is that hydrophobic amino acids not bound within the coil itself are left exposed to the polar solvent. This instability is what drives free

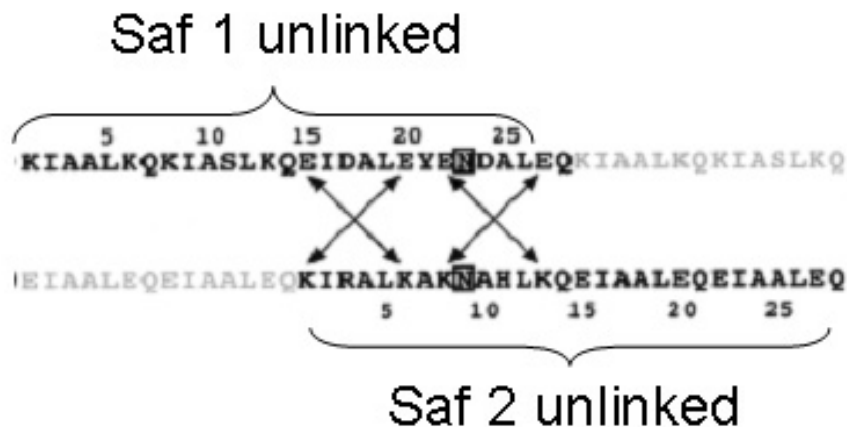
complementary helices to bind to the hydrophobic ends of the growing fiber referred to this mechanism as 'sticky end' fibrillogenesis (**Figure 9B**).

Sequence SAF1: KIAALKQKIASLKQEIDALEYENDALEQ

Sequence SAF2: KIRALKAKNAHLKQEIAALEQEIAALEQ

Figure 8- Amino-acid sequence of unlinked peptides. These 28 amino-acid subunits have alpha helical structures and form long, coiled-coil fibers when combined in solution

A)



B)

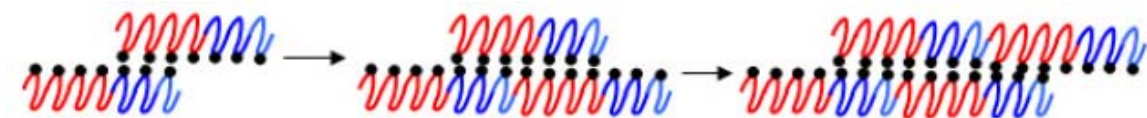
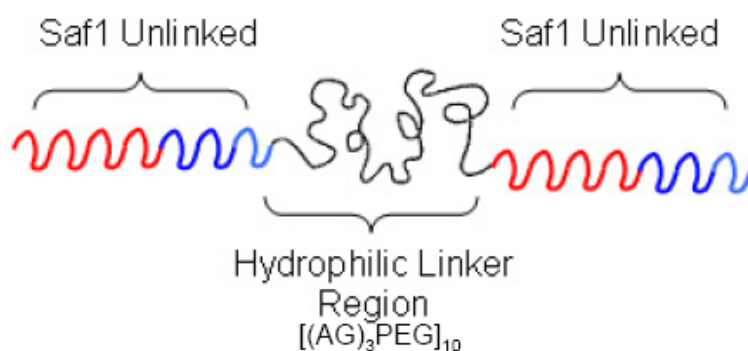


Figure 9- Elongation of unlinked peptide units into fibers is initiated by (A) formation of a staggered dimeric structure and propagated by (B) "Sticky End" fibrillogenesis by progressive addition of unlinked monomers.

Linked Subunit Structure

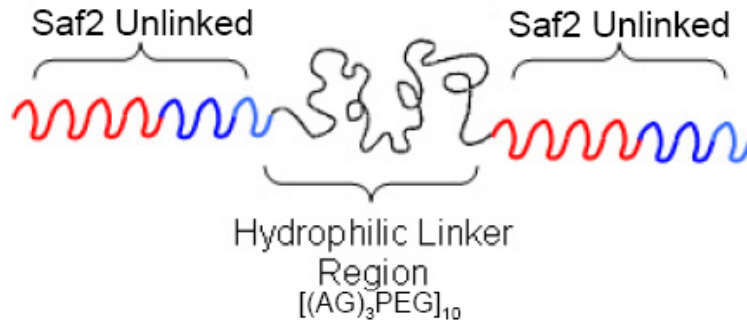
The two linked modules are two SAF units joined by Tirrell's long hydrophilic linker region, $[(AG)_3PEG]_{10}$, to make the triblock (**Figure 10-11**). The general purpose of this 173 amino-acid module is to allow for extensive water association that will foster the formation of a hydrogel, rather than the insoluble mesh of coiled-coil fibers that would result from the addition of the unlinked units solely.



Sequence: KIAALKQKIASLKQEIDALEYENDALEQ MRGS**HHHHHH**HGS
 ASYRDPMGAGAGAGPEGAGAGAGPEGAGAGAGPEGAGAGAGPEGA
 GAGAGPEGAGAGAGPEGAGAGAGPEGAGAGAGPEGAGAGAGPEGA
 GAGAGPEGARMPTSW KIAALKQKIASLKQEIDALEYE NDALEQ

Figure 10- Schematic triblock structure and sequence of the linked protein SAF1L.

The 6x histidine purification tag is highlighted in red.



Sequence: KIRALKAKNAHLKQEIAALEQEIAALEQ MRGS**HHHHHH**HGS
 ASYRDPMGAGAGAGPEGGAGAGPEGAGAGAGPEGAGAGAGPEGA
 GAGAGPEGAGAGAGPEGAGAGAGPEGAGAGAGPEGAGAGAGPEGA
 GAGAGPEGA RMPTSW KIRALKAKNAHLKQEIAALEQEIAALEQ

Figure 11- Schematic triblock structure and sequence of the linked-protein SAF2L.

The 6x histidine tag is highlighted in red.

Since fiber propagation occurs via the non-covalent, “sticky-end” fibrillogenesis described above, the addition of the unlinked units could stabilize the coiled-coil crosslinks by virtue of their fibrillogenic capacity (**Figure 12**). This would result in a network of elongated, fibrous endblocks interrupted by the hydrophilic random coil region of the triblock (**Figure 13**).

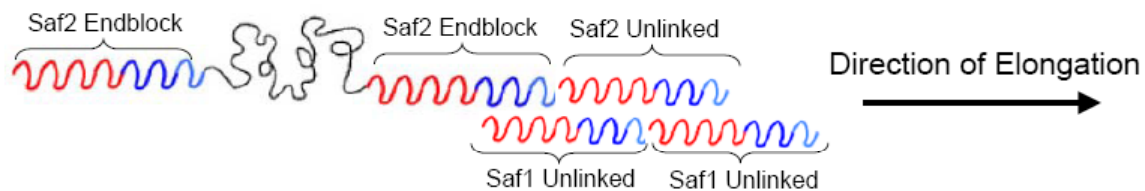


Figure 12- Due to the fibrillogenic nature of the Saf system of peptides, endblocks could be elongated by addition of unlinked subunits whose sequence is complimentary to that of the endblocks.

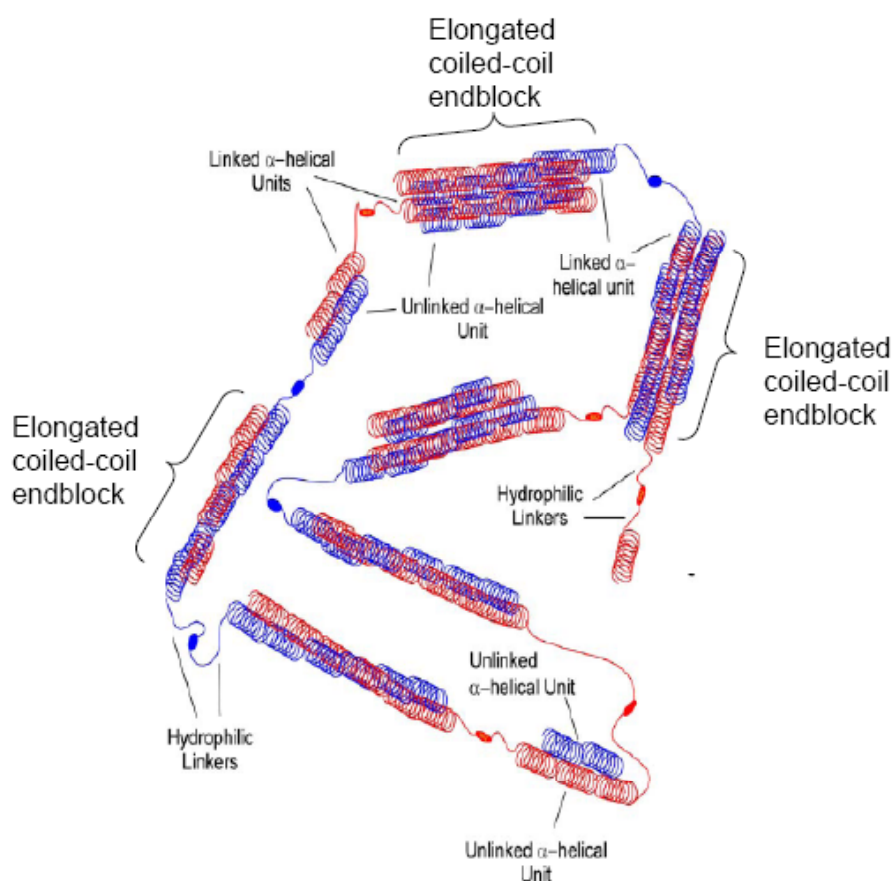


Figure 13- Cartoon of hydrogel network structure

Materials and Methods

Linked Gene Design and Expression Vectors

Overview

The linked proteins (Saf1L and Saf2L) were expressed in *E. coli* from artificial genes encoding the triblock in a single open reading frame. The genes were synthesized at DNA 2.0 (Menlo Park, CA) with codon usage optimized for expression in *E. coli*. Improvements were made to the original gene encoding the $[(AG)_3PEG]_{10}$ region used in Petka et al. [13] by eliminating some of the repetitive codon usage inherent to the AGAGAG peptide sequence. This was to minimize RNA secondary structures that would interfere with efficient translation.

We used two strategies for expression of the linked genes: 1) Expression as a Maltose Binding Protein (MBP) fusion and 2) Expression of the proteins alone without a fusion partner. Both strategies have their advantages and potential pitfalls.

Maltose binding protein serves not only as an affinity tag, but also aids solubility and refolding of recombinant proteins [30, 31]. MBP fusion proteins are purified via amylose column chromatography. The pMAL vector system (New England Biolabs, Ipswich, MA) contains restriction sites for insertion of the foreign gene and encodes a sequence of amino acids on the N-terminus of the target protein for proteolytic cleavage

after purification (**Figure 14**). The vectors use the tac promoter and MBP translation initiation sequences for high-level expression.

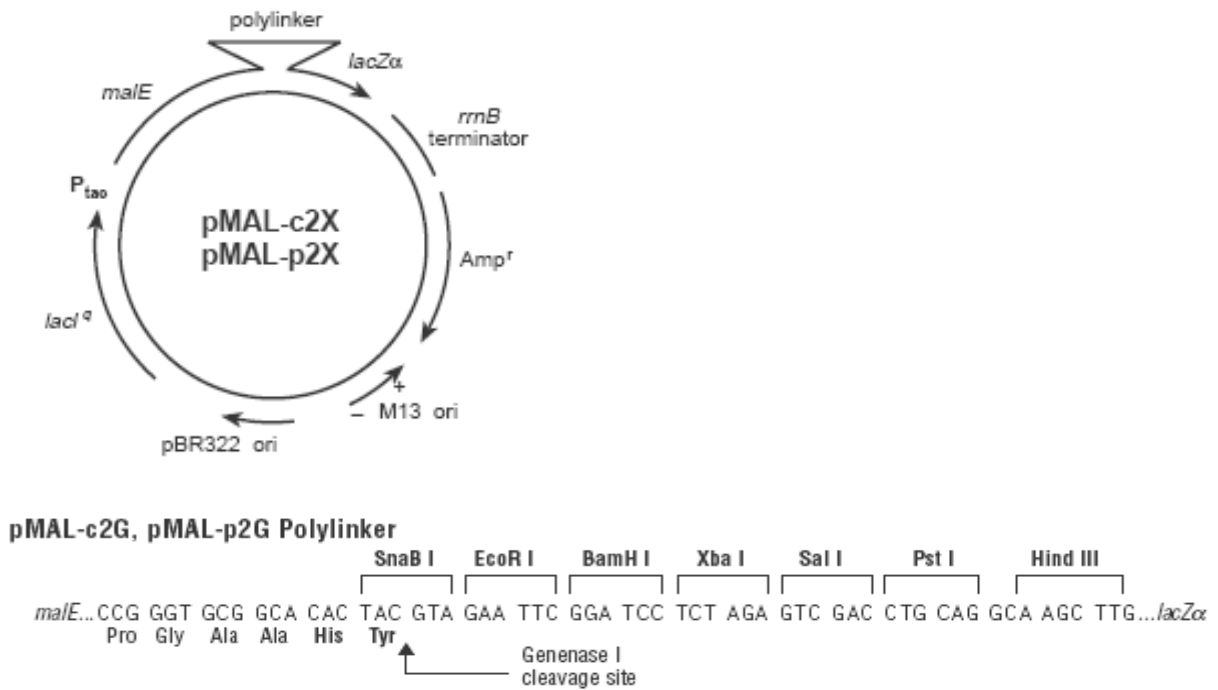


Figure 14- Schematic representation of the pMal plasmid vector with multiple cloning site highlighted (taken from [54]).

The second expression strategy, protein without a fusion partner, relies on the presence of a hexahistidine tag (6x His Tag) for purification. High-level expression is achieved due to the use of the T7 viral promoter to drive expression of the linked genes. There are several advantages to using this tag. First, purification is relatively straightforward using immobilized metal affinity chromatography (IMAC). In this case, chelated divalent metal ions such as Ni^{2+} , Co^{2+} and Zn^{2+} serve as the affinity ligands.

IMAC chromatography occurs under mild conditions and can be performed under both native and denaturing conditions. Adding a small amount of imidazole to the loading buffer can minimize non-specific binding to the column. Second, histidine tags have very little, if any, effect on protein conformation and solubility. This is mostly due to their small size. IMAC is not recommended for proteins containing metal ions. Furthermore, bacterial proteins that are naturally histidine rich may also bind to IMAC columns and appear as unwanted contaminants [32].

Modern protein expression has been overtaken by utilization of the T7 system of pET vectors (Merck/EMD). This promoter system is a hybrid of the T7 RNA polymerase promoter and the lac operator. Expression is driven by transcription of the highly processive T7 RNA polymerase, the gene for which lies on a prophage (λ DE3) and is under the control of the IPTG inducible lacUV5 promoter. Recombinant protein accumulations using T7 driven transcription is ~40-50% of total cell protein [33].

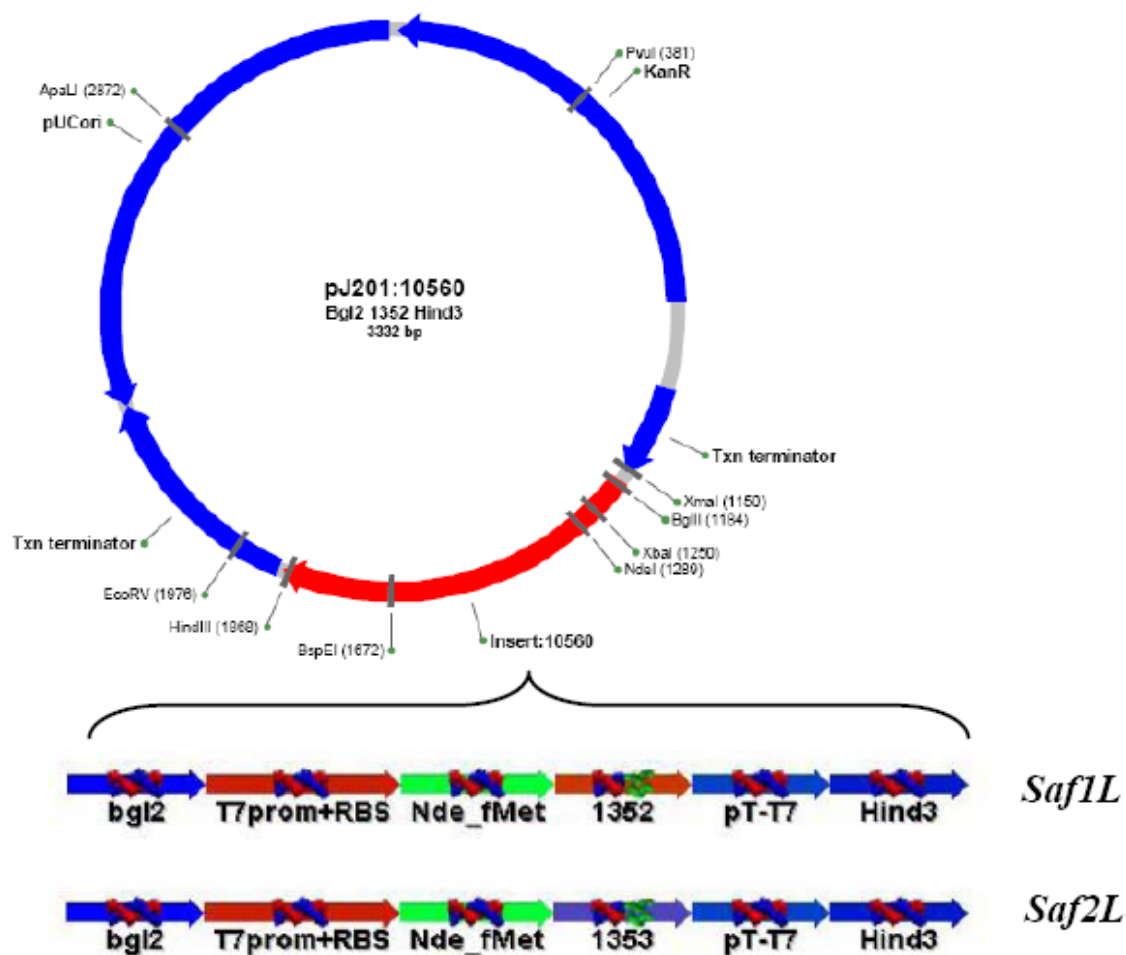


Figure 15- Schematic representation of the pJ201 vector with Saf1L and Saf2L genes placed between Bgl2 and HindIII restriction sites. T7 viral promoter drives transcription of linked constructs.

Detailed Methods: Cloning Expression and Purification of Linked Proteins

Linked proteins fused to MBP

Cloning

For expression as MBP fusion, Saf1L and Saf2L were subcloned from pBAD vector (DNA 2.0) into pMal-c2G (NEB, Ipswich, MA) between the SnaBI and HindIII (NEB, Ipswich, MA) restriction sites and ligated into the pMal vector with T4 DNA ligase (NEB). These recombinant vectors are denoted c2g-1 and c2g-2 for plasmids containing the Saf1L and Saf2L constructs respectively. The cloning host XL-1 Blue (Stratagene, La Jolla, CA) was used to amplify pMal constructs on LB plates containing ampicillin. Resistant colonies were selected and pDNA amplified by growth of single colonies in 250 mL of LB Broth. Plasmid DNA was subsequently purified using the MaxiPrep protocol (Qiagen, Valencia, CA).

Expression

Each gene was expressed separately in the expression host BL-21 Gold. 5ng of each plasmid was used to transform bacteria and transformants selected on ampicillin LB/Agar plates. Eight colonies from each plate were selected and tested for expression of the MBP-SAF1 fusion using 100mL cultures of LB Broth containing ampicillin (100ug/mL). Expression was induced during the mid-log phase of bacterial growth ($OD_{600} \sim 0.5$) by addition of 1 mM isopropylthiogalactoside (IPTG). Cells were grown for 4 hours and 25 uL aliquots centrifuged, dissolved in SDS-PAGE loading buffer and

subsequently analyzed for expression by SDS-PAGE by the presence of the ~58 kD band for each construct.

Colonies positive for expression were grown in 1L cultures containing 100 ug/mL of ampicillin. Expression was induced during the mid-log phase of bacterial growth ($OD_{600} \sim 0.5$) by addition of 1 mM isopropylthiogalactoside (IPTG) in three 1L LB suspension cultures. Cells are harvested by centrifugation at 5000 x g.

Purification

Pelleted bacteria were resuspended in Tris-EDTA buffer and frozen at -20°C overnight to cause freeze-thaw rupture of bacterial cell membranes. Bacteria were sonicated on ice to break open cell membranes and centrifuged at 9000 x g in order to obtain a crude extract of cytoplasmic proteins. These crude extracts were loaded into a 2.5 x 10cm column packed with 10 mL of amylose resin in order to isolate MBP fused to the Saf1L and Saf2L proteins. The basis of this separation is that the amylose resin binds the Maltose Binding Protein. The column is eluted with 10mM Maltose. 15mL of column resin generally yields 45mg fusion protein/L of culture (NEB).

The eluted fusion protein fractions were checked for their absorbance at 280 nm, the protein fractions pooled and then concentrated in Amicon Ultra (Millipore Corp, Bedford) centrifugal filters with a 10,000 molecular weight cut-off (MWCO). The retentates containing the fusion protein were then cleaved using the protease Genenase I at a concentration of 1mg protease/ 100 mg fusion protein for 7 hours at room temperature. Cleaved protein was then analyzed via mass spectrometry (MALDI-TOF) to determine protein purity and extent of cleavage.

Linked Proteins With 6x His Tag

No cloning steps were required as both genes encoding the linked proteins Saf1L and Saf2L were placed between the BglII and HindIII sites of expression plasmid pJ201 at DNA 2.0 (**Figure 15**).

Expression

The expression host BL-21 (DE3) (Stratagene, La Jolla, CA) was transformed with the recombinant pJ201 plasmids as described above except that kanamycin resistance was used as the selective marker for positive transformants on LB/Kanamycin plates (Kanamycin 30 ug/mL) (Plates- Teknova, Hollister, CA). Since both Saf1L and Saf2L were found to be expressed at low levels, expression studies proceeded directly to large-scale expression.

Single kanamycin resistant colonies were picked and used to inoculate individual 2L cultures of Terrific Broth (Sigma Aldrich, St. Louis, MO) without kanamycin. Cells were grown at 37°C in baffled flasks with constant shaking. When culture growth had achieved mid log phase ($OD_{600} \sim 0.5-0.7$) expression was induced by addition of 1mM IPTG (Sigma-Aldrich, St. Louis, MO). Cultures were incubated for four hours and cells subsequently harvested by centrifugation at 4°C and 4000xg. The cell free broth was removed and cells frozen overnight at -20°C.

Purification

Protein was purified under native conditions. Cells were thawed and resuspended by pipetting up and down in 5mL/g wet weight cells column Equilibration/Wash Buffer (50 mM sodium phosphate, 300 mM sodium chloride, 10 mM imidazole; pH 7.4) (Pierce

Biotechnology, Rockford, IL) on ice supplemented with 0.5mM PMSF (phenylmethylsulphonyl fluoride)(Sigma-Aldrich, St. Louis, MO). Resuspended cells were then sonicated on ice at medium amplitude with 2-second pulses for 2 minutes until sonicates became slightly transparent. Crude extracts were prepared by centrifugation at 12,000xg for 45 minutes at 4⁰C utilizing a fixed angle rotor.

Column purification was carried out at 4⁰C. The resulting crude extract from each batch (~30 mL) was loaded into pre-equilibrated 3mL His Pur Cobalt Spin Columns (Pierce Biotechnology) and allowed to move through the resin under the influence of gravity alone. Columns were washed with wash buffer containing a slightly higher concentration of imidazole (50 mM sodium phosphate, 300 mM sodium chloride, 30 mM imidazole; pH 7.4) to remove non-specifically bound bacterial proteins. 6x His tagged protein was eluted from the column by addition of elution buffer (50 mM sodium phosphate, 300 mM sodium chloride, 150 mM imidazole; pH 7.4). Eluted protein was frozen at -80⁰C and then lyophilized overnight.

Extent of cleavage and protein purity was determined using mass spectrometry (MALDI-TOF).

Characterization of Triblock/Linked Peptides

The molecular weights of the linked peptides produced from as MBP-fused and non-fused 6x-His tag were determined by MALDI-TOF MS. For MALDI characterization, one batch of both 6x his-tagged unlinked proteins were purified in a single column to determine whether they could be resolved as distinct peaks.

Synthesis of Unlinked Proteins

The unlinked proteins (Saf1, Saf2 and Saf2-p2a) were made at the Custom Peptide Services at Chi Scientific (Maynard, MA). Peptides were delivered desalted and at approximately 70-75% purity.

Hydrogel Formation

Falling Ball

The temperature at which hydrogels formed was tested by the falling ball method [34]. Only the 6x His tagged linked-proteins were used for making hydrogels. 400-500ug of unlinked and linked proteins were weighed on an analytical balance and dissolved in a 500uL eppendorf tube in 50 uL of 10mM MOPS buffer (pH 7.4) (Sigma Aldrich, St Louis, MO). A 0.5mm ruby ball was placed in the dissolved mixture of peptides and the tube containing the peptides heated in a water bath. Beginning at room temperature, the tube was inverted every 5⁰C rise in temperature to examine the movement of the ruby ball through the peptide mixture. When the ball ceased to move, gelation had been reached.

Rheometric Measurements

Protein solutions were prepared as described above and kept on ice until measurements were about to be taken. To assess whether changing the nature of the cross-link changed gelation behavior, the original protein mixture containing saf1-unlinked/saf2-sunlinked/Saf1L/Saf2L (denoted G1 mixture) was compared to an

additional protein mixture was made replacing the saf2-unlinked peptide with a second generation, saf-p2a unlinked peptide (denoted G2 mixture). Rheological studies were performed to determine the gel point, storage modulus (G'), and loss modulus (G'') via use of a stress-controlled rheometer (Bohlin). Parallel plates (15mm in diameter) with a 0.5 mm gap were employed for all samples. Liquid peptide solution (~100uL) was pipetted onto the lower plate of the rheometer and the top plate lowered until the proper gap distance was achieved. A temperature sweep was then performed from 25⁰C to 85⁰C by constant heating (peltier device) of the lower plate at 5⁰C/min. Measurements of G' and G'' were made under constant stress (1Pa) and constant frequency of 1 rad/s.

After gelation, the gelled samples were subjected to a frequency sweep (0.1-10 Hz) at 37⁰C and constant stress (1 Pa) to assess the viscoelastic character of the gels.

Results

Linked Proteins as MBP Fusions

High expression of Saf1L-MBP and Saf2L-MBP fusion proteins was observed with average yields between 45-50 mg of fusion protein per liter of cells cultured. Typical yields range between 10-40 mg/liter culture [35]. This corresponds to a theoretical yield of 12-14mg of each unlinked protein per liter of culture.

Purity of the amylose purified fusion constructs was not tested quantitatively. However, eluted fractions collected from the column yielded only one major absorption peak at 280nm (data not shown). The identity of these peaks as MBP-SafL fusion proteins was confirmed by MALDI-TOF.

Both linked proteins migrated through acrylamide gels much more slowly than expected, with apparent molecular weights approximately 10,000 Da heavier than

calculated (**Figure 16**). This also affected migration of the MBP-SafL fusion proteins. In fact, this was cause for some delay in our progress, as we believed β -galactosidase was being expressed with the fused protein. The β -galactosidase gene (*lacZ α*) is directly downstream of the MBP-SafL open reading frame. It was not until mass spectrometry was performed did we realize that the proper gene products were being expressed.

As can be seen in **Figure 16**, satisfactory cleavage of both Saf1L and Saf2L was achieved utilizing Genenase I. However, it is also clear that from the gel in **Figure 16** that additional cleavage fragments are generated. The likely cause is from Genenase I proteolysis at non-canonical amino-acid sequences within the linked proteins themselves. This is evidenced by the presence of single, decomposition fragments in the lanes for both proteins immediately after the four hour cleavage reaction (**Figure 16**). Once started, cleavage of the linked proteins by Genenase I proceeded rapidly to completion.

While both Saf2L-MBP and Saf2L (**Figure 17-18**) could both be detected using mass spectrometry, only the Saf1L-MBP fusion protein could be detected in our experiments. Cleavage produced no detectable MALDI peaks in the predicted regions, only peaks in the low molecular weight region (data not shown). These fragmentation peaks also appeared in the case of Saf2L (**Figure 19**).

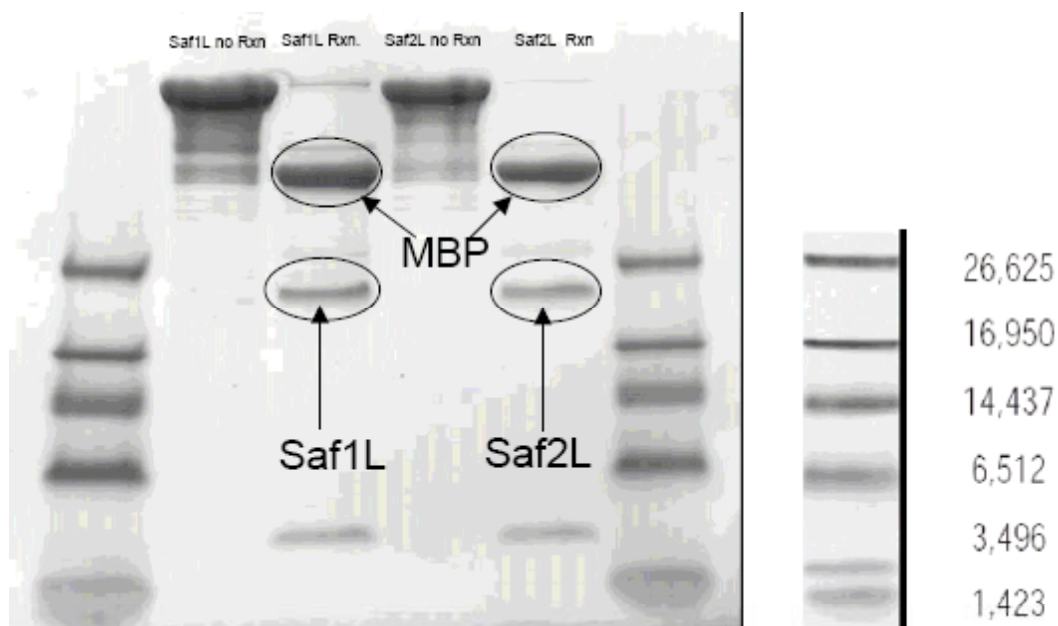


Figure 16- 10% SDS-PAGE of samples treated with and without Genenase 1 for four hours. Single fragments in the low molecular weight region can be seen in the Genease 1 treated lanes due to non-specific cleavage.

Table 1- MALDI-TOF Results of Expression of MBP-SafL Fusion Proteins

<u>Identity</u>	<u>Calculated MW (g/mol)</u>	<u>Found MW (g/mol)</u>
Saf1L-MBP	58927	58888
Saf2L-MBP	58830	58776
Saf1L	Not Detected	Not Detected
Saf2L	16210	16214

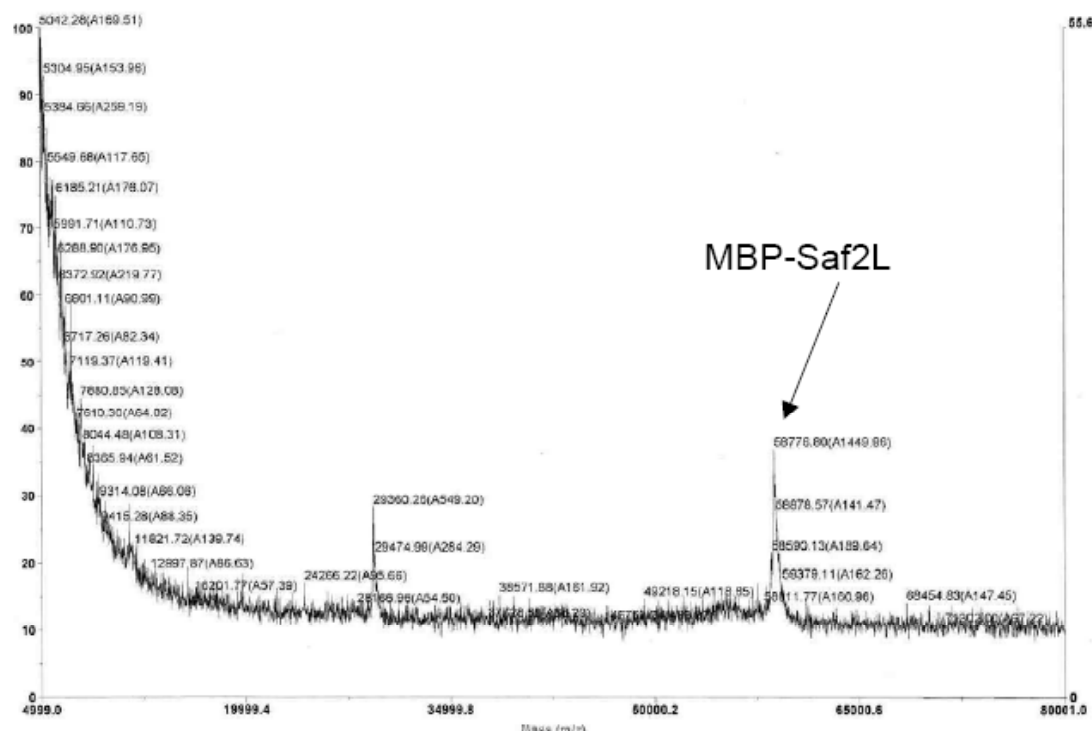


Figure 17- MALDI-TOF spectrum of Saf2L-MBP showing a peak at 58776 Da.

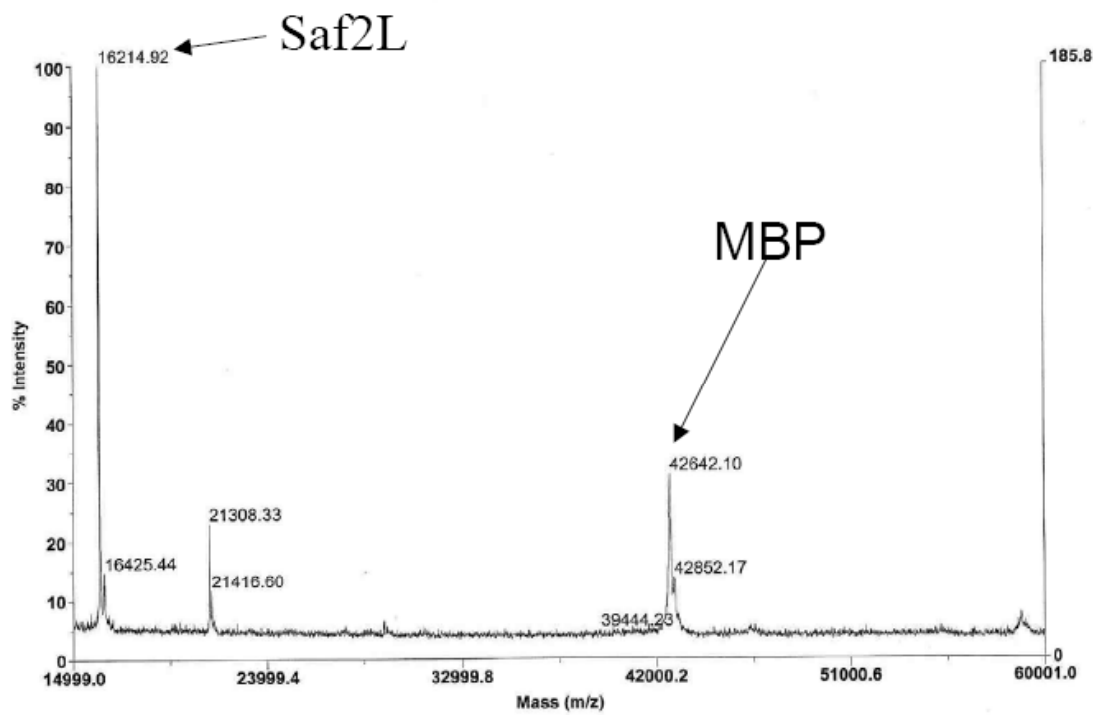


Figure 18- MALDI-TOF spectrum of seven-hour cleavage reaction of Saf2L-MBP.

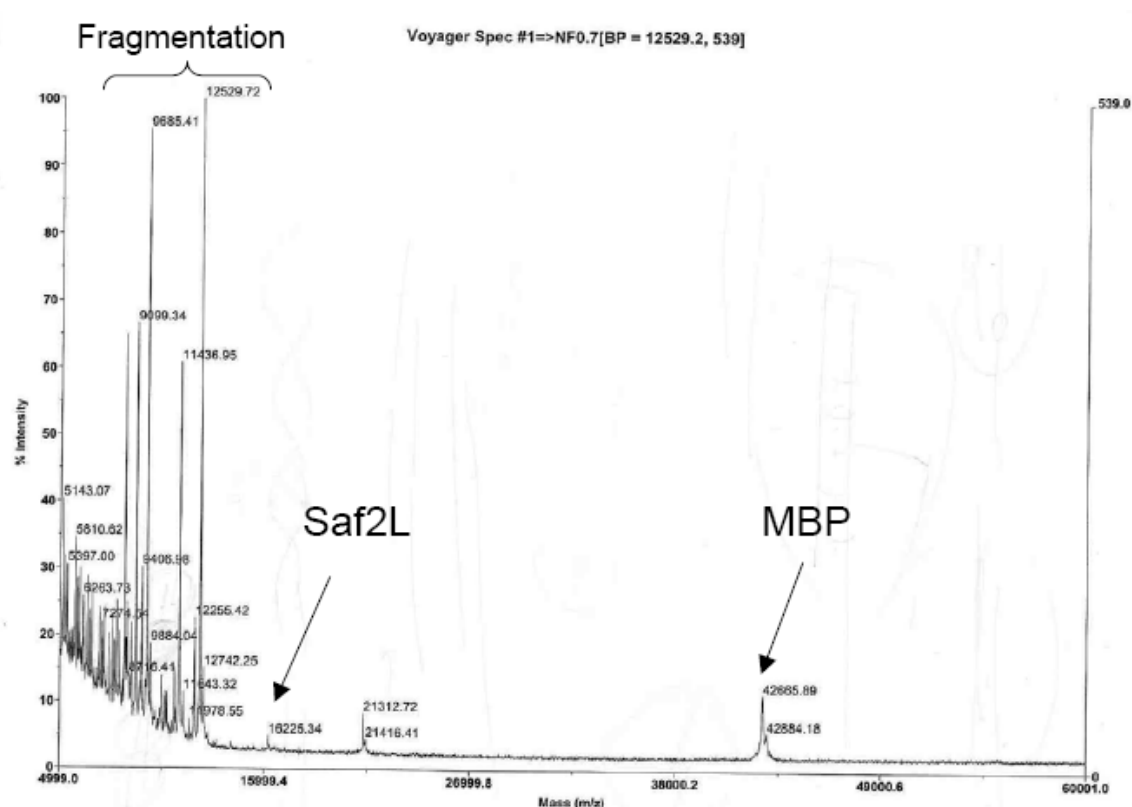


Figure 19- MALDI-TOF spectrum of the same mixture as in Figure 18 with diminished peak at 16.2 kDa and fragmentation peaks in the low molecular weight region.

Expression and Purification of 6x His Tagged Linked Proteins

Polyhistidine tagged linked proteins were expressed at ~4mg/liter of culture- a significantly lower level than their MBP fusion counterparts. Their expression was confirmed by MALDI-TOF (Saf1L: calc: 16112, found: 16187; Saf2L: calc: 16022, found: 16051) (**Figure 20**). However, the purity of the samples was very poor and time did not allow for further purification steps. Mass spectra of cobalt column eluates also revealed the presence of a 10-kDa peak in all samples tested (**Figure 21**). Whether this

peak is due to a histidine rich *E. coli* native protein, a truncated expression product, or the result of protease activity was not determined. Interestingly, both linked proteins were only successfully purified in the native state. Attempts at purification under denaturing conditions (8M Urea) yielded only the 10-kDa peak

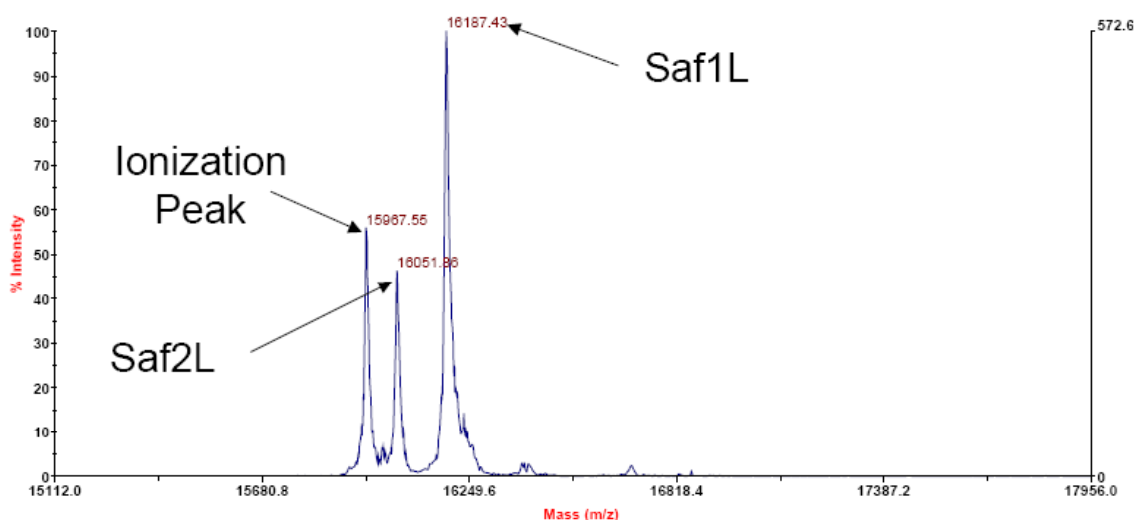


Figure 20- Co-purified samples of his-tagged Saf1L and Saf2L to determine that the linked proteins could be resolved as separate peaks using MALDI-TOF.

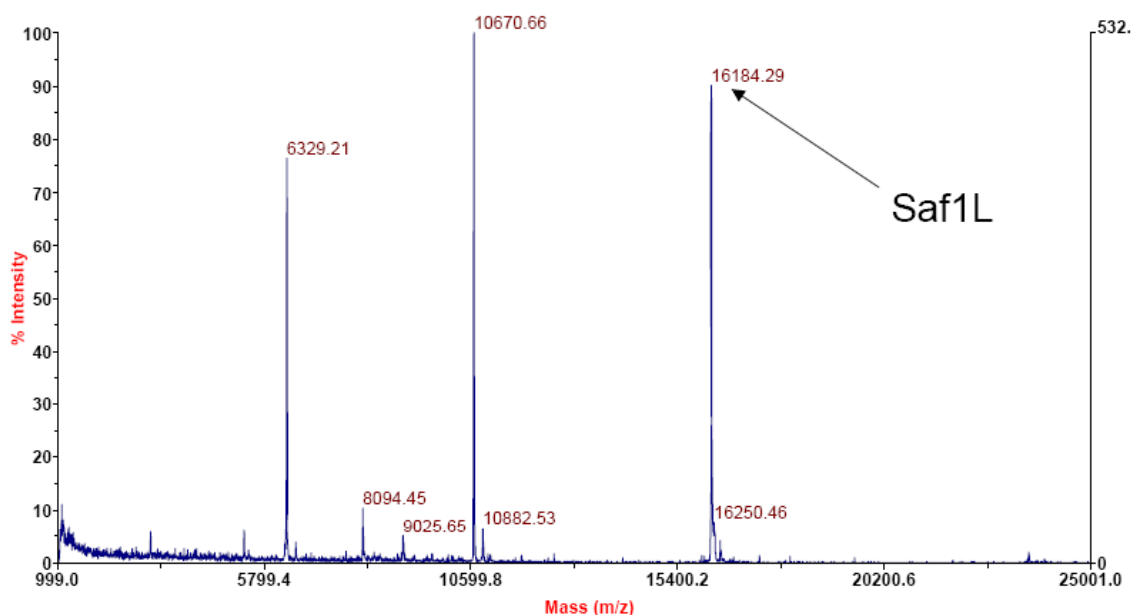


Figure 21- MALDI-TOF spectrum demonstrating the presence of several impurities, including a peak at 10670 Da. This peak was observed in all samples tested.

The peak at 10670 was analyzed using the program Aldente [36]. This program analyzes protein sequences for common serine-protease induced cleavage fragments and identifies fragments based on mass spectra peaks.

Hydrogel Formation: Determination of Gelation Temperature

Falling Ball

Hydrogel formation was tested under conditions of high protein concentration (20 $\mu\text{g}/\mu\text{L}$) and low ionic strength (MOPS) (**Figure 22**). In order for hydrogel formation to proceed, the solution of proteins had to be heated. We found that all four peptides were

necessary for hydrogel formation. This was surprising because, theoretically, the unlinked proteins alone should be able to form a hydrogel due to the presence of the helical endblocks. However, this was found not to be the case as mixtures of the linked peptides alone at 20 $\mu\text{g}/\mu\text{L}$ were unable to gel.

Gelation point was extremely variable, with some solutions from the same batch found to gel at relatively low temperatures (between 35-45 $^{\circ}\text{C}$) and others between 65-80 $^{\circ}\text{C}$ (**Figure 23**). In all cases, the gelation process was composed of appearance of a viscous liquid followed rapidly (<3sec) by complete gelation. Gels formed from the mixture of peptides were clear and, once formed, stable for several weeks at room temperature or 37 $^{\circ}\text{C}$.



Figure 22- An inverted 500uL eppendorf tube of a hydrogel formed by mixing the four protein subunits.

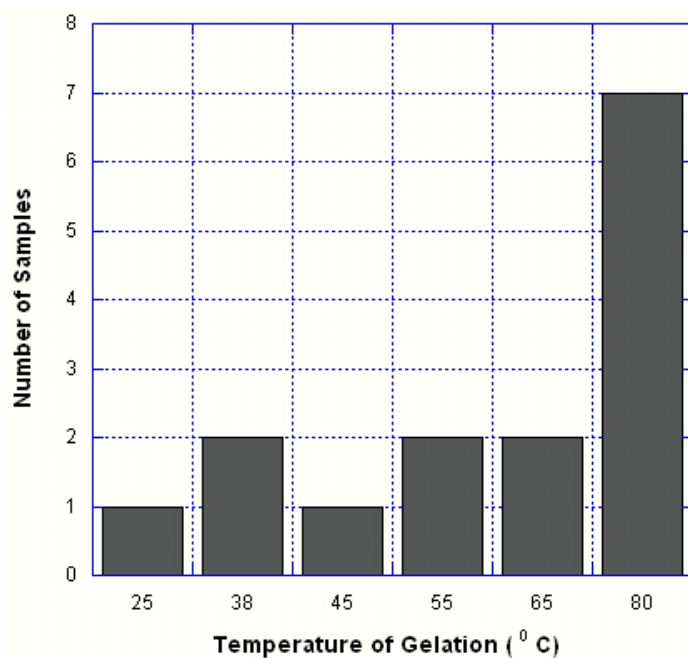


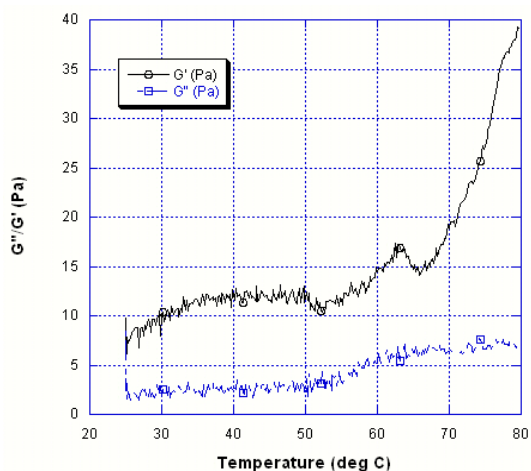
Figure 23- The distribution of gelation temperatures obtained by falling ball experiments was variable.

Rheometry

The exact point of gelation could not be assessed using microrheometer as the temperatures higher than 80°C would have caused too much water loss and significantly altered the results of the subsequent frequency sweep tests. Nevertheless, there is an abrupt rise in the storage modulus of both the G1 and G2 peptide mixtures beginning ~60°C and proceeded without plateau until the last measurement at 80°C (**Figure 24**).

Only one measurement could be made for each hydrogel, thus comparisons between the two samples are difficult to make.

A)



B)

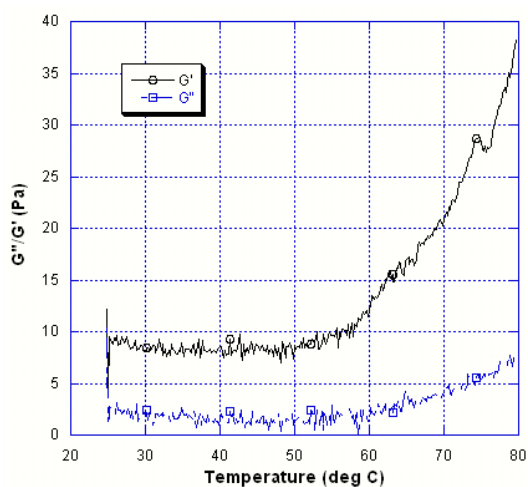
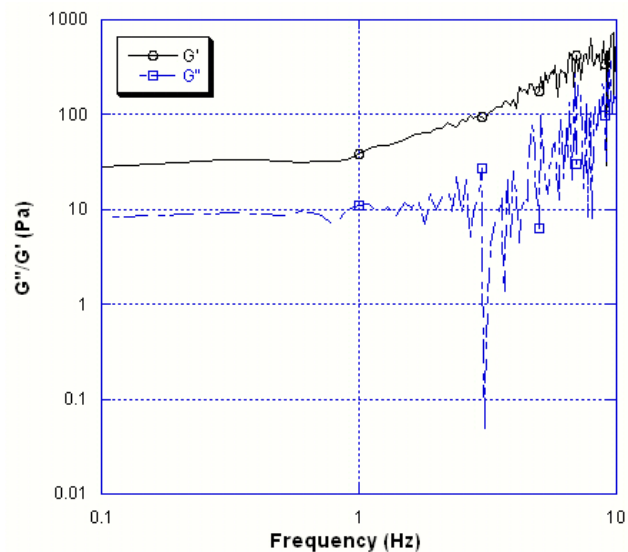


Figure 24- Parallel Plate Rheometry data with temperature sweep showing gelation behavior of (A) Generation 1 mixture and (B) Generation 2 mixture.

The variation of G' and G'' with frequency for the G1 and G2 hydrogel mixtures is presented in **Figure 25**. In both mixtures, G' and G'' are relatively flat until 1Hz is

reached. At that point both G' and G'' exhibit a modest rise with increasing frequency. The G'' plots in the measured frequency range were consistently lower than the G' plots. However, there are significant fluctuations in the G'' data for both hydrogel mixtures after approximately 4Hz was reached possibly indicating hydrogels lost their physical integrity.

A)



B)

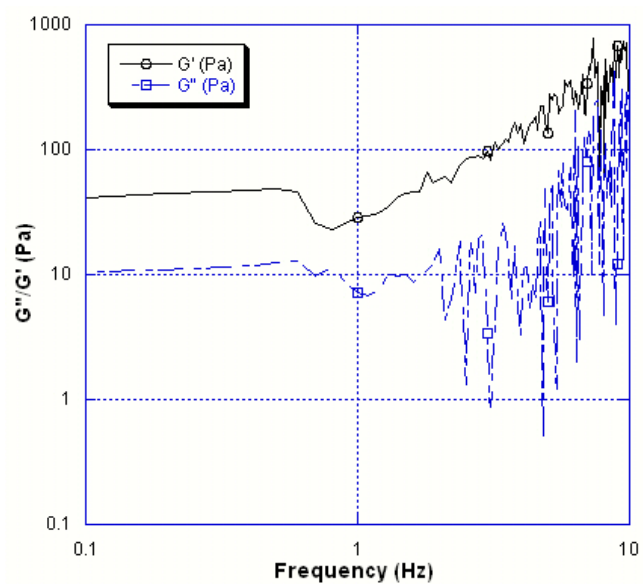


Figure 25- Parallel Plate Rheometry data with frequency sweep showing viscoelastic behavior of gels formed from (A) Generation 1 Mixture and (B) Generation 2 mixture.

Discussion

Expression of Linked Proteins

Not every protein can be expressed/over-expressed in *E. coli*. Of the 58,806 known targets cloned, only 17,350 have actually been purified (~30%) [32]. Furthermore, since each protein is unique, no a priori assumptions can be made as to which expression vectors, expression strategies, and purification protocols will be successful in producing the protein of interest. Thus, refining recombinant protein expression and purification in *E. coli* rests in large part on empirical observations and trial and error. This underscores the necessity of using multiple approaches when expressing heterologous proteins in *E. coli*, or any other expression host for that matter. This is particularly true when the target protein is not naturally occurring (i.e. the product of an artificial gene).

The [(AG)₃PEG]₁₀ hydrophilic linker has been expressed in *E. coli* as described in Petka et al.[13]. In their study, the [(AG)₃PEG]₁₀ region was expressed as a 6x His tagged protein. The investigators reported expression of 26 mg/liter of growth medium, much higher than what we obtained. However, compared to other proteins expressed in their study it was indeed poorly expressed. In fact, the much smaller ~56 amino acid helical region was expressed at approximately 5 times that of the linker region. Thus, significantly many more moles of the helical region were expressed than the linker region alone. The most likely cause of this was the repetitive sequence of nucleotides that encode the linker region. Repetitive DNA sequences are infamous for being unstable and prone to rearrangement/deletion in *E. coli* [37]. Furthermore, these repetitive sequences foster the formation of RNA with secondary structures that are incompatible with

efficient ribosomal binding and thus diminish expression of the heterologous construct [38].

Neither one of our expression methods yielded as much linked protein as was achieved by Petka et al. The fusion protein method would have yielded only 50% and the 6x His tagged protein approximately seven fold less per liter of culture than the yield reported by those investigators. It is difficult to assess the reason for why this is so, especially in light of removing the repetitive sequences from the linked region of our triblock constructs. It is tempting to surmise that the genes encoding the 6x His tagged linked proteins were poorly expressed. The fact that samples of whole cell lysate from small-scale cultures in both expression systems did not yield enough protein for detection by SDS-PAGE seems to support this (data not shown). However, culture conditions and degree of aeration severely impact expression level and solubility [32]. Small-scale cultures are notoriously poorly aerated and thus false negatives can be expected. Furthermore, MBP fusions were expressed quite well in 1L cultures. Nevertheless, enough linked proteins were eventually purified via IMAC to make a limited number of hydrogels.

MBP-Fusion Proteins

Expression of the linked proteins as MBP fusions initially looked very promising given the yields and purity obtained via amylose chromatography. Unfortunately, it appears as though Genenase I-mediated proteolysis quickly digested the linked proteins after cleaving at the proper sequence preceding the N-termini. This is not a new phenomenon as sequence dependent proteases may cleave at non-canonical sites in

certain proteins [35]. This can be very difficult to predict in the case of novel proteins due to their absence from databases containing the proteolytic activity of enzymes such as Genenase I.

Given the initial success of the fusion protein strategy and the low yield obtained using the 6x His tagged proteins, it would seem desirable to express the linked proteins as another type of fusion protein. One candidate would be expression as a small ubiquitin modifying protein (SUMO) fusion. The advantage of using SUMO is that it has properties similar to MBP such as enhancing solubility and refolding of recombinant proteins [30]. Its distinct advantage, however, is that cleavage of the SUMO purification tag via the SUMO protease is based on its tertiary structure rather than dependent on an amino acid sequence. Thus, target proteins are not cleaved by SUMO protease.

6x His Tagged Proteins

While the strategy for expressing the 6x His tagged proteins was successful, the purity was suboptimal. We estimate purity in the region of 50-60% as judged from mass spectrometry. MALDI consistently showed a peak at approximately 10.67 kDa in both Saf1L and Saf2L preparations. When recombinant proteins are expressed at low levels, such as the case here, endogenous histidine-rich *E. coli* proteins can bind to and subsequently elute from metal affinity columns [39]. However, it is more likely to be the result of serine-protease activity within *E. coli* itself. When the amount of crude extract loaded onto the Co^{2+} column was doubled, the peak at 10.67 kDa did not diminish. This indicates that a contaminating *E. coli* protein is unlikely to be that peaks identity, as the increase in hexa-histidine rich target protein would have displaced at least some of an endogenous protein. While we tried to perform purification under denaturing conditions

(8M urea) and in the presence of protease inhibitors, the strategy was ineffective. It should be noted, however, that some serine-proteases still retain activity under denaturing conditions such as 8M urea and relatively high concentrations of the anionic detergent Sodium-Dodecyl-Sulfate (SDS) [40].

Table 2 below shows two proposed sequences that match the mass spectra peaks at 10.67 kDa within 2 Da. These proposed sequences have the 6x His tag more towards the N-termini and would explain their relatively strong affinity for the metal affinity columns. The full-length protein has the 6x His tag approximately in the middle and would presumably exhibit diminished ability to bind to the IMAC column. Furthermore, their presence could help explain the relatively high protein concentrations needed to form the hydrogels, as these fragments have no endblocks with which to propagate network formation.

Both Saf1L-frag1 and Saf2L-frag1 have pI's of approximately 5.05 while Saf1L has a pI of 4.84 and Saf2L a pI of 6.11. Thus, ion-exchange chromatography could have served as a means of separation of the fragments from the full-length products if time had allowed. However, the Saf1L and Saf1L-frag1 separation would have been difficult without a gradient maker given the pI similarity between the two species. Alternatively, reverse phase HPLC would have been able to resolve the fragments from the full-length products.

Table 2- Proposed identity of 10.67 kDa peaks found in mass spectra.

<u>Identity</u>	<u>MW of</u> <u>Fragment</u>	<u>Sequence of Proposed Fragment</u>
		LEQMRGSHHHHHHGSASYRDPMGAGAGAGPEGAGAGAGPEGA
Saf1L-frag1	10668.24	GAGAGPEGAGAGAGPEGAGAGAGPEGAGAGAGPEGAGAGAGP EGAGAGAGPEGAGAGAGPEGAGAGAGPEGARMPTSWKIAAL IAALEQMRGSHHHHHHGSASYRDPMGAGAGAGPEGAGAGAGP
Saf2L-frag1	10668.24	EGAGAGAGPEGAGAGAGPEGAGAGAGPEGAGAGAGPEGAGAG AGPEGAGAGAGPEGAGAGAGPEGAGAGAGPEGARMPTSWKI

Hydrogel Formation

These studies revealed that formation of hydrogels was possible only by the mixing of all four protein subunits. Any formulation lacking both the linked and unlinked subunits, even at 20 ug/uL, resulted in no hydrogel formation, strongly suggesting that hydrogelation proceeded in accord with our proposed mechanism. There were two interesting features of the hydrogel produced in this study, 1) the protein mixture required (in a majority of cases) heating to relatively high temperatures produce the hydrogel and 2) a relatively high protein concentration was needed to achieve the hydrogel state.

The α -helical secondary structures as well as the coiled-coil structure itself are unusually stable under conditions of high heat. Furthermore, other protein-based hydrogels have been produced that require heating in order to achieve gelation [41].

Studies of fiber formation from isolated peptides indicate that some degree of denaturation of the peptides aids in fibrillogenesis [7]. Lysozyme based hydrogels were formed on heating the protein to 85⁰C [42]. Furthermore, some de novo designed coiled-coils have exhibited stability at temperatures at 80-85⁰C [13, 14, 22]. Others demonstrated stability at temperatures as high as and 108⁰C and hydrogels produced from these peptides stable up to 70⁰C depending on the length of the endblocks employed in the hydrogel design [43].

As discussed above, the presence of the fragments produced from serine protease induced cleavage may well have played a role in the necessity of high protein concentrations needed for hydrogel formation. Because of the presence of impurities, more lyophilized proteins had to be weighed to attain the proper number of moles of full-length product. Other genetically engineered coiled-coil triblock hydrogels formed at 0.1-0.3 ug/uL [14, 22], 0.08 ug/uL [17] and 0.04 ug/uL [13].

A falling ruby ball to test gelation points of peptide mixtures was done to determine gelation as a function of temperature. This method was not meant to yield an exact measure of gel-sol properties, only an estimate of the gelation point. As can be seen in **Figure 23**, there was considerable variation in this property of the gels. This variation was observed even within the same batch of peptides. Two possible explanations include the rate that the mixtures were heated and the presence of impurities in both the linked and unlinked proteins. Peptide mixtures were heated in a water bath with heating controlled by a hotplate. It was difficult to achieve precise control of the rate of heating with this experimental set-up. Heating rate was found to effect both temperature of gelation as well as storage modulus in hydrogels composed of PLLA-PEG-PLLA triblock

[44] and di-acrylated pluronic F-27 [45]. This may be due, in part, to the fact that network formation in physical hydrogels is a nucleation-controlled process [46]. Rates of nucleation can often directly correspond with rates of temperature change.

The unlinked proteins were at ~70% purity and the linked at approximately 50-65% purity. The presence of impurities on gelation temperature of hydrogels is not well documented and thus is unclear how their presence affected this property. However, the impurities in the unlinked proteins are likely to be truncated versions of the full-length, 28 amino-acid peptides due to the way peptides are synthesized. If enough amino-acids are in place on the N-terminal end and are lacking towards the C-terminal end, binding could occur to the endblock but disrupt its growth because there are no hydrophobic amino-acids to propagate “sticky-end” fibrillogenesis. This would be roughly equivalent to disrupting seed formation.

Rheology

As can be seen in **Figure 24**, both the generation 1 and 2 peptide mixtures exhibit a sharp increase in the storage modulus at 65 and 60⁰C respectively while the loss moduli stay almost constant or increase only modestly. In each case, G' is greater than G'' which indicates these materials are behaving as solid-like gels [47]. Note that for the fluid state, G'' will normally be greater than G' . However, for protein solutions this is not always the case and thus the G' , G'' relationship in the low temperature region [48].

We were unable to allow measurements to be taken after 80⁰C and thus were not able to allow the G' to achieve the plateau state. It is possible that some of the increase in the storage modulus is from loss of water at the relatively high temperatures. However, after rheological tests were done, the gels still seemed quite hydrated.

Data from oscillatory studies on viscoelastic materials can be interpreted by modeling the material as a number of elastic springs in series and/or parallel with a number viscous dashpots or pistons [49]. At low frequencies the springs are able to elongate with simultaneous extension of the dashpots to a degree that exceed those of the springs. When the deformation of the springs reaches a point of equilibrium and their deformation ceases, the dashpots will deform continuously under the oscillatory shear. It is during this phase that energy is dissipated and the sample behaves as a liquid. At the higher frequencies, the situation is reversed: the springs can elongate under the oscillatory shear but the dashpots have no time to react to the high frequency. During this phase, energy is stored and the material will behave as an elastic solid. At intermediate frequencies, both spring and dashpot elements will contribute and viscoelastic behavior is observed.

Frequency sweep data suggests that there was formation of a fairly extensive network, presumably mediated by protein-protein interactions. As can be seen in **Figure 25**, there is no increase with oscillatory shear until approximately 1 Hz, at which point both the storage and loss moduli increase with a gradual slope. This again demonstrates the hydrogels viscoelastic nature. This is to be expected from a hydrogel material consisting of both water and a continuous network of molecular associations such as outlined in **Figure 7**. Furthermore, since the G'' plots are consistently lower than G' plots, the hydrogel exhibits a predominant elastic character since there is a smaller dissipation energy.

Conclusion

We have fabricated a hydrogel by the mixing of four protein subunits. Hydrogel formation was observed only when all four subunits were mixed at a concentration of 20 ug/uL. It is likely that this high concentration of protein was necessary due to the presence of truncated protein products produced by the action of serine proteases. The result of protease cleavage produced fragments that would not incorporate into the hydrogel network but could still be purified by IMAC chromatography due to the presence of the 6x His tag in these truncated products. Nevertheless, using a high concentration of subunits could overcome their presence and network formation was evidenced by rheological studies. The expression of His tagged proteins was very low, yet their MBP fusion protein counterparts were expressed at very high levels. However, MBP fusion constructs were degraded by the protease that cleaved the linked proteins from the MBP tag. Proteins expressed as fusion constructs often do not fit conveniently into protease active sites. Therefore, the best way to express the linked proteins would be to express them as fusion proteins with the small-ubiquitin-modifying protein (SUMO). This system allows the expression of the linked proteins as fusion products but obviates their proteolytic cleavage due to the fact that the SUMO protease recognizes the tertiary structure of SUMO rather than a small sequence of amino acids. This circumstance makes it more likely that the linked proteins could be expressed at high levels without protease cleavage and thus increases their final yield.

References

1. Rao, K.P., Recent developments of collagen-based materials for medical applications and drug delivery systems. *J Biomater Sci Polym Ed*, 1995. **7**(7): p. 623-45.
2. Yu, X., G.P. Dillon, and R.B. Bellamkonda, A laminin and nerve growth factor-laden three-dimensional scaffold for enhanced neurite extension. *Tissue Eng*, 1999. **5**(4): p. 291-304.
3. Park, K.H. and K. Na, Effect of growth factors on chondrogenic differentiation of rabbit mesenchymal cells embedded in injectable hydrogels. *J Biosci Bioeng*, 2008. **106**(1): p. 74-9.
4. Wang, S., et al., Three-dimensional primary hepatocyte culture in synthetic self-assembling peptide hydrogel. *Tissue Eng Part A*, 2008. **14**(2): p. 227-36.
5. Peppas, N.A., et al., Hydrogels in pharmaceutical formulations. *Eur J Pharm Biopharm*, 2000. **50**(1): p. 27-46.
6. Tang, A., et al., The coiled coils in the design of protein-based constructs: hybrid hydrogels and epitope displays. *J Control Release*, 2001. **72**(1-3): p. 57-70.
7. Aggeli, A., et al., Exploiting Peptide Self-assembly to Engineer Novel Biopolymers: Tapes, Ribbons, Fibrils and Fibres, in *Self-Assembling Peptide Systems in Biology, Medicine and Engineering*. 2002. p. 1-17.
8. Mahmoud, R., et al., In vivo evaluation of whey protein-based biofilms as scaffolds for cutaneous cell cultures and biomedical applications. *Biomedical Materials*, 2007(1): p. S38.
9. Hwang, D.-C. and S. Damodaran, Synthesis and properties of fish protein-based hydrogel. *Journal of the American Oil Chemists' Society*, 1997. **74**(9): p. 1165-1171.
10. Stegemann, J.P., S.N. Kaszuba, and S.L. Rowe, Review: advances in vascular tissue engineering using protein-based biomaterials. *Tissue Eng*, 2007. **13**(11): p. 2601-13.
11. McGrath, K.P., et al., Chemical and biosynthetic approaches to the production of novel polypeptide materials. *Biotechnol Prog*, 1990. **6**(3): p. 188-92.
12. Cappello, J., et al., Genetic engineering of structural protein polymers. *Biotechnol Prog*, 1990. **6**(3): p. 198-202.
13. Petka, W.A., et al., Reversible hydrogels from self-assembling artificial proteins. *Science*, 1998. **281**(5375): p. 389-92.
14. Xu, C., V. Breedveld, and J. Kopecek, Reversible hydrogels from self-assembling genetically engineered protein block copolymers. *Biomacromolecules*, 2005. **6**(3): p. 1739-49.
15. Shen, W., et al., Tuning the erosion rate of artificial protein hydrogels through control of network topology. *Nat Mater*, 2006. **5**(2): p. 153-8.
16. Wright, E.R. and V.P. Conticello, Self-assembly of block copolymers derived from elastin-mimetic polypeptide sequences. *Adv Drug Deliv Rev*, 2002. **54**(8): p. 1057-73.
17. Jing, P., et al., Self-Assembling Peptide-Polymer Hydrogels Designed From the Coiled Coil Region of Fibrin. *Biomacromolecules*, 2008. **9**(9): p. 2438-2446.

18. Yu, Y.B., Coiled-coils: stability, specificity, and drug delivery potential. *Adv Drug Deliv Rev*, 2002. **54**(8): p. 1113-29.
19. Kopecek, J., Hydrogel biomaterials: A smart future? *Biomaterials*, 2007. **28**(34): p. 5185-5192.
20. Lao, U.L., et al., Genetic Engineering of Self-Assembled Protein Hydrogel Based on Elastin-like Sequences with Metal Binding Functionality. *Biomacromolecules*, 2007. **8**(12): p. 3736-3739.
21. Graddis, T.J., D.G. Myszka, and I.M. Chaiken, Controlled formation of model homo- and heterodimer coiled coil polypeptides. *Biochemistry*, 1993. **32**(47): p. 12664-71.
22. Xu, C. and J. Kopecek, Genetically engineered block copolymers: influence of the length and structure of the coiled-coil blocks on hydrogel self-assembly. *Pharm Res*, 2008. **25**(3): p. 674-82.
23. Woolfson, D.N. and M.G. Ryadnov, Peptide-based fibrous biomaterials: Some things old, new and borrowed. *Curr Opin Chem Biol*, 2006. **10**(6): p. 559-67.
24. AM Smith, E.B., WR Edwards, MJ Pandya, DN Woolfson, Engineering Increased Stability into Self-Assembled Protein Fibers. *Advanced Functional Materials*, 2006. **16**: p. 1022-30.
25. Andrew M. Smith, E.F.B., Wayne R. Edwards, Maya J. Pandya, Derek N. Woolfson, Engineering Increased Stability into Self-Assembled Protein Fibers. *Advanced Functional Materials*, 2006. **0000**(00): p. 1-9.
26. Papapostolou, D., et al., Engineering nanoscale order into a designed protein fiber. *Proceedings of the National Academy of Sciences*, 2007. **104**(26): p. 10853-10858.
27. Papapostolou, D., et al., Electrostatic Control of Thickness and Stiffness in a Designed Protein Fiber. *J. Am. Chem. Soc.*, 2008. **130**(15): p. 5124-5130.
28. Pandya, M.J., et al., Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis. *Biochemistry*, 2000. **39**(30): p. 8728-34.
29. Stevens, M.M., et al., Molecular level investigations of the inter- and intramolecular interactions of pH-responsive artificial triblock proteins. *Biomacromolecules*, 2005. **6**(3): p. 1266-71.
30. Arnau, J., et al., Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expression and Purification*, 2006. **48**(1): p. 1-13.
31. Nallamsetty, S. and D.S. Waugh, Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. *Protein Expression and Purification*, 2006. **45**(1): p. 175-182.
32. Graslund, S., et al., Protein production and purification. *Nat Methods*, 2008. **5**(2): p. 135-46.
33. Hartley, J.L., Cloning technologies for protein expression and purification. *Current Opinion in Biotechnology*, 2006. **17**(4): p. 359-366.
34. David, C.L., Y. Bernard, and A.L. Noshir, Mechanical Properties of a Reversible, DNA-Crosslinked Polyacrylamide Hydrogel. *Journal of Biomechanical Engineering*, 2004. **126**(1): p. 104-110.
35. Biolabs, N.E., pMal Protein Fusion and Purification System. Catalog#E800S, 2007.

36. Expasy, ALDENTE : PEPTIDE MASS FINGERPRINTING TOOL.
37. Bzymek, M. and S.T. Lovett, Instability of repetitive DNA sequences: The role of replication in multiple mechanisms. *Proceedings of the National Academy of Sciences of the United States of America*, 2001. **98**(15): p. 8319-8325.
38. Care, S., et al., The translation of recombinant proteins in *E. coli* can be improved by in silico generating and screening random libraries of a -70/+96 mRNA region with respect to the translation initiation codon. *Nucleic Acids Res.*, 2008. **36**(1): p. e6-.
39. Bolanos-Garcia, V.M. and O.R. Davies, Structural analysis and classification of native proteins from *E. coli* commonly co-purified by immobilised metal affinity chromatography. *Biochim Biophys Acta*, 2006. **1760**(9): p. 1304-13.
40. Blumentals, II, A.S. Robinson, and R.M. Kelly, Characterization of sodium dodecyl sulfate-resistant proteolytic activity in the hyperthermophilic archaeobacterium *Pyrococcus furiosus*. *Appl Environ Microbiol*, 1990. **56**(7): p. 1992-8.
41. Yin Liang Hsien, J.M.R., Modeling Gelation of Egg Albumen and Ovalbumin. *Journal of Food Science*, 1992. **57**(4): p. 856-861.
42. Yan, H., et al., Thermoreversible protein hydrogel as cell scaffold. *Biomacromolecules*, 2006. **7**(10): p. 2776-82.
43. Wang, C., R.J. Stewart, and J. Kopecek, Hybrid hydrogels assembled from synthetic polymers and coiled-coil protein domains. *Nature*, 1999. **397**(6718): p. 417-20.
44. Tomoko Fujiwara, T.M.T.Y.H.Y.S.S.Y.K., Novel Thermo-Responsive Formation of a Hydrogel by Stereo-Complexation between PLLA-PEG-PLLA and PDLA-PEG-PDLA Block Copolymers. *Macromolecular Bioscience*, 2001. **1**(5): p. 204-208.
45. Lee, S.Y., G. Tae, and Y.H. Kim, Thermal gelation and photo-polymerization of di-acrylated Pluronic F 127. *J Biomater Sci Polym Ed*, 2007. **18**(10): p. 1335-53.
46. Xiong, J.Y., et al., Understanding of hydrogel network formation and its application in the architecture of significantly enhanced hydrogel. *Applied Physics Letters*, 2006. **89**(8): p. 083106.
47. Eun S. Gil, R.J.S.S.M.H., Effect of β -Sheet Crystals on the Thermal and Rheological Behavior of Protein-Based Hydrogels Derived from Gelatin and Silk Fibroin. *Macromolecular Bioscience*, 2005. **5**(8): p. 702-709.
48. Gosal, W.S. and S.B. Ross-Murphy, Globular protein gelation. *Current Opinion in Colloid & Interface Science*, 2000. **5**(3-4): p. 188-194.
49. Rudraraju, V.S. and C.M. Wyandt, Rheology of Microcrystalline Cellulose and Sodiumcarboxymethyl Cellulose hydrogels using a controlled stress rheometer: part II. *International Journal of Pharmaceutics*, 2005. **292**(1-2): p. 63-73.

Chapter 3- Molecular Modeling and MD Simulation of Early Oligomeric Intermediates Formed During Fibrillogenesis of Saf Peptides

Introduction

Peptides are finding wide use as materials for fibrous, tissue engineering substrates [1]. One distinct advantage of using peptides for such purposes is that they help replicate the in vivo environment that cells grow in. This is particularly true of the extracellular matrix that is primarily composed of the protein. As materials, peptides offer distinct advantages. Any variety of short amino acid sequences can be rapidly synthesized using modern, solid state technology. Furthermore, some peptides have the ability to spontaneously adopt well-studied configurations such as β - sheets or α - helices. In nature, it is these structures that coalesce into extensively structured functional entities.

Recently, attention has been turned towards exploiting these elements of secondary structure for potential use in fibrous biomaterials. Some of these include β -sheet moieties such as the RAD 16 peptides [2, 3] as well as 28 [4], 34 [5] and 35-mer [6] α -helices that aggregate into long coiled-coil fibers. It would be of tremendous advantage to develop methods to aid in rational modification of these peptide subunits such that their aggregation properties could be modulated in a controlled fashion. In doing so, peptides could be optimized to aggregate under specific conditions of pH and

temperature. It may eventually be possible to control fibers' physical properties such as diameter, critical concentration for assembly, and kinetics of aggregation.

One impediment to such rational design, however, is the lack of thermodynamic and kinetic data to help describe the aggregating system of peptides. Without such data, it is difficult to propose and test fiber assembly mechanisms or identify key structural intermediates that arise early on during aggregation. Such data could be used to identify points of control and potentially exploit these structures to affect downstream aggregation processes and thus the fiber properties mentioned above.

Formation of fibrils by aggregating protein monomers generally proceeds after monomer concentrations reach some critical value. There are two main steps: 1) nucleation/Oligomerization and 2) Mature fiber formation [7]. This is based on the observation that there is usually a lag phase between addition of monomers and fiber formation, where monomers assemble into prefibrillar oligomers that seed fibrillogenesis. In experiments done with β -amyloid, "seeding" monomer mixtures with oligomeric aggregates eliminated the lag phase [8]. The existence of soluble oligomeric species has been reported in several protein and peptide fibrillogenic systems including intermediate filaments, insulin, prion-like proteins, β -amyloid, and α -synuclein [9-15]. Once seeded, long fibrils grow rapidly and, at some point, begin to associate with each other to form mature fibers of well-defined diameters.

Due to the limitations of current experimental techniques, it is difficult to obtain thermodynamic and kinetic data on early peptide aggregation oligomers. To fill this knowledge gap, investigators are now employing Molecular Dynamics (MD) simulations that are capable of generating thermodynamic data on these early aggregation events of

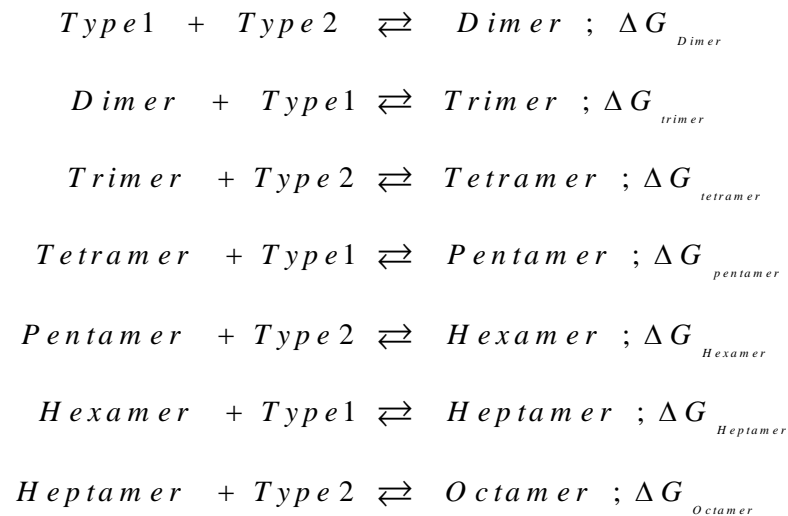
fiber forming peptides. For example, conformational thermodynamics of critical dimeric intermediates arising during the aggregation of the amyloid β ($A\beta$) have been assessed using all-atom MD simulations [16]. To date, however, we know of no MD simulations carried out on potential biomaterial aggregating systems. This may be partly due to the difficulty in generating the initial set of internal coordinates necessary to launch MD. For $A\beta$, and other naturally fiber-forming peptides, the initial set of internal coordinates that describe the position of each atom are documented from NMR studies and their coordinates found in the Protein Data Bank [17]. However, these data are not available for novel, artificial peptides.

In this chapter, we build all-atom MD models of the SAF, alpha helical peptide system developed by Woolfson et al. In this system two different alpha-helical peptides are combined in solution to produce long, coiled-coil fibers as evidenced by linear and circular dichroism [4, 6]. Fiber formation is seeded by a dimeric intermediate in which the complimentary alpha helices are offset such that hydrophobic residues are left exposed to solvent. The consequence of this staggered arrangement is that complimentary monomers can add to these hydrophobic, “sticky ends” and thus propagate fiber formation. We chose this system because 1) We believe it has great potential for biomaterial applications, 2) The constituent monomers and early oligomeric intermediates are small enough to be effectively modeled using MD, 3) The authors provide evidence for an aggregation mechanism that includes a staggered dimer as a seed for aggregation, and 4) monomer amino-acid sequences have been selectively engineered to produce fibers with different physical characteristics. We base the initial internal coordinates on a segment of tropomyosin, a naturally occurring alpha-helical coil-coil.

We use MD simulations coupled with the MM-PBSA post-processing module of AMBER 9 to explore the thermodynamics of some proposed early aggregation structures. We suggest that the methods we outline could be used to propose structures in any aggregating system of peptides. Furthermore we show that MD can be a predictive tool in analyzing the effect of amino acid replacements on oligomer structural stability.

Materials and Methods

We studied the free energy changes accompanying dimerization as well as subsequent monomer additions to several theoretical oligomers that would arise at the initiation of fibrillogenesis in the Saf system. Using MM-PBSA, we explore the ΔG_{bind} for the following reactions:



where Type1 refers to the monomers Saf-p1, Saf-p1-ext, and Type2 refers to Saf-p2, Saf-p2a, and Saf-p2a-ext. Three different generations of fibers were produced from mixing three different combinations of Type1 and Type2 monomers as follows: **Generation 1**

(G1): Saf-p1/Saf-p2, **Generation 2 (G2):** Saf-p1/Saf-p2a, and **Generation 3 (G3):** Saf-p1-ext/Saf-p2-ext. Table 3 gives the amino acid sequences of each monomer.

Table 3- Amino-Acid Sequences of Type1 and Type2 Monomers

Identity	Amino Acid Sequence
Saf-p1 (Type 1)	KIAALKQKIASLKQEIDALEYENDALEQ
Saf-p2 (Type 2)	KIRALKAKNAHLKQEIAALEQEIAALEQ
Saf-p2a (Type 2)	KIRRLKQKNARLKQEIAALEYEIAALEQ
Saf-p1-ext (Type 1)	KIAALKQKIASLKQEIDALEYENDALEQKIAALEQ
Saf-p2-ext (Type 2)	KIRRLKQKNARLKQKIAALEQEIAALEYEIAALEQ

Generation of Molecular Models

A molecular model of each oligomeric species was necessary to provide the initial coordinates for MD production runs. Since the Saf peptides are synthetic in nature, there exist no naturally occurring analogues to them and thus no known crystallized structures. As a consequence, the protein data bank cannot serve as a direct source of these coordinates. To circumvent this potential difficulty, we generated the molecular models via the replacement of side chains on tropomyosin with those of the peptides of the Saf system- a process known as threading [18, 19]. A similar method has been published elsewhere [20]. The coordinates for the crystal structure of tropomyosin at 7 angstrom resolution were obtained from the Protein Data Bank (PDB identifier 1C1G) [17, 21].

Tropomyosin is a protein that exhibits the “leucine zipper” motif and, as such, comprises two α -helical segments that adopt a coiled-coil structure. The structure of α -helices in coiled-coils is such that every seventh residue occupies the same angular position on the helical surface. Thus each 28 aa monomer (G1 and G2) accounts for four complete turns of the coiled-coil left-handed helix. The G3 monomers each have 5 turns. This seven-residue periodicity lends a sound structural foundation for the threading of the Saf monomers- all of which contain a multiple-of-seven number of residues.

The side chains on the A and B chains of tropopmysin were replaced with those from type-1 and type-2 monomers respectively for each generation as indicated in **Table 4**. The Swiss PDB Viewer software was used to replace the side-chains in the original pdb file[22]. The relative locations of these amino-acid substitutions were chosen such that atoms were replaced so the orientation of the R-groups with respect to one another were as outlined by Woolfson et. al. [4] (**Figure 26**).

A “master file” for each coiled-coil generation was made by grafting eight Type1 and eight Type2 repeating monomers onto the A and B chains respectively. This master file served two purposes: 1) A “minifiber” that served as an approximation of a coiled-coil of infinite length composed of each generation's monomeric units and 2) A file of eight octameric monomers could be edited to yield oligomers of the lengths (i.e.number of monomeric units) that were used in the simulations. In this latter case, TER cards (non-covalently bonded C-termini) were placed at the C-terminus of each monomer in the oligomer to simulate non-covalent formation of the nascent coiled-coil.

Table 4- Side-Chain Placement on Crystal Structure of Tropomyosin (PDB Identifier 1C1G)

GENERATION	REPEATING MONOMER	SEQUENCE	AA REPLACED
1	Saf1	K IAALKQK IASLKQE IDALEYE NDALEQ	Chain A: 3-226
	Saf2	K IRALKAK NAHLKQE IAALQEQ IAALQEQ	Chain B: 301-524
2	Saf1	K IAALKQK IASLKQE IDALEYE NDALEQ	Chain A: 3-226
	P2a	K IRRLKQK NARLKQE IAALQEQ IAALQEQ	Chain B: 301-524
3	P1-ext	K IAALKQK IASLKQE IDALEYE NDALEQK IAALQEQ	Chain A: 3-282
	P2a-ext	K IRRLKQK NARLKQK IAALQEQ IAALQEQ IAALQEQ	Chain B: 301-580

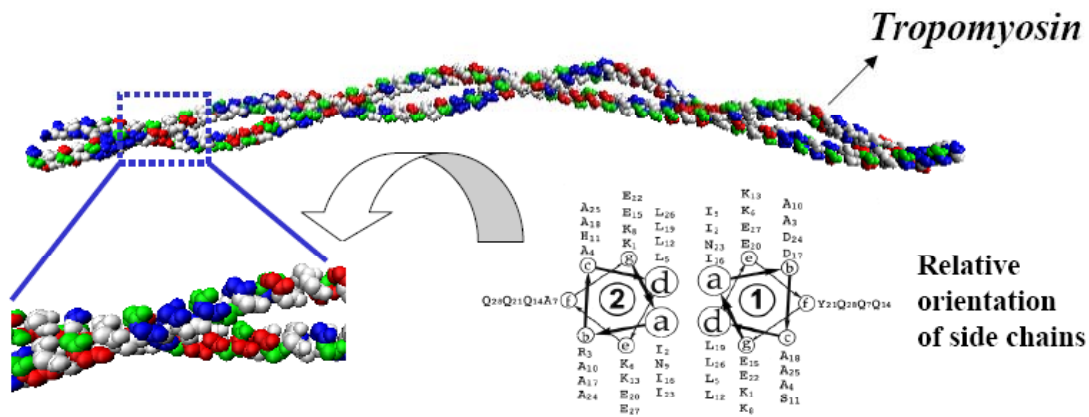
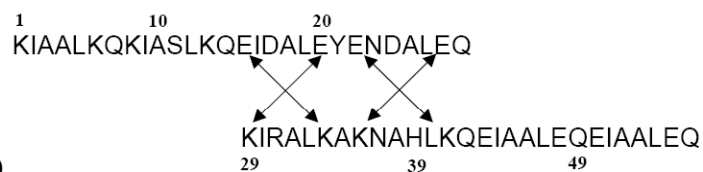


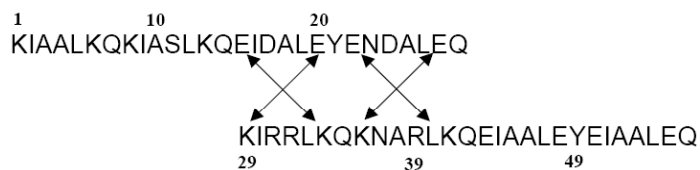
Figure 26-Side chain locations of the G1, G2 and G3 monomers were threaded onto the structure of tropomyosin to match those in the helical wheel diagram specified by Woolfson et al..

Four sets of oppositely charged glutamate-lysine ordered pairs help stabilize the staggered dimeric structure of the G1 and G2 dimers. The G3 dimer has an additional 2 sets (**Figure 27c**). Due to the periodicity of the coiled-coil structure, each of these ordered pairs were repeated along its length every 28 (G1 and G2) and 35 (G3) amino acids. This periodicity was accounted for in the minifiber by assigning each ordered pair to a repeating block (**Figure 28**).

A)



B)



C)

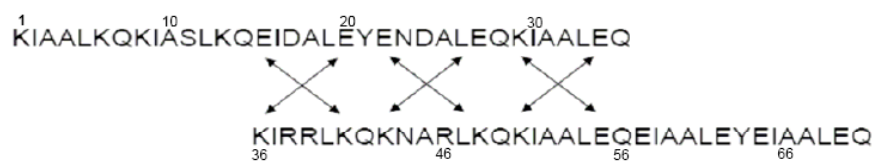


Figure 27- Glutamate-Lysine ordered pairs help stabilize the staggered dimeric structure: (A) and (B) E15-K34, E20-K29, E22-K41, E27-K36. (B) E15-K41, E20-K36, E22-K48, E27-K43, K29-E55, E35-K50.

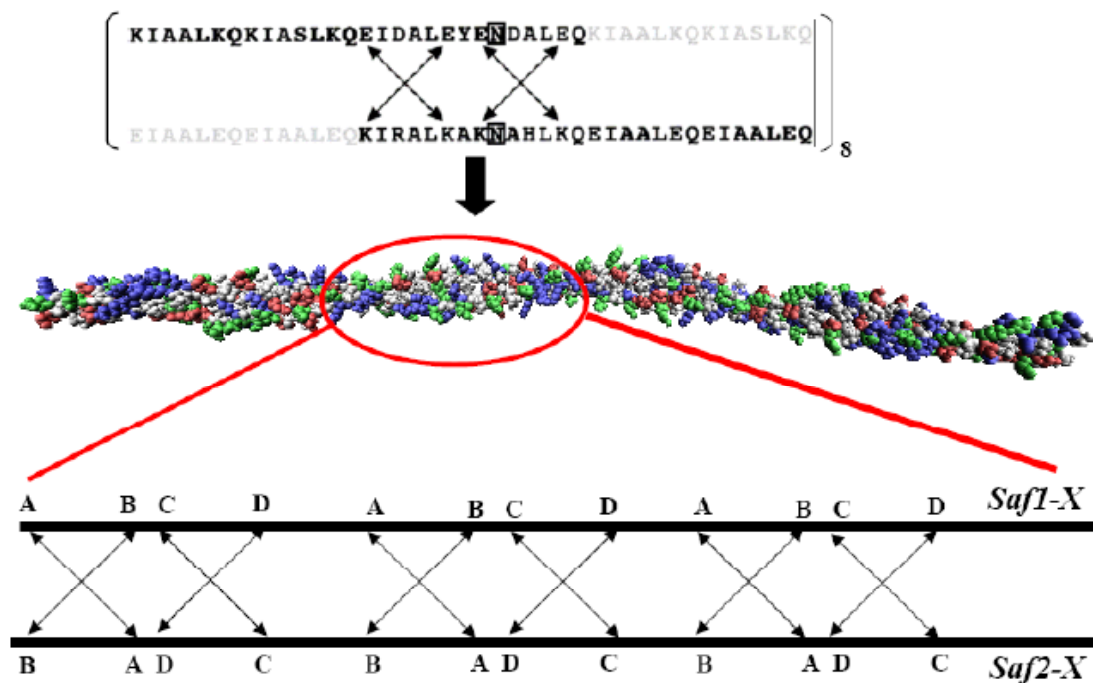


Figure 28- Eight monomer units were grafted onto tropomyosin to simulate a small fiber ('minifiber') comprised of many such repeating units. Glu-Lys ordered pairs are divide into 4 repeating blocks based on their periodicity along the coiled-coil structure.

Molecular Dynamics

The Amber 9 software package was used for all minimization steps, MD production runs and subsequent MM-PBSA post-processing. Each oligomeric intermediate was neutralized with the appropriate counter-ion, when necessary, and subsequent topology and coordinate files made for each solvated system using the xLeap (ff99 parameter set) module of Amber 9. Models were placed at the center of a rectangular simulation box and surrounded with an 8.5 nm layer of water on all sides.

The TIP3P water model explicitly represented the solvent. The Sander module of AMBER 9 was used for energy minimization to relieve unfavorable clashes in the modified crystal structure of the tropomyosin. All production simulations were performed at 300K. Particle Mesh Ewald [23] method was used to treat the long range electrostatic interactions. SHAKE [24] was used to constrain all bonds connecting to hydrogen atoms (time step of 2 fs). Temperature was controlled at 300 K using Berendsen's algorithm [25] with a coupling constant of 2 ps.

Equilibration of the solvated complex was initiated by minimization of the water-oligomer system in which the oligomer was restrained. Equilibration MD runs were then performed via 50ps of heating followed by 50 ps of density equilibration with weak restraints on the oligomer followed by 500 ps of constant pressure equilibration at 300K. Production runs were carried out for 2ns during which a snapshot of each oligomer was obtained every picosecond. These snapshots were then used by the MM-PBSA module of Amber 9 to estimate the free energy of binding of a type-1 or type-2 monomer to oligomers sizes ranging from dimer to octamer (See **Table 3**). Note that only a single trajectory was used for the MM-PBSA analysis. Due to the results obtained in this phase of the simulations, two additional production runs were performed to simulate protofibril addition. Specifically, the joining of two G2 pentamers to make a G2 decamer and two G3 tetramers to produce a G3 octamer.

Since the amino acid side chains of the Saf peptides were superimposed on an existing structure (tropomyosin), the structural integrity of each staggered dimer was evaluated under simulation conditions to demonstrate that each of the dimers remained intact over a comparatively long simulation time. Since the ordered pairs mentioned

above are critical for the maintenance of the staggered dimer orientation, the distance between these pairs was measured during the course of a 10ns production simulation. These distances were compared to the same distances that would exist in a small fiber made of eight repeating units of the monomers (denoted a ‘minifiber’) (**Figure 28**). The advantage of using a minifiber for an ordered-pair control is that each ordered pair is repeated periodically down the length of the coiled-coil, thus a cumulative average of such distances can be calculated. Furthermore, the minifiber should represent a stable structure due to the cumulative, attractive van-der Waals, ionic and hydrophobic forces present along the length of the coiled-coil.

Ordered pair distances were calculated by measuring the interhelical distance between the alpha-carbons in the pair. The rms deviation of the C-C-N backbone from the energy-minimized structure (structure of oligomer in the first frame of production runs) of each oligomeric species was calculated using the Ptraj module of Amber 9. The average secondary structure for each oligomeric intermediate was obtained from the average structure as calculated using the runningaverage command in Ptraj. This structure was then used as the basis to calculate the secondary structure using the secstruc command in Ptraj.

Results

Dimer Stability

Several inter-helical, α -carbon atom-pair distances were measured in the G1, G2 and G3 dimers. Compared to the higher order oligomers, structural stability of the dimers was a concern. Since each dimer was placed into the staggered conformation, only

approximately half of each monomer overlaps with its counterpart. In contrast to other naturally occurring, blunt-ended dimers such as GCN4, this arrangement confers much less potential stability. Measurement of glutamate-lysine ordered pairs was chosen since these atom pairs are believed to be critical for orienting the dimer the staggered configuration [4]. These interhelical positions were compared to those in the mini-fiber controls. These controls comprised eight monomeric type I and type II peptides units threaded onto chains A and B of the crystal structure of tropomyosin. With the exception of the interhelical pairs coinciding with blocks B and E in the generation 3 minifibers, interhelical distances in the 10ns dimer simulations did not differ significantly from those in the corresponding minifibers (**Figures 29 and 30**). For the G1 and G2 dimers, these values lie between 10 and 15 angstroms. Studies of the crystal structure of tropomyosin have shown average interhelical distances of 8 and 10 angstroms [26]. Molecular dynamics studies of tropomyosin have yielded average values of approximately 11.3 angstroms[27]. Similar interhelical distances have been measured for other coiled-coil proteins such as troponin and calmodulin [28].

The change in free energy on formation (ΔG_{bind}) of the G1, G2 and G3 staggered dimers from their respective constituent monomers were estimated using the single trajectory approach for the MM-PBSA calculation. The results show that dimerization from the monomers pairs is thermodynamically favorable, with values of -11, -33 and -41 kcal/mol for the G1, G2 and G3 dimers (**Figure 31**). However, it must be noted that the free energy values obtained from MM/PBSA calculations should not be interpreted in the same manner as experimental, calorimetric data. This is to say one cannot use these values to calculate respective equilibrium constants. Rather, these values serve as a

means of comparing the stabilities of the dimers with respect to one another. Therefore, the ΔG_{bind} value obtained for the G1 dimer is approximately 30% and 25% of that obtained for the G2 and G3 dimers, respectively. Surprisingly, the difference between the ΔG_{bind} values between G2 and G3 dimers was not significant. This is despite the presence of two extra ordered pairs that could potentially stabilize the G3 dimer.

Stability of oligomers

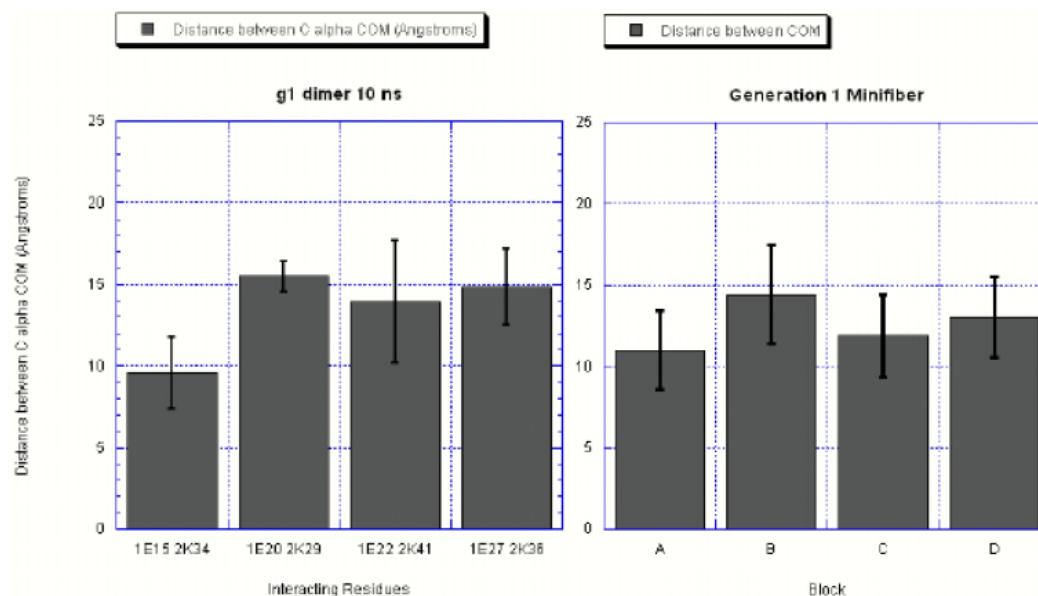
As with the formation of the G1, G2 and G3 dimers, all of the oligomeric combinations studied displayed negative ΔG_{bind} values (**Figure 32**). While there was a general trend towards more negative ΔG_{bind} values with increasing oligomerization number, there were notable energy peaks and troughs, especially in the G2 and G3 species after the tetrameric stage (**Figure 32**). Monomer addition during G2 oligomerization demonstrated a local energy minimum at the pentameric state and a relative energy maximum at the heptameric state. In contrast G3 oligomerization was characterized by local energy minima at the tetrameric, hexameric and octameric states while displaying relative energy maxima at the pentameric and hexameric states. These effects were much less pronounced in the G1 oligomers where oligomerization due to monomer addition produced species of approximately the same energy. Data also indicate a bias against type1 monomers after the tetrameric state during monomer addition to form the G3 oligomers (**Figure 33**)

The RMSD values of the G1 and G2 trimers and tetramers differ markedly from the original, energy-minimized structures in the first frame of the production simulation runs. However, these structures seem to stabilize after the tetrameric state and level off at

approximately 5 angstroms (**Figure 34**). In contrast, the G3 oligomers do not display varying RMSD values with oligomerization number.

Oligomerization number did not significantly affect secondary structure in any of the species studied. All of the oligomers maintained primarily (α -helical structure in all simulation runs with the generation 2 and 3 oligomers slightly more α -helical (~50%) than the generation 1 (~40%)). The 3-10 helix and turn secondary structures made up the combined balance of secondary structures in each oligomer (data not shown).

A)



B)

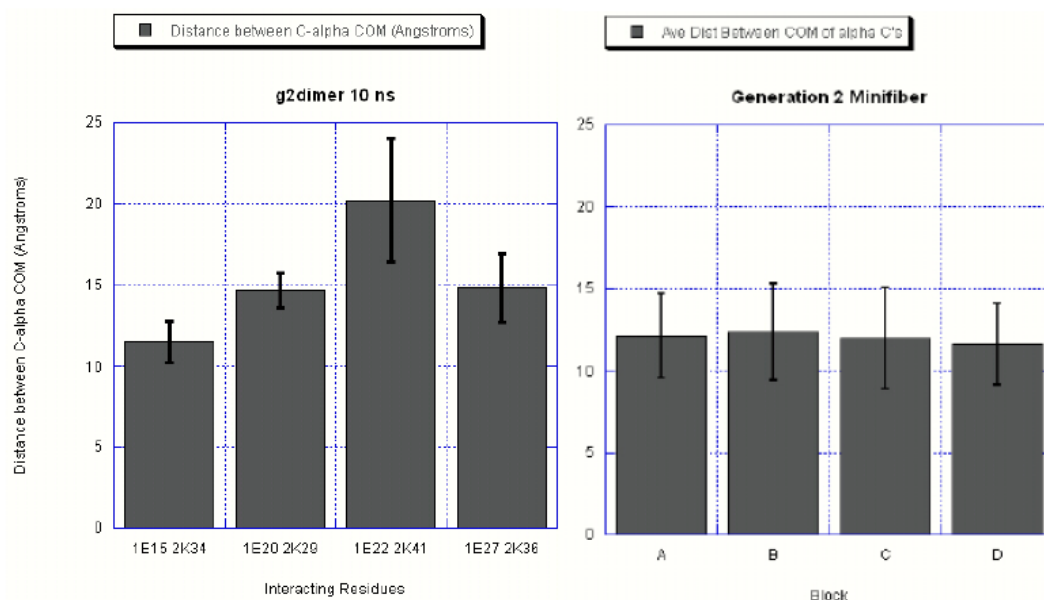


Figure 29- Distance between alpha carbon centers of mass (COM) amongst indicated ordered pairs in the dimer (left) and averaged ordered pairs in the minifiber (right) (A) Generation 1 (B) Generation 2.

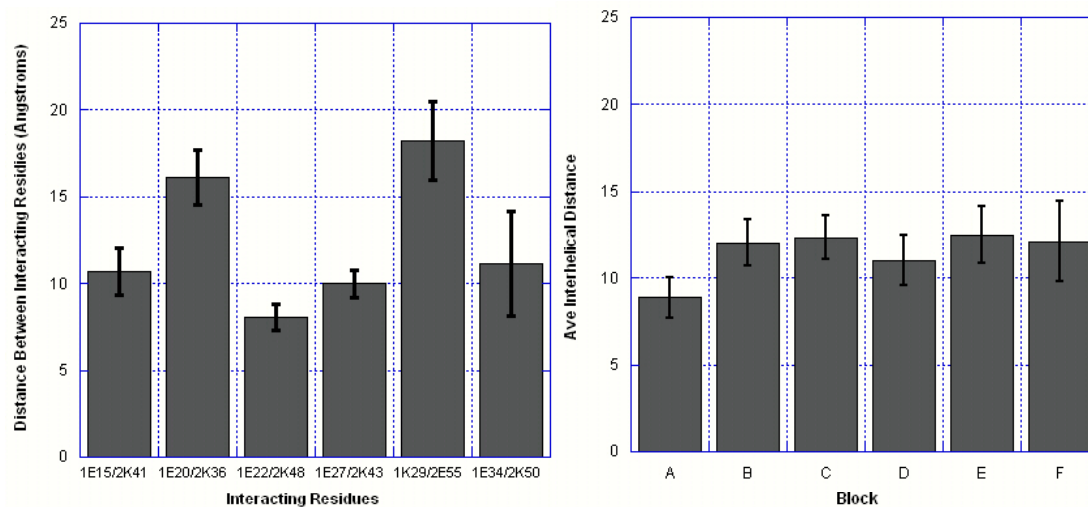


Figure 30- Ordered pair distances in generation 3 dimer (left) vs. minifiber (right).

Error bars indicate standard deviation of ordered pair distances with time.

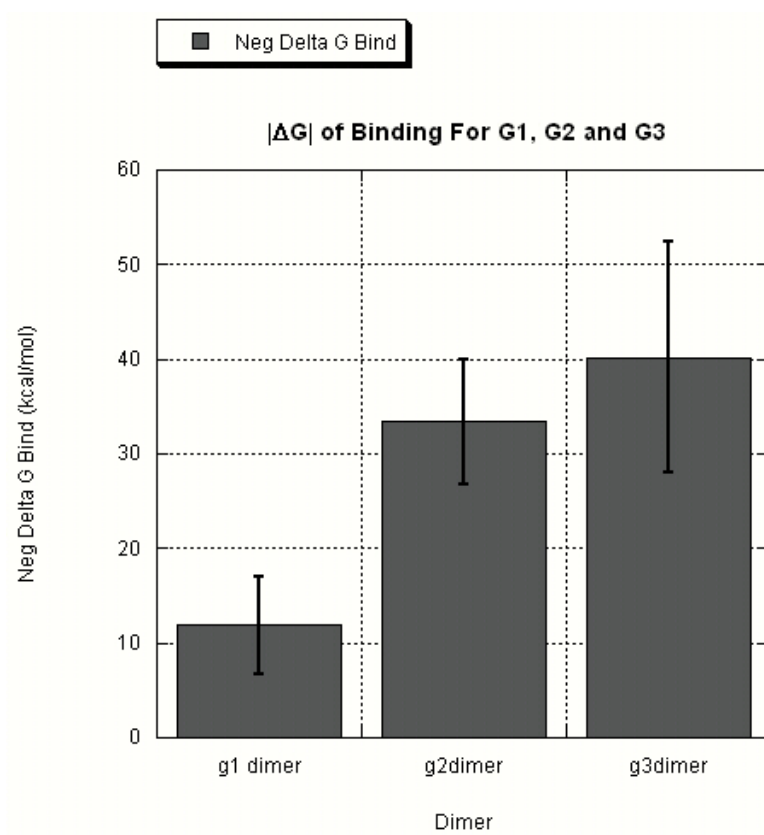


Figure 31- Progressive engineering of Saf system monomers produced dimers with different stabilities in simulations.

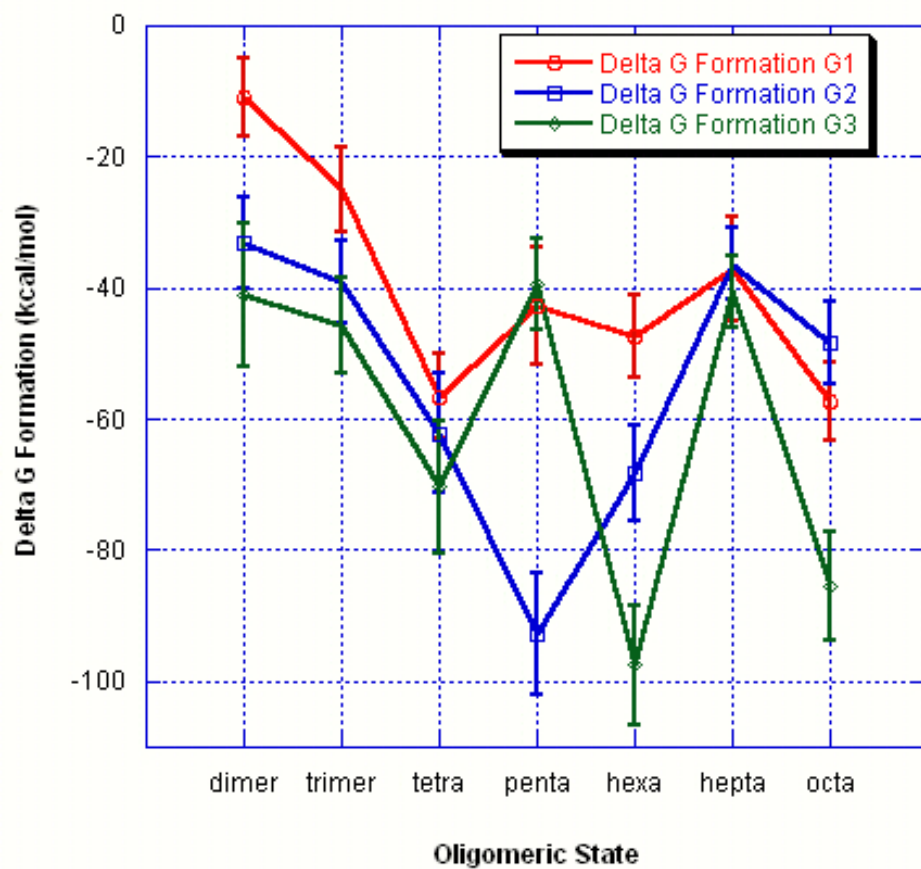


Figure 32- Thermodynamic behavior of the three generations of oligomers is significantly different. G3 oligomerization proceeds with undulations in energetics whereas G2 displays a single energy minimum at the pentameric state.

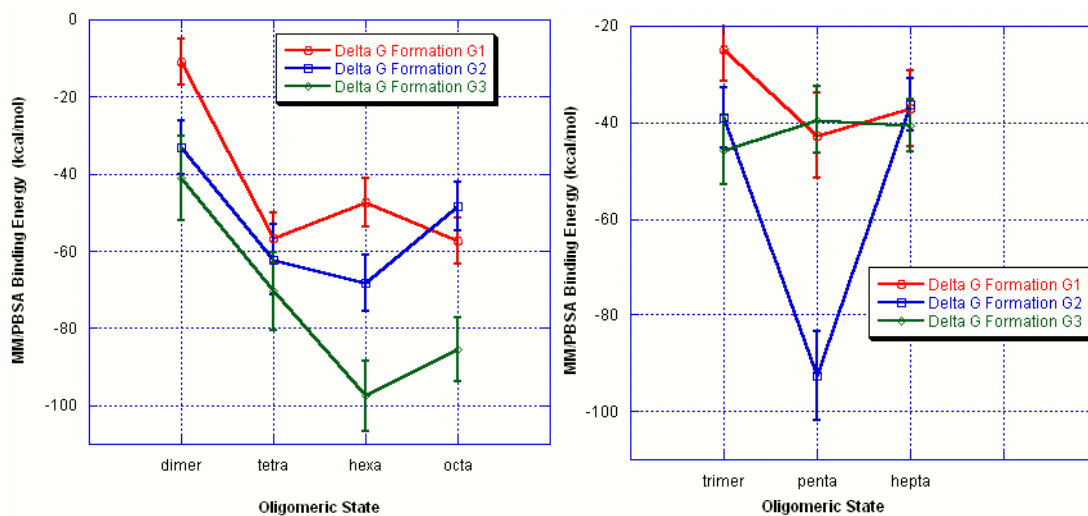


Figure 33- A bias against addition of Type1 and Type2 monomers exists in G3 oligomerization evidenced by static energy changes in the formation of odd numbered oligomers (right) vs. even numbered oligomers (left).

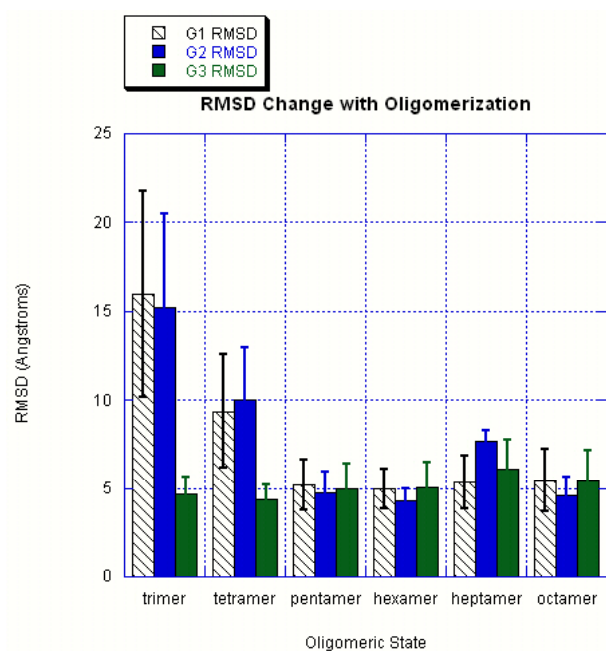


Figure 34- G3 oligomers are structurally sound throughout the oligomerization process whereas G1 and G2 oligomers do not attain structural stability until the pentameric state.

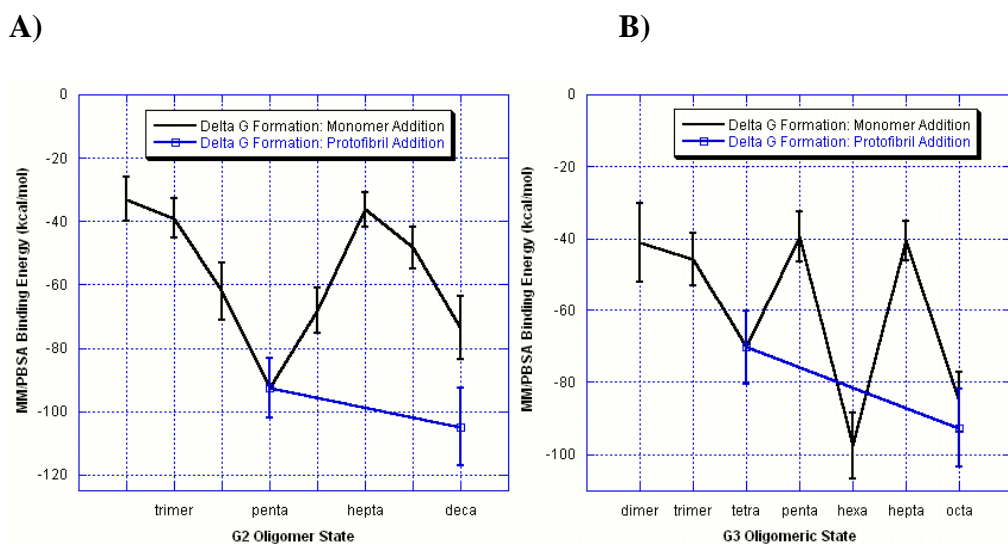


Figure 35- Protofibril addition is a plausible mechanism as simulated by the joining of (A) 2 pentameric units to form the decamer (blue) vs monomer addition (black) in G2 oligomerization. B) 2 tetrameric units to form the octamer (blue) vs monomer addition (black) to form the octamer in G3 oligomerization.

Discussion

We have analyzed the thermodynamics of formation and structural features of several theoretical oligomeric intermediates that would arise during the early stages of the aggregation of Saf monomers into fibers. We assume that there are two critical phases in early Saf aggregation: 1) formation of the staggered dimeric intermediate and 2) axial growth of nascent fibers by progressive monomeric addition of type 1 and type 2 monomers to the original dimer. Here we studied axial growth in one direction only, i.e. addition of a type 1 monomer to the staggered dimer. While mature Saf fibers do exhibit

a preferential assembly direction (polar assembly) this phenomenon has not been described for early axial growth [29].

Molecular dynamics on artificial, non-naturally occurring peptides presents a challenge because no internal coordinates are available as a starting point for the simulations. While it is possible to use any number of molecular modeling programs to generate an amino-acid sequence with defined secondary structure along with corresponding PDB coordinates, it can be difficult to assign that secondary structure into the proper tertiary structure (i.e. coiled-coil). X-ray diffraction studies have shown that Saf aggregates adopt coiled-coil structures [30], thus we chose to model Saf oligomers in the coiled-coil conformation. Other authors have used MD to obtain the coiled-coil structure of the yeast transcriptional activator GCN4 based on its C_{α} backbone [31]. We used a similar strategy but instead chose tropomyosin as a suitable substrate structure on which to replace amino side-chains with those of the Saf peptides. Tropomyosin is a classic coiled-coil protein and pdb coordinates were obtained from a high-resolution (7 angstroms) crystal structure. The structure was sufficiently large to accommodate all of the oligomer models as well as the mini-fibers, which contained 8 monomeric units joined end-to-end on the A and B chains.

Extended MD runs of 10ns performed on each dimer demonstrated maintenance of structural integrity with all ordered pair distances similar to those of the minifiber controls. Minifibers ordered pair distances, rather than corresponding distances in tropomyosin itself, were chosen as controls because coiled-coil diameters have been shown to be dependent on the identity of the hydrophobic amino acids at the a and d

positions [26]. These amino acids change along the length of tropomyosin and as a result change its diameter.

The findings here indicate that the formation of the staggered dimeric complex is a spontaneous event. This “seeding” step is likely thermodynamically controlled because only two, short peptides are involved in the complex and their conformations need not change drastically to accommodate the dimeric form. Thus, dimers with more negative Gibbs free energy changes are most likely more effective seeds for axial growth. Both the G2 and G3 dimers had ΔG_{bind} values 3 to 4 times more negative than the G1 dimer. The major structural difference between the G1 and G2 monomers is found in their type 2 monomers, p2 and p2a (**Table 1**). The change between these monomers is addition of more positively charged residues arginine and glutamine to the p2a monomer. Interestingly, the results show that the ΔG_{bind} values for the G2 and G3 dimers were not significantly different, which is surprising because the G3 peptides were designed to yield more stable staggered dimers because there are two extra oppositely charged ordered-pairs to help stabilize the complex [32].

The simulated stabilities of the proposed Saf dimeric complexes are on the order of naturally occurring ones. Human defensins, which are anti-microbial peptides that are naturally occurring components of the innate immune response, are dimeric complexes consisting of 30-45 amino acids. Studies of these dimers using MMPBSA calculations yielded binding energy values of approximately -22 kcal/mol [33].

Oligomerization binding energies in all generations roughly reflected those of dimerization until a tetrameric state, after which oligomerization behavior between the generations changes drastically. Of primary interest is the appearance of energy minima

and maxima during G2 and G3 oligomerization. With regard to G3 oligomers, there seems to be a bias against the formation of odd numbered mers after the tetrameric stage. In our modeling of monomer addition, odd numbered oligomers are the result of a type 1 monomer pendant to an even numbered chain, which indicates relative energy barriers to the formation of the pentamer and heptamer states. We wondered whether the structure of the even numbered oligomers was inherently more stable than those of their odd-numbered counterparts. Examination of the RMSD and secondary structure as a function of oligomerization showed no changes. In fact, the G3 oligomers remained structurally consistent throughout oligomerization. From this, we conclude that the energy barriers arise due to unfavorable molecular mechanical and/or solvation energies calculated by MMPBSA. The G2 oligomers have one major relative energy minimum at the pentameric stage and do not display preferential addition of type 1 or 2 monomers. However, their structure does not stabilize until the pentameric state as judged by the RMSD plot. After this point, G2 structure remains largely unchanged but the binding energy continues to rise.

Although the origins of the energy maxima at different oligomeric states are unknown, they have potentially important implications on the mechanism of fibrillogenesis. The MMPBSA results suggest that specific oligomeric 'protofibril' states are favored. We simulated protofibril addition by combining two G2 pentameric units to form a G2 decamer and two G3 tetramers to form the G3 octamer. Both processes were spontaneous and both compared favorably to the G2 decamer and G3 octamer formed by monomer addition (**Figure 9**). Thus, protofibril addition is feasible from a thermodynamic standpoint, and we posit that axial growth of fibers takes place, at least

initially, in two distinct phases. First, the staggered dimer forms and elongates by monomer addition to “sticky ends” until oligomers of defined sizes, or protofibrils, are produced. Second, fibers lengthen by protofibril rather than monomer addition. Thus, we propose a switch from monomer addition to protofibril addition during fibrillogenesis of Saf fibers. Presumably, this switch would occur at relative energy minima, as these would be the predominant structures at equilibrium. Thus, the G2 pentamer and the G3 tetramer would serve as these switch points in our simulated system.

Aggregation of monomers to protofibrils that in turn form fibers (or other structures) is a relatively common process among naturally occurring proteins [34, 35] that often includes a critical length dependence on protofibril aggregation [36]. A range of dimeric to octameric oligomers has been found to be stable seeds in the aggregation of β -amyloid [37-39]. With regard to coiled-coil systems, dimeric, tetrameric and octameric intermediates have been shown to predominate during the early stages of intermediate filament formation [40-42]. Fibrin has also been shown to aggregate from protofibrils of various sizes [43-45].

Recently, Woolfson et al. have demonstrated the existence of soluble helical oligomers in a newer generation of fiber-forming α -helical peptides using Analytical Ultracentrifugation (AU) [46]. The results of that study showed a mixture of dimers and trimers in the soluble fraction where lower molecular weight species are usually found. The structure of these oligomers, however, was not elucidated.

In conclusion, we have examined the thermodynamics of oligomer formation during fiber formation in the Saf system of peptides. We have found that progressive steps taken to engineer peptides that produce thicker fibers had significant effects on the

stability of the putative, staggered-dimer intermediate. Furthermore, these modifications also had an impact on the free energy of formation of early oligomers resulting from monomer addition to the staggered dimers. The results suggest that there is a switch during axial elongation from monomer addition to protofibril addition. Relative energy troughs in modeled oligomers that result from monomer addition evidence this. In addition, oligomers associated with energy minima could be used as seeds for protofibril addition and thus increase the assembly rate of fibers.

References

1. Aggeli, A., et al., Exploiting Peptide Self-assembly to Engineer Novel Biopolymers: Tapes, Ribbons, Fibrils and Fibres, in *Self-Assembling Peptide Systems in Biology, Medicine and Engineering*. 2002. p. 1-17.
2. Holmes, T.C., et al., Extensive neurite outgrowth and active synapse formation on self-assembling peptide scaffolds. *Proc Natl Acad Sci U S A*, 2000. **97**(12): p. 6728-33.
3. Anand, P., F.S. Nandel, and U.H. Hansmann, The Alzheimer's beta amyloid (Abeta1-39) monomer in an implicit solvent. *J Chem Phys*, 2008. **128**(16): p. 165102.
4. Pandya, M.J., et al., Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis. *Biochemistry*, 2000. **39**(30): p. 8728-34.
5. Melnik, T.N., et al., Shift of fibril-forming ability of the designed alpha-helical coiled-coil peptides into the physiological pH region. *Protein Eng*, 2003. **16**(12): p. 1125-30.
6. Andrew M. Smith, E.F.B., Wayne R. Edwards, Maya J. Pandya, Derek N. Woolfson, Engineering Increased Stability into Self-Assembled Protein Fibers. *Advanced Functional Materials*, 2006. **0000**(00): p. 1-9.
7. Hamley, I.W., Peptide fibrillization. *Angew Chem Int Ed Engl*, 2007. **46**(43): p. 8128-47.
8. Kane, M.D., et al., Evidence for Seeding of beta -Amyloid by Intracerebral Infusion of Alzheimer Brain Extracts in beta -Amyloid Precursor Protein-Transgenic Mice. *J. Neurosci.*, 2000. **20**(10): p. 3606-3611.
9. Kaye, R., et al., Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science*, 2003. **300**(5618): p. 486-9.
10. Serio, T.R., et al., Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science*, 2000. **289**(5483): p. 1317-21.
11. Fink, A.L., The aggregation and fibrillation of alpha-synuclein. *Acc Chem Res*, 2006. **39**(9): p. 628-34.

12. Hong, D.P. and A.L. Fink, Independent heterologous fibrillation of insulin and its B-chain peptide. *Biochemistry*, 2005. **44**(50): p. 16701-9.
13. Khurana, R., et al., Partially folded intermediates as critical precursors of light chain amyloid fibrils and amorphous aggregates. *Biochemistry*, 2001. **40**(12): p. 3525-35.
14. Fink, A.L., Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Fold Des*, 1998. **3**(1): p. R9-23.
15. Herrmann, H., et al., Characterization of distinct early assembly units of different intermediate filament proteins. *J Mol Biol*, 1999. **286**(5): p. 1403-20.
16. Urbanc, B., et al., Molecular dynamics simulation of amyloid beta dimer formation. *Biophys J*, 2004. **87**(4): p. 2310-21.
17. Berman, H.M., et al., The Protein Data Bank. *Nucleic Acids Res*, 2000. **28**(1): p. 235-42.
18. Bowie, J.U., R. Luthy, and D. Eisenberg, A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, 1991. **253**(5016): p. 164-70.
19. Xu, J., F. Jiao, and L. Yu, Protein structure prediction using threading. *Methods Mol Biol*, 2008. **413**: p. 91-121.
20. Ding, F., J.J. LaRocque, and N.V. Dokholyan, Direct Observation of Protein Folding, Aggregation, and a Prion-like Conformational Conversion. *J. Biol. Chem.*, 2005. **280**(48): p. 40235-40240.
21. Whitby, F.G. and G.N. Phillips, Jr., Crystal structure of tropomyosin at 7 Angstroms resolution. *Proteins*, 2000. **38**(1): p. 49-59.
22. Guex, N. and M.C. Peitsch, SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, 1997. **18**(15): p. 2714-23.
23. UL Essmann, T.B., H Darden, H Lee, LG Pedersen A smooth Particle Mesh Ewald method. *J Chem Phys*, 1995. **103**: p. 8577-8593.
24. J Ryckaert, G., Ciccotti, HJC Berendsen, Numerical Integration of the Cartesian eqations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys*, 1977. **23**: p. 327-341.
25. Berendsen, H., JPM Postma, WF van Gunsternen, A DiNola and JR Haak, Molecular Dynamics with coupling to an external bath. *J Chem Phys*, 1984. **81**: p. 3684-2690.
26. Brown, J.H., et al., Deciphering the design of the tropomyosin molecule. *Proceedings of the National Academy of Sciences of the United States of America*, 2001. **98**(15): p. 8496-8501.
27. Minakata, S., et al., Two-Crystal Structures of Tropomyosin C-Terminal Fragment 176-273: Exposure of the Hydrophobic Core to the Solvent Destabilizes the Tropomyosin Molecule. *Biophys. J.*, 2008. **95**(2): p. 710-719.
28. KuniHiko, G., et al., Ca²⁺ binding sites in calmodulin and troponin C alter interhelical angle movements. *FEBS letters*, 2004. **561**(1): p. 51-57.
29. Smith, A.M., et al., Polar assembly in a designed protein fiber. *Angew Chem Int Ed Engl*, 2004. **44**(2): p. 325-8.

30. Papapostolou, D., et al., Engineering nanoscale order into a designed protein fiber. *Proceedings of the National Academy of Sciences*, 2007. **104**(26): p. 10853-10858.
31. Nilges, M. and A.T. Brunger, Automated modeling of coiled coils: application to the GCN4 dimerization region. *Protein Eng.*, 1991. **4**(6): p. 649-659.
32. AM Smith, E.B., WR Edwards, MJ Pandya, DN Woolfson, Engineering Increased Stability into Self-Assembled Protein Fibers. *Advanced Functional Materials*, 2006. **16**: p. 1022-30.
33. Suresh, A. and C. Verma, Modelling study of dimerization in mammalian defensins. *BMC Bioinformatics*, 2006. **7 Suppl 5**: p. S17.
34. Ali, M.H. and B. Imperiali, Protein oligomerization: how and why. *Bioorg Med Chem*, 2005. **13**(17): p. 5013-20.
35. Carrotta, R., et al., Conformational characterization of oligomeric intermediates and aggregates in {beta}-lactoglobulin heat aggregation. *Protein Sci*, 2001. **10**(7): p. 1312-1318.
36. Khare, S.D., et al., Molecular Origin of Polyglutamine Aggregation in Neurodegenerative Diseases. *PLoS Computational Biology*, 2005. **1**(3): p. e30.
37. Ma, B. and R. Nussinov, Molecular dynamics simulations of alanine rich {beta}-sheet oligomers: Insight into amyloid formation. *Protein Sci*, 2002. **11**(10): p. 2335-2350.
38. Antwi, K., et al., Cu(II) organizes {beta}-2-microglobulin oligomers but is released upon amyloid formation. *Protein Sci*, 2008: p. ps.073249008.
39. Hartley, D.M., et al., Protofibrillar Intermediates of Amyloid beta -Protein Induce Acute Electrophysiological Changes and Progressive Neurotoxicity in Cortical Neurons. *J. Neurosci.*, 1999. **19**(20): p. 8876-8884.
40. Herrmann, H. and U. Aepli, Intermediate filament assembly: fibrillogenesis is driven by decisive dimer-dimer interactions. *Curr Opin Struct Biol*, 1998. **8**(2): p. 177-85.
41. Er Rafik, M., J. Doucet, and F. Briki, The Intermediate Filament Architecture as Determined by X-Ray Diffraction Modeling of Hard {alpha}-Keratin. *Biophys. J.*, 2004. **86**(6): p. 3893-3904.
42. Athlan, E.S. and W.E. Mushynski, Heterodimeric Associations between Neuronal Intermediate Filament Proteins. *J. Biol. Chem.*, 1997. **272**(49): p. 31073-31078.
43. Paul, A.J., Kinetics of formation of fibrin oligomers. I. Theory. *Biopolymers*, 1982. **21**(11): p. 2253-2264.
44. Bark, N., Z. Foldes-Papp, and R. Rigler, The incipient stage in thrombin-induced fibrin polymerization detected by FCS at the single molecule level. *Biochem Biophys Res Commun*, 1999. **260**(1): p. 35-41.
45. Li, X., D. Galanakis, and D.A. Gabriel, Transient Intermediates in the Thrombin Activation of Fibrinogen. *J. Biol. Chem.*, 1996. **271**(20): p. 11767-11771.
46. Papapostolou, D., et al., Electrostatic Control of Thickness and Stiffness in a Designed Protein Fiber. *J. Am. Chem. Soc.*, 2008. **130**(15): p. 5124-5130.

Chapter 4- Discussion

Biomaterials that use proteins as their constituent units are now being researched at an ever-increasing rate [1]. This is primarily because a vast proportion of the physiologic milieu is composed of protein. Materials, such as tissue engineering scaffolds, which are designed to support viability and function of cells may well display enhanced functionality if made of protein units. However, such materials would have to be amenable to rational modification so they could be tailored to the specific requirements of each cell type. Even though proteins form a vast array of functionally distinct and complex structures, rational engineering of their properties has remained elusive. Thus, efforts to manufacture new protein biomaterials will be hampered until a better understanding is achieved in fields such as protein-protein interactions, protein aggregation, protein folding and conformation. Furthermore, advances in scientific computing (hardware and software), total gene synthesis, cloning, expression and purification of recombinant constructs will also aid in the design of new proteins with targeted physical properties.

In this thesis we have attempted to address some of these knowledge gaps in the creation of a new protein based biomaterials in three ways. First, we have designed a hydrogel based on a modular protein system utilizing a triblock, ABA motif supplemented with small peptides that can elongate endblocks via non-covalent forces. Such a design opens new avenues to rational engineering because each of the module's amino-acid sequences can be altered separately to affect their functionality in the

aggregated state. For example, the hydrophilic B block can be modified to bind metals [2], and include cell attachment motifs such as RGD sequences [3, 4]. The helical endblocks have been modified with respect to their amino acid sequences and length to produce hydrogels with different gelation properties [2, 5, 6].

Second, we have explored the expression and purification of two novel recombinant proteins in *E. coli* by two separate methods. Advances in total gene synthesis have now made available the possibility of constructing genes that encode almost any protein sequence at a relatively low cost. However, the expression and purification of these genes is far from an exact science and will almost certainly involve a great deal of trial-and-error in every case. Furthermore, investigators must be prepared for the possibility that a particular gene construct cannot be expressed at acceptable levels in the bacterial host. In fact, such an enterprise involves an approximate failure rate of 70% [7]. That being the case, it is critical to design expression studies that offer multiple routes to the desired product such as expression as a number of fusion proteins along with other types of purification tags. The methods we presented here were expression of the linked triblocks as MBP fusion proteins and as non-fused, 6x His tagged proteins. It was clear from our studies that the expression of the MBP fusion would have resulted in better yield and superior purity of our target proteins when compared to that of the 6x His tagged protein. Unfortunately, cleavage of the MBP tag also resulted in cleavage of the target, linked proteins as well due to proteolysis at non-canonical sequences within the target. We were able to purify linked proteins expressed as the 6x His tag, albeit at very low levels. Again, proteolysis was suspected but this time as the action of serine proteases native to the *E. coli* host. Undoubtedly, another round of expression would be

required to produce the protein at acceptable levels and purity. We suggest expression as a SUMO fusion because the SUMO protease recognized the tertiary structure of the SUMO tag rather than a short sequence of amino-acids. Furthermore, expression as a fusion protein may shield the linked proteins from native *E. coli* proteases [8].

Overall, producing a hydrogel from expressed protein subunits is an inefficient process. Several years were required to find the correct vectors and purification conditions to produce the triblock unlinked units. This was done at great cost in terms of research dollars. Furthermore, yields of recombinant proteins are in the mg range and must be scaled-up in large fermenting units if they are to have any commercial value at all [9]. Services are now available that can express and purify proteins from recombinant genes, however the cost is still prohibitive (\$5,000-10,000) for approximately 10mg of purified protein.

Third, we have use all-atom MD simulations to describe the thermodynamics of the early aggregation steps of the Saf system of peptides. To our knowledge, such modeling has not been performed for an aggregating, biomaterial system. Our results suggest the presence of (at least) a two-phase mechanism of aggregation of these peptides from monomer addition to protofibril addition, and strongly indicate specific intermediates that depend on the specific sequence of the monomers. This is evidenced by the existence of potential energy barriers with oligomerization number if oligomerization had proceeded by monomer addition alone. Implicit in this model is the assumption that these steps are under thermodynamic control. At the very least, kinetics of monomer addition in this early phase stays relatively constant. Such assumptions have

been employed in the modeling of the oligomerization of amyloid A β_{16-22} and yeast prion protein Sup 35 [10, 11].

It would be ideal to verify the protofibril mechanism by obtaining empirical results in the laboratory. For example, the MD simulations suggest G2 monomers aggregate via addition of pentameric units. Thus, $n \times 5$ mers would be predominant in these early aggregates. Though lifetime of early oligomers is likely to be fleeting, it is possible to isolate them using Analytic Ultracentrifugation (AUC). Woolfson et al. have already performed this method to isolate early oligomeric states of their most recent generation of fiber forming peptides [12]. This method relies on the difference between sedimentation rates of molecules of different sizes. Alternatively, Multiangle Light Scattering can be used to study the molecular weight of macromolecular intermediates as outlined by Bernacco et al [13]. These investigators were able to measure the molecular weights of rod-like monomers in the aggregation of fibrin using a stop-flow mixer linked to a commercial multiangle light scattering photometer.

The modeling results also suggest that we may achieve improved endblock/unlinked protein interhelical association if both the endblocks and unlinked proteins are changed with respect to their length. For example, if the generation 3 helices are used as endblocks, we could alter the coding region for the A blocks to be tetramers (as four repeating 35 amino-acid generation 3 helices). This follows from the result that there is a relative energy minima at the tetrameric state. Similarly, generation 2 endblocks and unlinked proteins could be extended to the pentameric level (five repeating generation 2 helices) as this was the main energy minimum found in the MD simulations.

Use of all-atom MD where the solvent is explicitly represented also has its limitations. Even 2ns of data obtained in our studies were produced from runs that could exceed 24 hours utilizing 16 processors running Amber 9. While this yielded insight into the thermodynamics of early aggregation steps of unlinked peptides, it cannot give any information on the final material properties of the aggregated state. Nor can it explain how many large protein subunits would interact or how they would be aligned in the final gel network. Implicit solvent MD methods are now available to reduce the computational cost of calculating the trajectory of each water molecule in the simulation system. These methods include those based on a solvation term based on solvent accessible surface area (SASA) [14], generalized Born model, and distance dependent dielectric (DDD) model [15]. Such modeling would be ideal for the simulation of the larger protein units indicated in the protofibril aggregation model that we have proposed. Once again, such simulations could glean no final material properties of the hydrogel but it would represent a step towards simulating the system with slightly more complexity. In so doing, we could then move on to create a course-grain model where much larger and many more protein units could be involved in the simulation.

References

1. Terech, R.G.W.a.P., ed. *Molecular Gels. Materials with Self-Assembled Fibrillar Networks*. 2006, Springer.
2. Lao, U.L., et al., Genetic Engineering of Self-Assembled Protein Hydrogel Based on Elastin-like Sequences with Metal Binding Functionality. *Biomacromolecules*, 2007. **8**(12): p. 3736-3739.
3. Mi, L., et al., Self-assembling protein hydrogels with modular integrin binding domains. *Biomacromolecules*, 2006. **7**(1): p. 38-47.

4. Zhiyuan Hu, F.L.Y.P.C.H.L.R.J.C.J.W.Y.Z., Arg-Gly-Asp (RGD) peptide conjugated poly(lactic acid)-poly(ethylene oxide) micelle for targeted drug delivery. *Journal of Biomedical Materials Research Part A*, 2008. **85A**(3): p. 797-807.
5. Xu, C., V. Breedveld, and J. Kopecek, Reversible hydrogels from self-assembling genetically engineered protein block copolymers. *Biomacromolecules*, 2005. **6**(3): p. 1739-49.
6. Xu, C. and J. Kopecek, Genetically engineered block copolymers: influence of the length and structure of the coiled-coil blocks on hydrogel self-assembly. *Pharm Res*, 2008. **25**(3): p. 674-82.
7. Graslund, S., et al., Protein production and purification. *Nat Methods*, 2008. **5**(2): p. 135-46.
8. Tang, W., et al., An efficient system for production of recombinant urokinase-type plasminogen activator. *Protein Expr Purif*, 1997. **11**(3): p. 279-83.
9. Jung, K.-H., Continuous production of recombinant interferon-[alpha] in *Escherichia coli* via the derepression of trp promoter using casamino acid. *Process Biochemistry*, 2006. **41**(4): p. 809-814.
10. Favrin, G., A. Irback, and S. Mohanty, Oligomerization of amyloid A{beta}16-22 peptides using hydrogen bonds and hydrophobicity forces. *Biophys. J.*, 2004: p. biophysj.104.046839.
11. Gsponer, J., U. Haberthur, and A. Caflisch, The role of side-chain interactions in the early steps of aggregation: Molecular dynamics simulations of an amyloid-forming peptide from the yeast prion Sup35. *Proc Natl Acad Sci U S A*, 2003. **100**(9): p. 5154-9.
12. Papapostolou, D., et al., Electrostatic Control of Thickness and Stiffness in a Designed Protein Fiber. *J. Am. Chem. Soc.*, 2008. **130**(15): p. 5124-5130.
13. Bernocco, S., et al., Polymerization of rod-like macromolecular monomers studied by stopped-flow, multiangle light scattering: set-up, data processing, and application to fibrin formation. *Biophys J*, 2000. **79**(1): p. 561-83.
14. Philippe Ferrara, J.A.A.C., Evaluation of a fast implicit solvent model for molecular dynamics simulations. *Proteins: Structure, Function, and Genetics*, 2002. **46**(1): p. 24-33.
15. Ting Wang, R.C.W., Implicit solvent models for flexible protein-protein docking by molecular dynamics simulation. *Proteins: Structure, Function, and Genetics*, 2003. **50**(1): p. 158-169.

Chapter 5- Future Directions

Improved Design and Expression of Triblock Proteins

The results of this research suggest that there is room for considerable improvement of the expression of the triblock linked protein units. As has been stated previously, this appears to be best accomplished by expressing the triblocks as fusion proteins as they are particularly vulnerable to protease cleavage. A myriad of fusion protein systems are available that could well perform this task. However, in the interest of time a SUMO fusion would be the best candidate. Fragment analysis performed in Chapter 2 revealed that the likely cleavage sites were in the endblocks themselves. It must be noted that these endblocks have not been expressed in *E. coli* and it is their expression that seems problematic. Use of the SUMO protease would greatly reduce the chance of the linked proteins being cleaved. However, there are also other systems that could be used that employ genetically engineered proteases that display enhanced specificity. One such example is the PreScission protease system (GE Healthcare, Piscataway, NJ) which utilizes the purification tag Glutathione-S-transferase (GST) and is purified using GST-antibody affinity columns. The advantage to this system is that the target protein can be cleaved from the GST tag while still bound to the affinity column [1]. This system has received positive reviews from several investigators (Claes Gaustoffsen, DNA2.0, personal communication).

Simulation results suggest alterations in endblock design. First, dimerization results reveal that better interhelical associations between the linked and unlinked can be

achieved by using the generation 2 and generation 3 helices. This follows from the result that dimerization stability increased three to fourfold as compared to the generation 1 helices. This could be important because strength of interhelical associations were shown to change the gelation point of other genetically engineered triblock hydrogels [2, 3]. Furthermore, since dimer formation is believed to seed fiber formation in the Saf system of peptides, this redesign feature could potentially yield enhanced elongation of the endblocks at the same time [4].

To correspond with the elongated endblocks, unlinked units of the same size would have to be expressed in *E. coli*. We would presumably express these as SUMO fusions as well. In addition to using these unlinked units in the newly re-designed hydrogels, they should also be assessed for their fibrillogenic potential on their own. This can be done using Circular Dichroism (CD) Spectroscopy. The ratio of the absorbencies at 220nm and 208nm has been used to indicate the degree of association of α -helices into their coiled-coil forms [5]. The closer this ratio is to 1, the greater of degree of association between the helices. We have already tested this method on the Generation 1 and Generation 2 fiber mixtures as in **Figure 36**. In accord with simulation results, the helical stability of fibers formed from G2 peptides was superior to that of the G1 mixture.

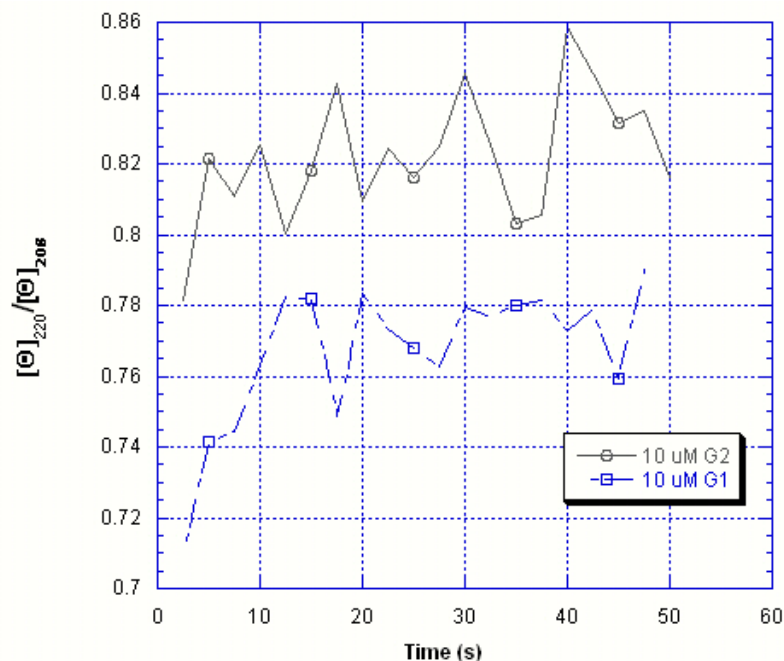


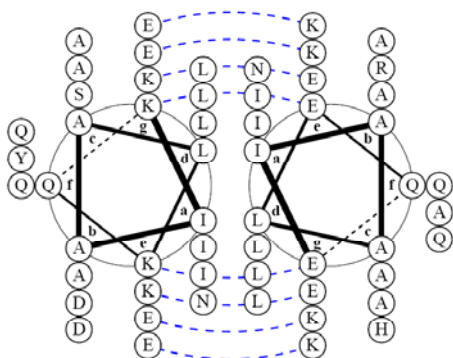
Figure 36- The G2 fiber mixture displays superior coiled-coil forming ability as compared to the G1 mixture.

Improving Dimer Stability

Since the strength of interhelical associations would seem to play a significant role in temperatures of gelation and melting, we propose placement of phenylalanine residues in the hydrophobic pocket (i.e. the ‘a’ and ‘d’ residues in the helical wheel in **Figure 37A**). Placement of phenylalanine residues may improve the stability of the staggered dimer form due to a phenomenon known as pi-stacking. These stabilizing forces arise due to favorable interactions between between the molecular orbitals of adjacent aromatic ring. This alignment stabilizes structures such as the DNA double helix as well as the interactions among newly developed peptides [6-9]. Initial simulations will

focus on the effect of adding phenylalanine residues to the hydrophobic pocket in the area of overlap of the staggered dimers (**Figures 37A and 37B**). Although phenylalanine is not as hydrophobic as leucine or isoleucine on the Kyte-Doolittle Scale [10], the stability of the pi-stacking interactions of several aligned phenyl groups could potentially offset this parameter. However, we are unsure of how the addition of relatively large aromatic groups into the hydrophobic core of the coiled-coil will affect the ordered-pair, charge-charge interactions that help stabilize staggered dimer formations. Tropomyosin diameter has been shown to be affected by the identity of the amino acids at the 'a' and 'd' positions [11]. Thus the measurement of the charge-charge ordered pair inter-helical distances may be done once more and compared to the results in **Chapter 2** to determine that these interhelical distances have not increased significantly.

A)



B)

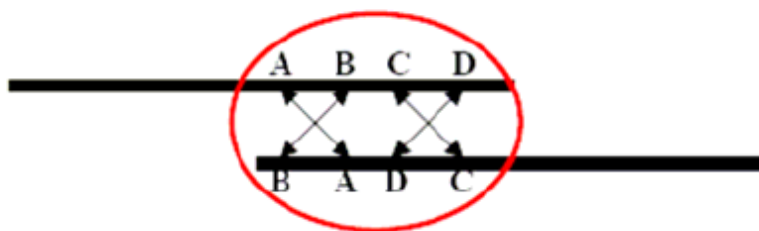


Figure 37- (A) Residues at the 'a' and 'd' positions in the helical wheel diagram represent the hydrophobic pocket region where phenylalanine residues will be placed. B) Staggered dimer with the overlap region highlighted.

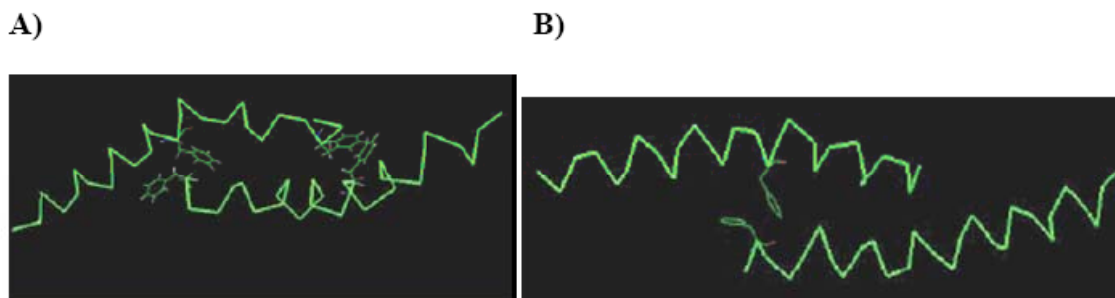


Figure 38- Example of placement of the phenylalanine residues in the hydrophobic, overlap region of the staggered dimer to promote pi-stacking in the dimer (A) Placement of 4 phenylalanine residues. B) Placement of two phenylalanine residues.

Developing Coarse-Grain Models of Peptide Aggregation and Hydrogelation

We would like to model peptide aggregation on a larger scale and for a greater period of time in order to more accurately predict how amino acid alterations in endblocks and unlinked subunits will effect such aggregation. We can work towards this goal by modeling our aggregating system in several phases. In each progressive phase, we are able to make more approximations and lose some degree of superfluous detail as suggested by the previous phase. For example, our atomistic models have indicated that fiber formation may proceed through a protofibril addition mechanism. Thus, in the next phase of modeling, peptides would not be represented as many 28 or 35 amino-acid helices but as fewer 140 amino-acid blocks. Furthermore, we could employ an implicit solvent model such as Generalized Born/ DDD to rid the simulations of the burden of calculating solvent trajectories. In so doing we can maximize the available computing power for the simulation of these larger units. Data obtained from fiber formation studies (CD and AU described above) could also be used to parametrize the coarse grain models as well. Eventually, we hope to reduce the number of interaction sites between the

macromolecular constructs and assign broad categories of charge, hydrophilicity and hydrophobicity to separate parts of these constructs.

Modeling hydrogel formation can benefit from the use of hierarchical strategy for model building. In such models, each monomer in a polymer is represented as spherical coarse grain unit [12]. Hydrophobic and hydrophilic groups are modeled as linear repeats which helps make parametrization slightly easier. As a first approximation, the ABA triblock entities that compose our hydrogel can be modeled with the A blocks as hydrophobic groups and the B block as a hydrophilic group.

References

1. Sigrell, J., Improving the purification of GST Fusion Proteins. International Biotechnology Laboratory, 2002.
2. Xu, C., V. Breedveld, and J. Kopecek, Reversible hydrogels from self-assembling genetically engineered protein block copolymers. *Biomacromolecules*, 2005. **6**(3): p. 1739-49.
3. Xu, C. and J. Kopecek, Genetically engineered block copolymers: influence of the length and structure of the coiled-coil blocks on hydrogel self-assembly. *Pharm Res*, 2008. **25**(3): p. 674-82.
4. Pandya, M.J., et al., Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis. *Biochemistry*, 2000. **39**(30): p. 8728-34.
5. Choy, N., V. Raussens, and V. Narayanaswami, Inter-molecular Coiled-coil Formation in Human Apolipoprotein E C-terminal Domain. *Journal of Molecular Biology*, 2003. **334**(3): p. 527-539.
6. Jahnke, E., et al., Functional, hierarchically structured poly(diacetylene)s via supramolecular self-assembly. *Macromol Biosci*, 2007. **7**(2): p. 136-43.
7. Jayawarna, V., et al., Three-dimensional cell culture of chondrocytes on modified di-phenylalanine scaffolds. *Biochem Soc Trans*, 2007. **35**(Pt 3): p. 535-7.
8. South, C.R., C. Burd, and M. Weck, Modular and dynamic functionalization of polymeric scaffolds. *Acc Chem Res*, 2007. **40**(1): p. 63-74.
9. Gazit, E., Self-assembly of short aromatic peptides: From amyloid disease to nanotechnology. *NanoBioTechnology*, 2005. **1**(3): p. 286-288.
10. Kyte, J. and R.F. Doolittle, A simple method for displaying the hydropathic character of a protein. *J Mol Biol*, 1982. **157**(1): p. 105-32.

11. Brown, J.H., et al., Deciphering the design of the tropomyosin molecule. Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(15): p. 8496-8501.
12. Steve, O.N., et al., Coarse grain models and the computer simulation of soft materials. Journal of Physics: Condensed Matter, 2004(15): p. R481.

Bibliography

1. Nakagawa, T. and J. Ito, Drug delivery systems for the treatment of sensorineural hearing loss. *Acta Otolaryngol Suppl*, 2007(557): p. 30-5.
2. Hamley, I.W., Peptide fibrillization. *Angew Chem Int Ed Engl*, 2007. **46**(43): p. 8128-47.
3. Dong, S. and X. Chen, Some new aspects in biosensors. *J Biotechnol*, 2002. **82**(4): p. 303-23.
4. Stapleton, F., et al., Silicone hydrogel contact lenses and the ocular surface. *Ocul Surf*, 2006. **4**(1): p. 24-43.
5. Martens, P.J., S.J. Bryant, and K.S. Anseth, Tailoring the degradation of hydrogels formed from multivinyl poly(ethylene glycol) and poly(vinyl alcohol) macromers for cartilage tissue engineering. *Biomacromolecules*, 2003. **4**(2): p. 283-92.
6. Weiss, R.G. and P.T. , *Molecular Gels: Materials with Self-assembled Fibrillar Networks*. 2005.
7. Lehning, E.J., et al., Biochemical and morphologic characterization of acrylamide peripheral neuropathy. *Toxicol Appl Pharmacol*, 1998. **151**(2): p. 211-21.
8. Stegemann, J.P., S.N. Kaszuba, and S.L. Rowe, Review: advances in vascular tissue engineering using protein-based biomaterials. *Tissue Eng*, 2007. **13**(11): p. 2601-13.
9. Daamen, W.F., et al., Elastin as a biomaterial for tissue engineering. *Biomaterials*, 2007. **28**(30): p. 4378-98.
10. Koide, T., Designed triple-helical peptides as tools for collagen biochemistry and matrix engineering. *Philos Trans R Soc Lond B Biol Sci*, 2007. **362**(1484): p. 1281-91.
11. Woolfson, D.N. and M.G. Ryadnov, Peptide-based fibrous biomaterials: Some things old, new and borrowed. *Curr Opin Chem Biol*, 2006. **10**(6): p. 559-67.
12. Dinerman, A.A., et al., Swelling behavior of a genetically engineered silk-elastinlike protein polymer hydrogel. *Biomaterials*, 2002. **23**(21): p. 4203-10.
13. Ehrick, J.D., et al., Genetically engineered protein in hydrogels tailors stimuli-responsive characteristics. *Nat Mater*, 2005. **4**(4): p. 298-302.
14. Shen, W., et al., Assembly of an Artificial Protein Hydrogel through Leucine Zipper Aggregation and Disulfide Bond Formation. *Macromolecules*, 2005. **38**(9): p. 3909-3916.
15. Shen, W., et al., Tuning the erosion rate of artificial protein hydrogels through control of network topology. *Nat Mater*, 2006. **5**(2): p. 153-8.
16. Petka, W.A., et al., Reversible Hydrogels from Self-Assembling Artificial Proteins. *Science*, 1998. **281**(5375): p. 389-392.
17. Wheeldon, I.R., S.C. Barton, and S. Banta, Bioactive proteinaceous hydrogels from designed bifunctional building blocks. *Biomacromolecules*, 2007. **8**(10): p. 2990-4.

18. AM Smith, E.B., WR Edwards, MJ Pandya, DN Woolfson, Engineering Increased Stability into Self-Assembled Protein Fibers. *Advanced Functional Materials*, 2006. **16**: p. 1022-30.
19. Pandya, M.J., et al., Sticky-end assembly of a designed peptide fiber provides insight into protein fibrillogenesis. *Biochemistry*, 2000. **39**(30): p. 8728-34.
20. Andrew M. Smith, E.F.B., Wayne R. Edwards, Maya J. Pandya, Derek N. Woolfson, Engineering Increased Stability into Self-Assembled Protein Fibers. *Advanced Functional Materials*, 2006. **0000**(00): p. 1-9.
21. Balgude, A.P., et al., Agarose gel stiffness determines rate of DRG neurite extension in 3D cultures. *Biomaterials*, 2001. **22**(10): p. 1077-84.
22. LeSar, A.R.a.R., Modeling and Simulation of Biomaterials. *Annu. Rev. Mater. Res.*, 2004. **34**: p. 279-314.
23. Rezaei-Ghaleh, N., M. Amininasab, and M. Nemat-Gorgani, Conformational changes of {alpha}-chymotrypsin in a fibrillation-promoting condition: A molecular dynamics study. *Biophys J*, 2008.
24. Melnik, T.N., et al., Shift of fibril-forming ability of the designed alpha-helical coiled-coil peptides into the physiological pH region. *Protein Eng*, 2003. **16**(12): p. 1125-30.
25. Anand, P., F.S. Nandel, and U.H. Hansmann, The Alzheimer's beta amyloid (Abeta1-39) monomer in an implicit solvent. *J Chem Phys*, 2008. **128**(16): p. 165102.
26. Stabenfeldt, S.E., A.J. Garcia, and M.C. LaPlaca, Thermoreversible laminin-functionalized hydrogel for neural tissue engineering. *J Biomed Mater Res A*, 2006. **77**(4): p. 718-25.
27. Yan, H., et al., Thermoreversible protein hydrogel as cell scaffold. *Biomacromolecules*, 2006. **7**(10): p. 2776-82.
28. Wang, C., R.J. Stewart, and J. Kopecek, Hybrid hydrogels assembled from synthetic polymers and coiled-coil protein domains. *Nature*, 1999. **397**(6718): p. 417-20.
29. Jung, J.P., et al., Modulating the mechanical properties of self-assembled peptide hydrogels via native chemical ligation. *Biomaterials*, 2008. **29**(13): p. 2143-51.
30. Rao, K.P., Recent developments of collagen-based materials for medical applications and drug delivery systems. *J Biomater Sci Polym Ed*, 1995. **7**(7): p. 623-45.
31. Yu, X., G.P. Dillon, and R.B. Bellamkonda, A laminin and nerve growth factor-laden three-dimensional scaffold for enhanced neurite extension. *Tissue Eng*, 1999. **5**(4): p. 291-304.
32. Park, K.H. and K. Na, Effect of growth factors on chondrogenic differentiation of rabbit mesenchymal cells embedded in injectable hydrogels. *J Biosci Bioeng*, 2008. **106**(1): p. 74-9.
33. Wang, S., et al., Three-dimensional primary hepatocyte culture in synthetic self-assembling peptide hydrogel. *Tissue Eng Part A*, 2008. **14**(2): p. 227-36.
34. Peppas, N.A., et al., Hydrogels in pharmaceutical formulations. *Eur J Pharm Biopharm*, 2000. **50**(1): p. 27-46.
35. Tang, A., et al., The coiled coils in the design of protein-based constructs: hybrid hydrogels and epitope displays. *J Control Release*, 2001. **72**(1-3): p. 57-70.

36. Aggeli, A., et al., Exploiting Peptide Self-assembly to Engineer Novel Biopolymers: Tapes, Ribbons, Fibrils and Fibres, in *Self-Assembling Peptide Systems in Biology, Medicine and Engineering*. 2002. p. 1-17.
37. Mahmoud, R., et al., In vivo evaluation of whey protein-based biofilms as scaffolds for cutaneous cell cultures and biomedical applications. *Biomedical Materials*, 2007(1): p. S38.
38. Hwang, D.-C. and S. Damodaran, Synthesis and properties of fish protein-based hydrogel. *Journal of the American Oil Chemists' Society*, 1997. **74**(9): p. 1165-1171.
39. McGrath, K.P., et al., Chemical and biosynthetic approaches to the production of novel polypeptide materials. *Biotechnol Prog*, 1990. **6**(3): p. 188-92.
40. Cappello, J., et al., Genetic engineering of structural protein polymers. *Biotechnol Prog*, 1990. **6**(3): p. 198-202.
41. Petka, W.A., et al., Reversible hydrogels from self-assembling artificial proteins. *Science*, 1998. **281**(5375): p. 389-92.
42. Xu, C., V. Breedveld, and J. Kopecek, Reversible hydrogels from self-assembling genetically engineered protein block copolymers. *Biomacromolecules*, 2005. **6**(3): p. 1739-49.
43. Wright, E.R. and V.P. Conticello, Self-assembly of block copolymers derived from elastin-mimetic polypeptide sequences. *Adv Drug Deliv Rev*, 2002. **54**(8): p. 1057-73.
44. Jing, P., et al., Self-Assembling Peptide-Polymer Hydrogels Designed From the Coiled Coil Region of Fibrin. *Biomacromolecules*, 2008. **9**(9): p. 2438-2446.
45. Yu, Y.B., Coiled-coils: stability, specificity, and drug delivery potential. *Adv Drug Deliv Rev*, 2002. **54**(8): p. 1113-29.
46. Kopecek, J., Hydrogel biomaterials: A smart future? *Biomaterials*, 2007. **28**(34): p. 5185-5192.
47. Lao, U.L., et al., Genetic Engineering of Self-Assembled Protein Hydrogel Based on Elastin-like Sequences with Metal Binding Functionality. *Biomacromolecules*, 2007. **8**(12): p. 3736-3739.
48. Graddis, T.J., D.G. Myszka, and I.M. Chaiken, Controlled formation of model homo- and heterodimer coiled coil polypeptides. *Biochemistry*, 1993. **32**(47): p. 12664-71.
49. Xu, C. and J. Kopecek, Genetically engineered block copolymers: influence of the length and structure of the coiled-coil blocks on hydrogel self-assembly. *Pharm Res*, 2008. **25**(3): p. 674-82.
50. Papapostolou, D., et al., Engineering nanoscale order into a designed protein fiber. *Proceedings of the National Academy of Sciences*, 2007. **104**(26): p. 10853-10858.
51. Papapostolou, D., et al., Electrostatic Control of Thickness and Stiffness in a Designed Protein Fiber. *J. Am. Chem. Soc.*, 2008. **130**(15): p. 5124-5130.
52. Stevens, M.M., et al., Molecular level investigations of the inter- and intramolecular interactions of pH-responsive artificial triblock proteins. *Biomacromolecules*, 2005. **6**(3): p. 1266-71.

53. Arnau, J., et al., Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expression and Purification*, 2006. **48**(1): p. 1-13.
54. Nallamsetty, S. and D.S. Waugh, Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. *Protein Expression and Purification*, 2006. **45**(1): p. 175-182.
55. Andre, P., et al., Splitting cell adhesiveness into independent measurable parameters by comparing ten human melanoma cell lines. *Cell Biophys*, 1990. **17**(2): p. 163-80.
56. Graslund, S., et al., Protein production and purification. *Nat Methods*, 2008. **5**(2): p. 135-46.
57. Hartley, J.L., Cloning technologies for protein expression and purification. *Current Opinion in Biotechnology*, 2006. **17**(4): p. 359-366.
58. David, C.L., Y. Bernard, and A.L. Noshir, Mechanical Properties of a Reversible, DNA-Crosslinked Polyacrylamide Hydrogel. *Journal of Biomechanical Engineering*, 2004. **126**(1): p. 104-110.
59. Biolabs, N.E., pMal Protein Fusion and Purification System. Catalog#E800S, 2007.
60. Expasy, ALDENTE : PEPTIDE MASS FINGERPRINTING TOOL.
61. Bzymek, M. and S.T. Lovett, Instability of repetitive DNA sequences: The role of replication in multiple mechanisms. *Proceedings of the National Academy of Sciences of the United States of America*, 2001. **98**(15): p. 8319-8325.
62. Care, S., et al., The translation of recombinant proteins in *E. coli* can be improved by in silico generating and screening random libraries of a -70/+96 mRNA region with respect to the translation initiation codon. *Nucleic Acids Res.*, 2008. **36**(1): p. e6-.
63. Bolanos-Garcia, V.M. and O.R. Davies, Structural analysis and classification of native proteins from *E. coli* commonly co-purified by immobilised metal affinity chromatography. *Biochim Biophys Acta*, 2006. **1760**(9): p. 1304-13.
64. Blumentals, II, A.S. Robinson, and R.M. Kelly, Characterization of sodium dodecyl sulfate-resistant proteolytic activity in the hyperthermophilic archaeobacterium *Pyrococcus furiosus*. *Appl Environ Microbiol*, 1990. **56**(7): p. 1992-8.
65. Yin Liang Hsien, J.M.R., Modeling Gelation of Egg Albumen and Ovalbumin. *Journal of Food Science*, 1992. **57**(4): p. 856-861.
66. Tomoko Fujiwara, T.M.T.Y.H.Y.S.S.Y.K., Novel Thermo-Responsive Formation of a Hydrogel by Stereo-Complexation between PLLA-PEG-PLLA and PDLA-PEG-PDLA Block Copolymers. *Macromolecular Bioscience*, 2001. **1**(5): p. 204-208.
67. Lee, S.Y., G. Tae, and Y.H. Kim, Thermal gelation and photo-polymerization of di-acrylated Pluronic F 127. *J Biomater Sci Polym Ed*, 2007. **18**(10): p. 1335-53.
68. Xiong, J.Y., et al., Understanding of hydrogel network formation and its application in the architecture of significantly enhanced hydrogel. *Applied Physics Letters*, 2006. **89**(8): p. 083106.

69. Eun S. Gil, R.J.S.S.M.H., Effect of β -Sheet Crystals on the Thermal and Rheological Behavior of Protein-Based Hydrogels Derived from Gelatin and Silk Fibroin. *Macromolecular Bioscience*, 2005. **5**(8): p. 702-709.
70. Gosal, W.S. and S.B. Ross-Murphy, Globular protein gelation. *Current Opinion in Colloid & Interface Science*, 2000. **5**(3-4): p. 188-194.
71. Rudraraju, V.S. and C.M. Wyandt, Rheology of Microcrystalline Cellulose and Sodiumcarboxymethyl Cellulose hydrogels using a controlled stress rheometer: part II. *International Journal of Pharmaceutics*, 2005. **292**(1-2): p. 63-73.
72. Holmes, T.C., et al., Extensive neurite outgrowth and active synapse formation on self-assembling peptide scaffolds. *Proc Natl Acad Sci U S A*, 2000. **97**(12): p. 6728-33.
73. Kane, M.D., et al., Evidence for Seeding of β -Amyloid by Intracerebral Infusion of Alzheimer Brain Extracts in β -Amyloid Precursor Protein-Transgenic Mice. *J. Neurosci.*, 2000. **20**(10): p. 3606-3611.
74. Kayed, R., et al., Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science*, 2003. **300**(5618): p. 486-9.
75. Serio, T.R., et al., Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science*, 2000. **289**(5483): p. 1317-21.
76. Fink, A.L., The aggregation and fibrillation of α -synuclein. *Acc Chem Res*, 2006. **39**(9): p. 628-34.
77. Hong, D.P. and A.L. Fink, Independent heterologous fibrillation of insulin and its B-chain peptide. *Biochemistry*, 2005. **44**(50): p. 16701-9.
78. Khurana, R., et al., Partially folded intermediates as critical precursors of light chain amyloid fibrils and amorphous aggregates. *Biochemistry*, 2001. **40**(12): p. 3525-35.
79. Fink, A.L., Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Fold Des*, 1998. **3**(1): p. R9-23.
80. Herrmann, H., et al., Characterization of distinct early assembly units of different intermediate filament proteins. *J Mol Biol*, 1999. **286**(5): p. 1403-20.
81. Urbanc, B., et al., Molecular dynamics simulation of amyloid β dimer formation. *Biophys J*, 2004. **87**(4): p. 2310-21.
82. Berman, H.M., et al., The Protein Data Bank. *Nucleic Acids Res*, 2000. **28**(1): p. 235-42.
83. Bowie, J.U., R. Luthy, and D. Eisenberg, A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, 1991. **253**(5016): p. 164-70.
84. Xu, J., F. Jiao, and L. Yu, Protein structure prediction using threading. *Methods Mol Biol*, 2008. **413**: p. 91-121.
85. Ding, F., J.J. LaRocque, and N.V. Dokholyan, Direct Observation of Protein Folding, Aggregation, and a Prion-like Conformational Conversion. *J. Biol. Chem.*, 2005. **280**(48): p. 40235-40240.
86. Whitby, F.G. and G.N. Phillips, Jr., Crystal structure of tropomyosin at 7 Angstroms resolution. *Proteins*, 2000. **38**(1): p. 49-59.

87. Guex, N. and M.C. Peitsch, SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, 1997. **18**(15): p. 2714-23.
88. UL Essmann, T.B., H Darden, H Lee, LG Pedersen A smooth Particle Mesh Ewald method. *J Chem Phys*, 1995. **103**: p. 8577-8593.
89. J Ryckaert, G., Ciccotti, HJC Berendsen, Numerical Integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys*, 1977. **23**: p. 327-341.
90. Berendsen, H., JPM Postma, WF van Gunsternen, A DiNola and JR Haak, Molecular Dynamics with coupling to an external bath. *J Chem Phys*, 1984. **81**: p. 3684-2690.
91. Brown, J.H., et al., Deciphering the design of the tropomyosin molecule. *Proceedings of the National Academy of Sciences of the United States of America*, 2001. **98**(15): p. 8496-8501.
92. Minakata, S., et al., Two-Crystal Structures of Tropomyosin C-Terminal Fragment 176-273: Exposure of the Hydrophobic Core to the Solvent Destabilizes the Tropomyosin Molecule. *Biophys. J.*, 2008. **95**(2): p. 710-719.
93. Kuniyiko, G., et al., Ca²⁺ binding sites in calmodulin and troponin C alter interhelical angle movements. *FEBS letters*, 2004. **561**(1): p. 51-57.
94. Smith, A.M., et al., Polar assembly in a designed protein fiber. *Angew Chem Int Ed Engl*, 2004. **44**(2): p. 325-8.
95. Nilges, M. and A.T. Brunger, Automated modeling of coiled coils: application to the GCN4 dimerization region. *Protein Eng.*, 1991. **4**(6): p. 649-659.
96. Suresh, A. and C. Verma, Modelling study of dimerization in mammalian defensins. *BMC Bioinformatics*, 2006. **7 Suppl 5**: p. S17.
97. Gazit, E., Self-assembly of short aromatic peptides: From amyloid disease to nanotechnology. *NanoBioTechnology*, 2005. **1**(3): p. 286-288.
98. Carrotta, R., et al., Conformational characterization of oligomeric intermediates and aggregates in {beta}-lactoglobulin heat aggregation. *Protein Sci*, 2001. **10**(7): p. 1312-1318.
99. Khare, S.D., et al., Molecular Origin of Polyglutamine Aggregation in Neurodegenerative Diseases. *PLoS Computational Biology*, 2005. **1**(3): p. e30.
100. Ma, B. and R. Nussinov, Molecular dynamics simulations of alanine rich {beta}-sheet oligomers: Insight into amyloid formation. *Protein Sci*, 2002. **11**(10): p. 2335-2350.
101. Antwi, K., et al., Cu(II) organizes {beta}-2-microglobulin oligomers but is released upon amyloid formation. *Protein Sci*, 2008: p. ps.073249008.
102. Hartley, D.M., et al., Protofibrillar Intermediates of Amyloid beta -Protein Induce Acute Electrophysiological Changes and Progressive Neurotoxicity in Cortical Neurons. *J. Neurosci.*, 1999. **19**(20): p. 8876-8884.
103. Herrmann, H. and U. Aebi, Intermediate filament assembly: fibrillogenesis is driven by decisive dimer-dimer interactions. *Curr Opin Struct Biol*, 1998. **8**(2): p. 177-85.
104. Er Rafik, M., J. Doucet, and F. Briki, The Intermediate Filament Architecture as Determined by X-Ray Diffraction Modeling of Hard {alpha}-Keratin. *Biophys. J.*, 2004. **86**(6): p. 3893-3904.

105. Athlan, E.S. and W.E. Mushynski, Heterodimeric Associations between Neuronal Intermediate Filament Proteins. *J. Biol. Chem.*, 1997. **272**(49): p. 31073-31078.
106. Paul, A.J., Kinetics of formation of fibrin oligomers. I. Theory. *Biopolymers*, 1982. **21**(11): p. 2253-2264.
107. Bark, N., Z. Foldes-Papp, and R. Rigler, The incipient stage in thrombin-induced fibrin polymerization detected by FCS at the single molecule level. *Biochem Biophys Res Commun*, 1999. **260**(1): p. 35-41.
108. Li, X., D. Galanakis, and D.A. Gabriel, Transient Intermediates in the Thrombin Activation of Fibrinogen. *J. Biol. Chem.*, 1996. **271**(20): p. 11767-11771.
109. Terech, R.G.W.a.P., ed. *Molecular Gels. Materials with Self-Assembled Fibrillar Networks*. 2006, Springer.
110. Mi, L., et al., Self-assembling protein hydrogels with modular integrin binding domains. *Biomacromolecules*, 2006. **7**(1): p. 38-47.
111. Zhiyuan Hu, F.L.Y.P.C.H.L.R.J.C.J.W.Y.Z., Arg-Gly-Asp (RGD) peptide conjugated poly(lactic acid)-poly(ethylene oxide) micelle for targeted drug delivery. *Journal of Biomedical Materials Research Part A*, 2008. **85A**(3): p. 797-807.
112. Tang, W., et al., An efficient system for production of recombinant urokinase-type plasminogen activator. *Protein Expr Purif*, 1997. **11**(3): p. 279-83.
113. Jung, K.-H., Continuous production of recombinant interferon-[alpha] in *Escherichia coli* via the derepression of trp promoter using casamino acid. *Process Biochemistry*, 2006. **41**(4): p. 809-814.
114. Favrin, G., A. Irback, and S. Mohanty, Oligomerization of amyloid A{beta}16-22 peptides using hydrogen bonds and hydrophobicity forces. *Biophys. J.*, 2004: p. biophysj.104.046839.
115. Gsponer, J., U. Haberthur, and A. Caflisch, The role of side-chain interactions in the early steps of aggregation: Molecular dynamics simulations of an amyloid-forming peptide from the yeast prion Sup35. *Proc Natl Acad Sci U S A*, 2003. **100**(9): p. 5154-9.
116. Bernocco, S., et al., Polymerization of rod-like macromolecular monomers studied by stopped-flow, multiangle light scattering: set-up, data processing, and application to fibrin formation. *Biophys J*, 2000. **79**(1): p. 561-83.
117. Philippe Ferrara, J.A.A.C., Evaluation of a fast implicit solvent model for molecular dynamics simulations. *Proteins: Structure, Function, and Genetics*, 2002. **46**(1): p. 24-33.
118. Ting Wang, R.C.W., Implicit solvent models for flexible protein-protein docking by molecular dynamics simulation. *Proteins: Structure, Function, and Genetics*, 2003. **50**(1): p. 158-169.
119. Sigrell, J., Improving the purification of GST Fusion Proteins. *International Biotechnology Laboratory*, 2002.
120. Choy, N., V. Raussens, and V. Narayanaswami, Inter-molecular Coiled-coil Formation in Human Apolipoprotein E C-terminal Domain. *Journal of Molecular Biology*, 2003. **334**(3): p. 527-539.
121. Jahnke, E., et al., Functional, hierarchically structured poly(diacetylene)s via supramolecular self-assembly. *Macromol Biosci*, 2007. **7**(2): p. 136-43.

122. Jayawarna, V., et al., Three-dimensional cell culture of chondrocytes on modified di-phenylalanine scaffolds. *Biochem Soc Trans*, 2007. **35**(Pt 3): p. 535-7.
123. South, C.R., C. Burd, and M. Weck, Modular and dynamic functionalization of polymeric scaffolds. *Acc Chem Res*, 2007. **40**(1): p. 63-74.
124. Kyte, J. and R.F. Doolittle, A simple method for displaying the hydropathic character of a protein. *J Mol Biol*, 1982. **157**(1): p. 105-32.
125. Steve, O.N., et al., Coarse grain models and the computer simulation of soft materials. *Journal of Physics: Condensed Matter*, 2004(15): p. R481.

Curriculum Vitae

CHRISTOPHER L. GAUGHAN

10L Lincoln Pl
 North Brunswick, NJ 08902
 (H) 732-762-6324
 (W) 732-445-6349
 clgaughan68@gmail.com

SUMMARY:

Finding solutions to modern biotechnological problems requires an integrated, multidisciplinary approach. I am a PhD. Biochemical Engineer with a multidisciplinary background that includes recombinant protein production and purification, novel biomaterial fabrication and extensive research in the neurosciences. My educational background is a combination of molecular biology, chemistry and engineering. I believe my broad experience makes me uniquely suited to solving today's complex biotechnical problems.

EDUCATION:

RUTGERS UNIVERSITY, New Brunswick, NJ
Doctor of Philosophy, Chemical and Biochemical Engineering (**October 2008**)

UNIVERSITY OF ALBANY, Albany, NY
B.S., Chemistry (Magnum Cum Laude), Department of Chemistry (**2001**)

SUNY, COLLEGE OF PURCHASE, Purchase, NY
B.S., Biology, Department of Biology (**1991**)

RESEARCH EXPERIENCE:

Chemical & Biochemical Engineering, Rutgers University, NJ (2001-present)

Fellow

Ph.D. Thesis: Fabrication and Modeling of Novel Proteins For Use In Self-Assembling Hydrogels

- Engineered a novel hydrogel composed of four self-assembling peptide units, which can modulate material compliance
- Designed and cloned artificial gene constructs into plasmid vectors in E. coli: optimized culture conditions & throughput
- Expressed and purified artificial proteins in E. coli through optimizing column chromatography
- Developed a method to model and simulate oligomeric intermediates in a peptide fiber-forming pathway using Molecular Dynamics
- Analyzed molecular threading of novel proteins onto pre-existing secondary structures

SKILLS:

Laboratory

- Genetic/ Protein engineering: Designed/cloned artificial gene constructs into plasmid vectors
- Protein purification/DNA purification, SDS Page, Agarose gel electrophoresis, Column Affinity chromatography (Ni²⁺, Co²⁺, Amylose)
- Optimized bacterial culture maintenance
- Animal Surgery (Isolation of primary rat neuronal cells, Sterotaxic surgery)
- Biomaterial generation/characterization: Microrheology and DSC Calorimetry of hydrogels
- UV/Vis, CD spectroscopy of peptides; preparation of peptide samples for electron microscopy

Management

- Supervised multiple undergraduate research assistants in the laboratory

Computer Skills

- Shell programming (bash, csh)
- Software Packages: GROMACS, AMBER, PyMOL, VMD

WORK EXPERIENCE:

Intracellular Therapies, New York, NY (Summer 2005)

Biotechnology Research Fellow Internship

Determined the amount of an experimental drug distributed to the rat Central Nervous System

- Developed pharmacokinetic studies of novel drug uptake and distribution of metabolites in the rat central nervous system
- Assisted in establishing an HPLC isolation scheme of metabolites from rat nervous tissue

Montefiore Medical Center, Department of Anesthesiology, New York, NY (1999-2001)

Research Assistant: Electron Microscopy

Contract Work: Obtained data on ion levels in rat hippocampal slices for use in journal publications.

- Utilized frozen microtomy to prepare frozen sections for electron microscopy analysis

Montefiore Medical Center/Albert Einstein College of Medicine, Department of Neurosurgery and Anesthesiology, New York, NY (1994-1998)

Research Technical Associate

Performed experiments to determine how muscarinic receptor functions changed during the course of Alzheimers disease induced in rats and how the rat hippocampus is damaged on exposure to neurotoxicants.

- Studied Alzheimer's disease and exposure to chronic neuro-toxicants
- Conducted Electron Microscopy: Frozen microtomy; frozen sectioning of rat neuronal tissue and subsequent analysis via TEM/STEM
- Performed Rat Surgery: Sterotaxic surgery on rat brain, isolation of rat hippocampus, striatum, cortex, tibial and femoral nerves
- Carried out experiments on the uptake and subsequent release of radio-isotopically labeled neurotransmitters in fresh dissected hippocampal tissue
- Developed and tested receptor/ligand binding experiments; neurotransmitter release experiments

TEACHING EXPERIENCE:

Department of Chemical and Biochemical Engineering, Rutgers University, Piscataway, NJ (2005)

Teaching Assistant

- Taught gifted high school seniors in the RISE/ISURF summer program. Supervision of biomedical research showcasing the importance of planning and strategically designing experiments, executing complex lab techniques, and performing data analysis.
- Teaching assistant Transport Phenomenon for the undergraduate Chemical and Biochemical engineering
- Ran Discussion Section (60 students), prepared and graded homework problems, mid-term and final exams.
-

NCEES FE Certification Course, Engineering Department, Rutgers University, Piscataway, NJ (2005-Present)

Instructor

- Taught Chemistry Section to post graduate students preparing to take Professional Engineering exam

AWARDS & HONORS:

- NIH Biotech Fellow (2004-2008)
- NJ Spinal Cord Fellowship (2000-2004)

PUBLICATIONS:

- **Gaughan, CL.**, Shreiber, DI. Thermodynamics of Early Oligomer Formation in Fibrillogenesis from Designer Peptides: A Molecular Dynamics Study (in prep)
- **Gaughan, CL.**, Shreiber, DI. A Novel Protein Hydrogel from Four Self-Assembling Peptide Subunits (in prep)
- **Gaughan, CL.**, Shreiber, DI. Molecular Dynamics Simulations of Early Aggregation Events: Implications for Biomaterials and Disease Research. In preparation for Chemical Biology. Nova Science Publishers (in prep)
- LoPachin, RM., Schwarcz, AI, **Gaughan, CL**, Mansukhani, S, Das, S., (2004) In vivo and in vitro effects of acrylamide on synaptosomal neurotransmitter uptake and release. *Neurotoxicology*, **25**(3): p. 349-63.
- LoPachin, RM., **Gaughan, CL**, Lehning, EJ, Weber, ML, Taylor, CP. (2001), Effects of ion channel blockade on the distribution of Na, K, Ca and other elements in oxygen-glucose deprived CA1 hippocampal neurons. *Neuroscience*, 2001. **103**(4): p. 971-83.
- Taylor, CP, Weber, ML, **Gaughan, CL**, Lehning, EJ, LoPachin, RM, (1999) Oxygen/glucose deprivation in hippocampal slices: altered intraneuronal elemental composition predicts structural and functional damage. *J Neurosci.*, **19**(2): p. 619-29.
- Potter, PE., **Gaughan, CL** and Assouline, Y., (1999). Lesion of septal-hippocampal neurons with 192 IgG-saporin alters function of M1 muscarinic receptors. *Neuropharmacology*, **38**(4): p. 579-86.
- LoPachin, RM and **Gaughan, CL.**, Electron probe X-ray microanalysis: A quantitative electron microscopy technique for measurement of elements and water in nervous tissue cells. In: J. Harry and H. A. Tilson, Editors, *Methods in Molecular Medicine*, Humana Press, Totowa (1999), pp. 289–299.
- LoPachin, RM., **Gaughan, CL**, Lehning, EJ, Kaneko, Y, Kelly, TM, Blight, A (1999)., Experimental spinal cord injury: spatiotemporal characterization of elemental concentrations and water contents in axons and neuroglia. *J Neurophysiol*, **82**(5): p. 2143-53.
- Lehning, EJ., **Gaughan, CL**, Eichberg, J, LoPachin, RM, (1997) Rubidium uptake and accumulation in peripheral myelinated internodal axons and Schwann cells. *J Neurochem.*, **69**(3): p. 968-77.
- Potter, PE., Thorne, B., and **Gaughan, CL.**, (1997). Modulation of hippocampal norepinephrine release by cholinergic agonists is altered by AF64A lesion. *Brain Res Bull*, **42**(2): p. 153-60.
- Lehning, EJ., **Gaughan, CL.**, and LoPachin, RM (1997) Acrylamide intoxication modifies in vitro responses of peripheral nerve axons to anoxia. *J Peripher Nerv Syst.*, **2**(2): p. 165-74.

CONFERENCE PRESENTATIONS

- **Gaughan, CL., Shreiber DI.** "Design of a Self-Assembling α -helical Peptide Hydrogel" Presented at the Third Annual Biomedical Engineering Showcase, Woodbridge, New Jersey: **March 2006.**
- **Gaughan, CL., Shreiber DI.** "Simulation of 28 amino acid α -helical peptides" Presented at the 2006 Annual BMES Conference, Poster Session 5, Chicago, Illinois: **October 2006.**

- **Gaughan, CL., Shreiber DI.** "Simulation of 28 amino acid α -helical peptides" Presented at the 2005 Annual BMES Conference, Poster Session 5, Baltimore, Maryland: **October 2005.**
- **Gaughan, CL., Shreiber DI.** "Simulation of 28 amino acid α -helical peptide dimers" Presented at the Mid-Atlantic Biochemical Engineering Consortium Conference, Raleigh, North Carolina: **May 2005.**
- **Gaughan, CL., Shreiber DI.** "Expression of Artificial Proteins in E. coli for the Fabrication of A Novel Hydrogel" Presented at the Mid-Atlantic Biochemical Engineering Consortium Conference, Raleigh, North Carolina: **May 2005.**