Structural Control Concepts in Protein Nanofiber Networks
Liang Chang, Ongi Englander, William Oates and Anant Paravastu
FAMU-FSU College of Engineering
Tallahassee, FL 32310
{englander, woates, paravastu}@eng.fsu.edu

Abstract: It is known that certain proteins self-heal after being exposed to ultrasonic radiation while others do not which has implications on guided assembly of nanoscale structural and electronic networks. For example, the RADA-16I protein contains hydrogen bonds between protein molecules which can be reversibly broken and re-formed. Most protein nanofiber systems (e.g., collagen or amyloid fibrils) are formed by self-assembly but are not known to self-repair without addition of new material. Here we explore fundamental differences in these cases from a thermodynamic continuum approach as well as experimental methods of nanofiber synthesis and characterization. Some governing relations are obtained for fracture and self-healing of hydrogen bonds and the stronger, irreversible covalent bonds. This provides relations to guide forces necessary to break hydrogen bonds while keeping each fibril intact. Correlation of molecular dynamic and continuum fracture mechanics provides insight on key parameters governing variations in intrinsic bonding for proteins that spontaneously self-heal after fracturing hydrogen bonds, and will be connected with experimental measurements of nano-scale mechanical properties and dynamic reassembly kinetics. Preliminary experimental progress includes protein synthesis and self-assembly and electron microscopy-based imaging of nanostructures.

Keywords: proteins, RADA-16I, self-healing, hydrogen bond, molecular dynamic simulations, nonlinear continuum fracture mechanics

1. Introduction
Naturally occurring self-assembled biomaterials such as the bones, shells, and teeth are some of the most impressive materials known in terms of their toughness, stability, and light weight. Their remarkable properties are a consequence of high-fidelity protein-mediated self-assembly of their (typically inorganic) nanostructured building blocks. As proteins are hetero-polymers of up to 20 different amino acids, they can be optimized (by rational design or evolution) for highly specific interactions with other proteins or across interfaces. In principle, a new generation of nanostructured material could be templated by an ordered matrix of self-assembled designer proteins. The versatility of proteins would allow an unlimited variety of protein-nanostructure interactions, and a bio-templated approach to construction could lead to an unlimited variety of achievable mechanical and electrical properties.

Our work is aimed at achieving long-range order in a new class of adaptive self-assembling protein scaffolds for next-generation aerospace materials. The proteins we study are known to self-assemble into nanomaterials themselves. The nanostructures they form are characterized by the “cross-β” structural motif [1]. Cross-β refers to filament formation (10 nm wide and microns long) through the arrangement of individual protein molecules into extended β-strand conformations, and self-assembly through intermolecular backbone hydrogen bonds and sidechain interactions. In addition to self-assembling domains, novel designer proteins could eventually include surface-recognizing domains for highly specific interfaces with other types of nanomaterials, including carbon nanotubes, metal nanowires, and silica particles [2-4].

Our proposed novel approach to scaffold alignment is inspired by the self-healing or adaptive self-assembly of proteins. Self-assembled protein filaments could be destroyed by mechanical forces, creating fragments which can themselves self-assemble to regenerate filaments. We hypothesize that the anisotropic response of a protein filament to mechanical forces or waves could drive macroscopic alignment of a self-assembled protein matrix to a field of standing waves (Figure 1). Thus, we seek to understand the optimal acoustic stimulation for guidance of self-assembly. While it is common laboratory practice to destroy protein assemblies by fluid sonication, we know of no systematic studies on the sensitivity of protein self-assembly to mechanical vibrations in solution. Recent reports that amyloid fibrils have an elastic modulus comparable to steel [5] with geometries on the order of 1 µm in length, diameter 10 nm, and density comparable to water, leads to a resonant frequency on the order of 100 kHz. This is about an order of magnitude larger than typical sonicators. Instead of random acoustic transmission using a sonicator, here we propose to develop a device that can generate a standing wave in water using piezoelectric transducers; see [6] for a similar system.
To date, our experimental work has focused on protein synthesis and imaging while the theory and computation effort has focused on nonlinear fracture mechanics of protein structures. This new knowledge will serve to guide the development of acoustic wave devices to create an anisotropic protein structure.

![Figure 1: a) A sample structure of a cross-beta protein filament [7], b) Protein self-assembly in bulk solution leads to randomly oriented filaments. c) Ultrasonic waves may preferentially disrupt protein assemblies oriented in specific directions. Fragments produced could self-assemble into new filaments oriented differently. d) Self-assembly in the presence of standing ultrasonic waves may lead to long-range orientational order.]

2. Background

Proteins are polypeptides where the degrees of freedom are restrained by the establishment of H-bonds. These H-bonds make the corresponding conformations more stable than other ones. Such stabilized conformations are called the “secondary structures” of the protein, the primary structures being defined as the succession of the constituent amino acids of the “polypeptide backbone” of the protein.

Two characteristic secondary conformations, α-helices and pleated β-sheets were defined in 1951 by Pauling et al [8, 9] after they had determined the structure of a particular protein. These secondary conformations have been later found in almost all other proteins. The presence along the succession of amino acids of ordered α-helices and β-sheets is the criterion that allows us to distinguish a protein from a polypeptide. It is acquired during the synthesis of proteins, following mechanisms that are not yet entirely understood. A given protein has a well-defined secondary structure. Breaking the secondary structures of a living protein, a process called denaturation, which may be provoked by changing the neutral pH of the bio-medium, heating it, dehydrating it, etc.

There are of two sorts of pleated β-sheets structure: parallel and antiparallel. H-bonds are in less constrained configurations in these antiparallel β-sheets than the H-bonds of parallel β-sheets, as they can be in their preferred linear configuration with bending. Furthermore, the “turns” that separate two sequences of such a conformation can be much shorter than the sequences that separate the sets of peptides that establish parallel β-sheets[10]. Therefore, the antiparallel β-sheets have a tendency to be consequently more stable and more easily formed than parallel ones. H-bonds still exist in the α-helices, the axes of helices are roughly parallel to H-bonds. Most helix proteins take the right-handed helix structures. These α-helix structures are more flexible than β-sheets but mechanically less resistant. β-sheets are much stiffer.

3. Experimental Results

Protein synthesis

We have synthesized three self-assembling proteins using fluorenylmethyloxycarbonyl (FMOC) chemistry: the Alzheimer’s β-amyloid protein (Aβ), RADA-16I, and MAX8. The ability of Aβ to exhibit triggered nanofiber self-assembly in the presence of pre-formed fibril seeds is critical to our proposed approaches to controlled nanostructure assembly. Furthermore, the unstructured N-terminal domain of Aβ could provide a versatile platform for surface decoration. RADA-16I is known to exhibit dynamic reassembly or self-healing. MAX8 is reported to exhibit salt-triggered assembly. Amino acid sequences and our achieved purities for these proteins are shown in Table 1.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Amino Acid Sequence</th>
<th>Achieved Purity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer's β-amyloid (Aβ)</td>
<td>DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV</td>
<td>80%</td>
</tr>
<tr>
<td>RADA16-1</td>
<td>Ac-RADARADARADARADARADARADA-NH₂</td>
<td>Unknown</td>
</tr>
<tr>
<td>MAX-8</td>
<td>VKVKVKVKP*PTKVEVKVKV-NH₂</td>
<td>&gt;95%</td>
</tr>
</tbody>
</table>

(Ac-): N-terminus is acetylated; (-NH₂): C-terminus is amidated; (*): Un-natural D-type chirality.
Protein synthesis was monitored by UV/VIS spectroscopy of FMOC protecting groups effluent to the peptide synthesis reaction vessel. Figure 1a shows data for our initial unsuccessful Aβ synthesis. Synthesis started from the C-terminus (residue #40). High absorbances in the green bars indicate that protein polymerization steps were successful. Other colored bars for the same amino acid indicate multiple deprotection attempts; ideally these signals would be low if first-attempt deprotections are successful. As polymerization proceeded towards the N-terminus (lower residue #), declining absorbance of the green bars indicates failed polymerization. We subsequently better optimized Aβ synthesis by lowering overall scale and using a synthesis resin with a reduced substitution level (Fig. 2b). Similar data for RADA-16I and MAX8 (not shown) indicate successful syntheses.

![Absorbance of FMOC protecting groups after each deprotection step](image)

**Fig. 2.** Absorbance of FMOC protecting groups after each deprotection step. Different colored bars for each amino acid refer to repeated deprotection attempts. Synthesis starts at the C-terminus (amino acid #40). a) Results for a failed synthesis; decay of absorbances indicated by green bars corresponds to declining polymerization efficiencies.

Synthesized proteins were purified using high performance liquid chromatography (HPLC). Figure 3(a) is a sample dataset from UV/VIS monitoring of protein exiting the HPLC column. Composition of HPLC fractions were determined using mass spectrometry (Figure 3(b)). We found significant amount of oxidized (most like at M35) Aβ in the product and were able to purify to only 80% desired product. We are currently exploring methods of reversing or avoiding M35 oxidation.

![High performance liquid chromatography data for synthesized Aβ](image)

**Fig. 3.** (a) High performance liquid chromatography data for synthesized Aβ. (b) Mass spectrometry data for HPLC purified Aβ. Signals marked by check marks correspond to desired product. There is also signal (+18 mass) corresponding to oxidized protein (20% impurity).

**Protein Characterization**

We have used TEM imaging to detect the dynamic reassembly of RADA-16I nanofiber networks, as shown in Figure 4. Sonication breaks nanofiber networks into small protein oligomeric particles (Fig. 4a). With time, these particles are able to reassemble into nanofibers without re-addition of fresh protein (Fig. 4b). We also found unexpectedly that nanofiber thickness also changes with time (Figure 4c). In this experiment, all three images represent the same sample at different points of time following sonication.
RADA-16I self-assembly and growth has been evaluated using light scattering techniques. Dynamic light scattering (DLS) techniques are typically used to determine the size of spherical particles in solution. In DLS, a laser beam is pointed onto a sample and intensity of scattered light of a sample undergoing Brownian motion is recorded. Light scattering data is well suited for monodispersed samples as the results are dominated by the largest species in solution. An example of light scattering data ($\Gamma$ coefficient) as a function of time is seen in Fig. 5. The $\Gamma$ coefficient is related to the diffusivity of the protein nanofibers ($D_t$) and a scattering vector ($q$) as described by Equation 1. Further, the diffusivity can be correlated to the changing length of the protein nanofibers as described by Equation 2 where $k$ is Boltzmann’s constant, $T$ is the solution temperature, $\eta$ is the viscosity of the suspending medium and $d$ is the diameter of the nanofibers as measured by TEM ($d \approx 5$nm).

We may conclude that the RADA-16I samples are polydisperse and include larger species and/or aggregated species. These results are consistent with the SEM images and reassembly studies performed by TEM. Ongoing experiments are focused on evaluating the effect of sample sonication and filtering processes (designed to the break up and remove larger species) on the DSL data and resulting nanofiber growth parameters.

\[
\Gamma = D_t q^2
\]

\[
D_t = \frac{kT}{3\pi\eta L} \left[ \ln \frac{L}{d} + 0.312 + 0.565 \frac{d}{L} + 0.100 \left( \frac{d}{L} \right)^2 \right]
\]
The following sequence of steps is used to prepare samples for SEM imaging. Small volumes (10-30µL) of protein samples are pipetted onto a glass slide. The samples are immediately placed in liquid nitrogen (2 seconds) to halt further changes to the sample (growth, aggregation). Next, the frozen glass slides are lyophilized (dry-freeze) overnight to dehydrate and remove the solvent from the samples. With the addition of a thin gold coating, the lyophilized samples can be imaged using the SEM. Since the samples may be frozen as a function of time, studies of sample dynamics can also be performed using this technique. An optical image of a prepared RADA-16I SEM sample is seen in Fig. 6 (without gold coating).

The results of the SEM imaging based characterization of the RADA-16I nanofibers are seen in Fig. 7. We note very fine features as well as the formation of larger aggregates and sheet-like features. The formation of the larger sheet-like features was not expected and may be due to the sample preparation process. Scanlon et. al. have suggested that during the lyophilization process, ice formation leads to the formation of large protein nanofiber accumulations at ice crystal edges as the nanofibers are forced to these locations [11, 12]. The results also illustrate three dimensional and out-of-plane aggregation which is not commonly characterized by AFM and TEM techniques. Upon imaging the samples at different focal planes, we observe that the smallest particulates are found at the base of the samples while larger features are found near the top (Fig. 7). We speculate that the smaller features can penetrate the dense networks.

We have focused on the evaluation of our sample preparation protocols as we are interested in understanding the aggregation of the proteins into thick aggregates and sheet-like structures. This is due to the observed aggregation in specimen preparation. Since we have previously limited our SEM samples to RADA-16I samples, we decided to test Aβ samples, prepared in the same manner. Aβ protein nanofibers have never been documented to form sheets and the current literature suggests that their self-assembly and growth mechanisms should not support the formation of sheet-like structures. SEM images of Aβ samples prepared via dry freezing and thin gold coating are seen in Fig. 8a. Large aggregates and sheet-like features are observed with the Aβ sample as well. We also suspected that the application of the gold coating could impact the sample. Therefore, we prepared RADA samples and coated with a thin iridium layer. The results are shown in Figure 8b, again suggesting the formation of large aggregates. As with the other samples, this sample was also prepared following sonication which has been documented to break proteins into fragments.

Since the dry freezing process may be responsible for the aggregation behavior, ongoing sample preparation processes include allowing the sample to dry in air and using a hot plate (at 30-50°C) to dehydrate the samples. Finally, since protein aggregation is known to take place as part of protein self-assembly and growth, it is possible that aggregation during sample preparation contributes as well. To investigate this possibility, the samples will be filtered, using 200nm-pore filter, immediately before the dry freezing process to ensure that large aggregates are not present within the sample.
3. Theory and Computational Modeling

Molecular Dynamic Modeling

To understand the molecular mechanism of self-assembly and reassembly of protein involving RADA16-I and MAX-8, we simulate the process of separation between β strands. The 16-residue peptide forms a stable β-sheet structure and undergoes molecular self-assembly into nanofibers. We asked whether the self-assembled nanofibers could undergo reassembly after mechanical breakage through sonication. Sonication usually breaks the weak chemical bonds including hydrogen, ionic bonds, and hydrophobic interactions, but not the covalent peptide bonds, thus this process may allow reassembly of the mechanically disassembled individual peptides [13]. It is known that the RADA16-I forms a very stable β-Sheet [14]. It is therefore meaningful to focus on the β-Sheet structure and its “fracture” mechanism.

A single protein titin I27 module is an ideal test model as shown in Fig. 8(left), due to its stable mechanical properties, the two beta strands could be taken from the protein and considered as the unit block acting as a mechanical clamp in a large protein structure. An assembly of H-bonds displayed as red dash lines in Fig. 9(right) is the smallest and simplest subsystem that can be considered in the unfolding of a protein domain. We model the two layer beta-strand protein building block as a polypeptide chain stabilized by H-bonds, force the alpha carbon 78 atom on three different directions, which is very similar to the classic three fracture modes as shown on Fig. 9(right). This provides a detailed look at the H-bond breaking mechanism, its effect on the whole protein system, and input for continuum fracture and damage modeling of micro to mesoscale protein networks. Note that all three of these modes would occur in the presence of random acoustic waves.

As shown in Fig. 9 (right), the alpha carbon atom 78 is pulled in three different directions which are almost perpendicular to each other, the constant pulling velocity is 1 Angstrom/ pico-second and lasts 30ps, the output trajectories obviously display the stretch and rupture processes. Displacement and force of the atom in three distinct modes are recorded. The cases of Mode I and II are shown in Fig. 10. Force reactions of these modes show obvious nonlinear responses and several rapid drops in force and corresponding increases in displacement which is indicative of bond rupture and “fracture” of hydrogen bonds. It should be noted that the force for bond rupture is lower for
Mode I than Mode II fracture. It is shown in the following section that these predictions correlate well with nonlinear continuum fracture mechanics.

Continuum Fracture Model

We correlate the molecular dynamic simulations with nonlinear continuum fracture mechanics to give additional insight on the differences in fracture strength with respect to load. Since these proteins are characterized by geometries with large aspect ratios of length versus diameter, we investigate driving forces to break the weak hydrogen bonds between proteins using a double cantilever beam model. We show that the driving force for crack propagation is strongly dependent on the direction of load. This provides important relations that can be used to guide self-assembly and alignment of these structures.

The fracture mechanics of solids is a well known and details on the theory is referred to the literature [15] [16] [17]. Prior to giving quantitative results, it is interesting to introduce Rice’s model which states that crack extension is governed by

\[ g \nu = (G - R )\nu \geq 0 \] (3)

where \( g \) is the crack extension force, \( G \) is the applied energy release rate, \( R \) is the fracture toughness, and \( \nu \) is the crack tip velocity. The inclusion of entropy via the second law of thermodynamics within this model provides a more general fracture relationship beyond the classic Griffith concept that states cracks will grow once \( G=R \). Here, kinetic relations are introduced which state that if \( g > 0 \), then \( \nu \geq 0 \) and the crack moves forward. If \( g < 0 \), then \( \nu \leq 0 \) and the crack moves backwards. This process places no restrictions on the rate of self-healing observed in certain proteins and can be used to identify regions of crack growth or healing as a function of load and geometry.

The double cantilever beam model is used to estimate the effect of separating two proteins attached by weak hydrogen bonds. A comparison of Mode I (opening model) and Mode II (shear mode) loading is given. It is shown that the driving force for fracture or separation of proteins is strongly dependent on the direction of loading, crack length, and aspect ratio of the proteins. A hyperelastic constitutive model is introduced to approximate the stress-
strain behavior of the proteins using a generalized elastomer model. The mechanical energy of the effective protein model is

\[
\psi(\lambda_i) = \sum_{p=1}^{N} \frac{\mu_p}{\alpha_p} \left( \lambda_1^{\alpha_p} + \lambda_2^{\alpha_p} + \lambda_3^{\alpha_p} - 3 \right) + p(I_3 - 1)
\]

(4)

where \(\lambda_i\) is the material stretch (\(i=1,2,3\)), \(\mu_p\) are the shear moduli, and \(\alpha_p\) is a set of fitting parameters. The third strain invariant is \(I_3 = \lambda_1 \lambda_2 \lambda_3\) and the undetermined Langrange multiplier is \(p\). This second term enforces incompressibility which is typical of elastomer or gel networks.

Since this energy function is nonlinear, we use the J-integral to quantify crack tip driving forces for Mode I and II loading. The J-integral is

\[
J = \int_{\Gamma_0} \left( \psi n_x - T_i u_{i,X_i} \right) dS
\]

(5)

where \(T_i\) is the traction and \(u_{i,X_i}\) is the displacement gradient in the \(X_i\) direction. Note that these relations are written in the reference configuration. Since this integral is path independent, closed form solutions can be obtained by selecting certain paths around the crack tip.

The following loading is used to quantify differences in crack tip driving forces as a function of loading. In Mode I, a uniform displacement is applied along the length of the proteins resulting in opening mode fracture between the two protein fibrils. In Mode II, displacement is applied on the protein ends in opposite directions resulting in shear loading at the crack tip. The differences in driving forces for these two cases are

\[
J_I = 2h \sum_{p=1}^{N} \frac{\mu_p}{\alpha_p} \left( 2 \lambda_2^{\alpha_p/2} + \lambda_2^{\alpha_p} - 3 \right)
\]

(6)

\[
J_{II} = 2h \left( \sum_{p=1}^{N} \frac{\mu_p}{\alpha_p} \left( 2 \lambda_2^{\alpha_p/2} + \lambda_2^{\alpha_p} - 3 \right) + \sum_{p=1}^{N} \mu_p \left( \lambda_1^{\alpha_p-1} - \lambda_1^{-(\alpha_p/2+1)} \right) \right)
\]

(7)

for Mode I and Mode II, respectively, where \(\lambda_1\) and \(\lambda_2\) is the stretch in the \(X_1\) and \(X_2\) directions, respectively. The parameter \(h\) is the nominal diameter of the protein and \(a\) is length of the crack which is limited to a length less than the proteins. Prior to plotting the driving force and load versus displacement for a given set of parameters, we linearized the constitutive relation by assuming linear elastic properties. If we enforce \(J_I > J_{II}\), this requires \(a/h > \sqrt{3}\) which is a reasonable constraint based on typical aspect ratios of protein fibrils.

Lastly, we implement a set of nonlinear material parameters to quantify differences in crack tip driving forces for Mode I and II. As illustrated in Figure 11, it is shown that the driving forces for Mode I fracture or hydrogen debonding as the material is loaded above a critical stretch value. This illustrates that if the H-bond strength is above this critical stretch, it will lead to significantly different forces necessary to break hydrogen bonds. The result correlates well with molecular dynamic simulations shown in Figure 10 since bonds break easier in Mode I than Mode II.

![Figure 11: (a) J-integral result for Mode I and II loading illustrating larger driving forces for Mode I when the displacement is greater than ~0.5 nm. (b) The corresponding stress versus displacement. Note that the nominal stress versus displacement correlates well with recent experiments given in the literature on similar proteins.](image-url)
4. Discussion and Future Work
The central aim of this research is to manipulate nanostructure and structural order at multiple length scales using ultrasonic excitation of protein nanofiber systems. With the anticipated level of control, we aim to develop a novel basis for construction of light, tough aerospace materials and sensors that depend on controlled arrangement of nanodevices. Towards this end, we have demonstrated our ability to a) synthesis proteins, b) measure and improve protein purities by HPLC and mass spectrometry, c) characterize 2-dimensional and 3-dimensional protein nanofiber structures in by TEM and SEM, respectively, d) measure self-assembly kinetics by TEM and light scattering, and e) model protein fracture using molecular dynamics and nonlinear continuum fracture mechanics. The next step will be to analyze the effects of anisotropic ultrasonic excitation on protein nanofiber systems using a custom made prototype ultrasonic resonator illustrated in Figure 12.

![Fig. 12 The custom made prototype ultrasonic resonator](image)

The theoretical and computational molecular dynamic simulations provide promising results on achieving anisotropic structures. For example, the continuum model illustrated in Figure 11 shows that the driving force to open a “crack” or separate two proteins is different depending on the mode of fracture. If the intrinsic toughness is above a critical value, Mode I fracture gives a significantly larger crack tip driving force. This driving force in Figure 11(a) is different for each mode of fracture for displacement > 0.5 nm. Note that this is based on the same applied force defined by the nonlinear force-displacement relation in Figure 11(b). It should be noted that the continuum fracture mechanics model only provides driving forces for crack propagation. It says nothing about bond rupture. We implement molecular dynamic simulations to quantify the bond rupture process. It is shown that lower external loads under Mode I loading are necessary to rupture H-bonds which correlates with the continuum model. The results are promising for future acoustic stimulation experiments. Ideal standing waves may not be necessary if a critical power can be obtained that leads to bond rupture from Mode I loading but does not break bonds from Mode II. Current work is focused on obtaining quantitative model predictions of molecular protein structures used in the experiments, Mode III modeling, and experiment acoustic modeling.

5. Acknowledgments
This work was supported by the Florida Center for Advanced Aero-Propulsion (FCAAP). We thank Dr. S. Ramakrishnan for access to and assistance with the light scattering system.

References