Summer 2021

Stimuli Responsive Dye-Containing PEG-PLA Block Copolymer Micelles and Computationally Assisted Design of a Stapled Peptide Bundle

Tyler L. Odom
Missouri State University, Odom515@live.missouristate.edu

STIMULI RESPONSIVE DYE-CONTAINING PEG-PLA BLOCK COPOLYMER MICELLES AND COMPUTATIONALLY ASSISTED DESIGN OF A STAPLED PEPTIDE BUNDLE

A Master's Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Chemistry

By

Tyler Lee Odom

July 2021

STIMULI RESPONSIVE DYE-CONTAINING PEG-PLA BLOCK COPOLYMER MICELLES AND COMPUTATIONALLY ASSISTED DESIGN OF A STAPLED PEPTIDE BUNDLE

Chemistry

Missouri State University, July 2021

Master of Science

Tyler Lee Odom

ABSTRACT

In this thesis, I report the preparation and characterization of dye-containing PEG-*b*-PLA block copolymer micelles and the computational design of a novel coiled-coil peptide bundle. The PEG-*b*-PLA micelles encapsulate hydrophobic molecules into their core and have strong potential as nanocontainers or delivery vesicles. In theory, these internalized molecules can be released upon exposure to mechanical forces that disrupt the micellar structures. This force-responsive nature is one of the inherent properties of micellar systems. However, there is a stark lack of research that utilize this property in applications. Along those lines, I have studied the behavior of dye-containing micelles upon exposure to mechanical stresses. In this work, two fluorescent dyes that undergo fluorescence resonance energy transfer (FRET) in close proximity have been used as reporters to gauge the structure of the micelles. Upon exposure to mechanical force *via* vortexing, the fluorescent signal of micelles decreases as a function of rotational speed and time of vortexing. In another project, I have worked on the design of a novel coiled-coil peptide bundle. Coiled-coil peptides have been widely used as templates for protein design. The specific goal of this project was to develop a photo-responsive hexameric coiled-coil bundle that can incorporate hydrophobic molecules into the interior core and reversibly catch and release small molecules. The azobenzene derivative, BSBCA was chosen as the linker to impart the photo-responsive properties to the peptide bundle. Using the PyRosetta computational suite, I have identified 3 top candidates of the mutated peptides that contain two cysteine residues to be linked with BSBCA and are anticipated to form stable hexamer bundles in aqueous environment.

KEYWORDS: force responsive materials, FRET, block copolymers, self-assembly, coiled-coils, protein design, stapled peptides

STIMULI RESPONSIVE DYE-CONTAINING PEG-PLA BLOCK COPOLYMERS AND COMPUTATIONALLY ASSISTED DESIGN OF A STAPLED PEPTIDE BUNDLE

By

Tyler Lee Odom

A Master's Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Chemistry

July 2021

Approved:

Keiichi Yoshimatsu, Ph.D., Thesis Committee Chair

Cyren Rico, Ph.D., Committee Member

Fei Wang, Ph.D., Committee Member

Christopher Lupfer, Ph.D., Committee Member

Julie Masterson, Ph.D., Dean of the Graduate College

ACKNOWLEDGEMENTS

There are many people I would like to thank for helping with the work presented in this thesis, both for the wise research advice and friendship they provided. First, I would like to thank my research advisor Dr. Keiichi Yoshimatsu. He has served as a wonderful mentor for the years I have worked with him. He has continually pushed me and challenged me to become a better student and researcher. He was always willing to sit down and help me with my teaching, classwork, or my endless research questions I asked. The assistance and advice with my graduate school applications was also an enormous help that I cannot be grateful enough for.

There are many other professors that have played a pivotal role in the progression of this work. I would like to thank all the member on my thesis committee: Dr. Cyren Rico, Dr. Fei Wang, and Dr. Christopher Lupfer, for taking their time to give me constructive and valuable feedback to my thesis. I would also like to thank Dr. Reza Herati and Dr. Gary Meints. Dr. Herati was always available to provide advice on organic synthesis and lent us some crucial equipment we needed. Dr. Meints gave us some valuable help and advice with NMR which I am very grateful for. Dr. Risheng Wang and Dr. Wenyan Liu were also gracious enough to help us with TEM imaging and let them use their facilities.

I am truly thankful for the friendship and help many of my peers provided. I would like to thank Collin Johnson and Megan Westwood for help with NMR as well as some programming advice. I would also like to thank the members of Yoshimatsu's lab for making the lab a fun and enjoyable place to spend my time over the years: Luckio Owoucha, Adjoa Adams, Devin Mart, Giselle Campos, and Alex Babel. Lastly, I want to shout out Devin and Giselle for helping me with the micelle project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

OVERVIEW

Materials that have the ability to predictably respond to external stimuli are commonly termed ‘smart materials’. In my Master thesis work, I worked on the development of two types of smart materials. The first chapter of this thesis describes the characterization and development of a force-responsive dye-containing block copolymer micelle system. The second chapter of this thesis focuses on the design of a photo-responsive coiled-coil peptide bundle. While the chemical composition of these compounds and the approaches used to characterize them are different, there are several fundamental similarities. Both compounds consist of several independent chains that spontaneously self-assemble; the self-assembly process is primarily driven by hydrophobic interactions. At last, both of these smart materials have the potential to incorporate small, hydrophobic molecules into their core, offering the potential for being used in broad application areas as a new class of nanocontainers or delivery vehicles. In this thesis, the efforts toward designing and developing each of these smart materials are described in each of chapters, respectively.

STIMULI RESPONSIVE DYE-CONTAINING PEG-PLA BLOCK COPOLYMER MICELLES

Introduction

Stimuli Responsive Polymers. Virtually an unlimited number of possible polymeric chemical structures and monomer compositions can be synthesized, which has given rise to an enormous number of applications.[1] Polymers that respond predictably to external stimuli have been growing in interest in recent years for their unique properties and applications spanning multiple disciplines. These stimuli-responsive polymers, also known as 'smart polymers', are polymer-based materials whose chemical or physical properties will change upon exposure to a stimulus such as pH, temperature, mechanical force, or electric fields.[2–5] Depending on the composition of the polymer and local environment they may show sensitivity to multiple stimuli simultaneously, or only respond to one.[6] These materials are growing in interest in engineering for their uses as self-healing materials and nanocontainers.[7–9] Stimuli-responsive materials have also been designed to mimic biological systems or interact with them and have been developed as biosensing nanoprobes and drug delivery vesicles.[1,10]

Force Responsive Polymers. Force-responsive polymers are a subset of stimuli responsive materials whose properties change upon exposure to mechanical force. External force is commonly applied to these materials through sonication, elongation, or flow fields.[11–13] Davis *et al.* has reported a polymeric material that will undergo selective covalent bond cleavage upon stretching of the polymer solid which ultimately induces a visible change in color.[12] As seen in Figure 1, applied force causes scission of the strained C-O bond in spiropyran which forms the highly conjugated merocyanine. The conjugated nature of merocyanine causes it to absorb light

in the visible region and gives it its characteristic deep purple color. This process is reversible upon exposure to UV light.[12] There are several clear applications of these materials in engineering.[14] For example, incorporating force-sensitive polymers into structural materials that will undergo visible change upon onset of structural degradation, indicating that repair is needed.

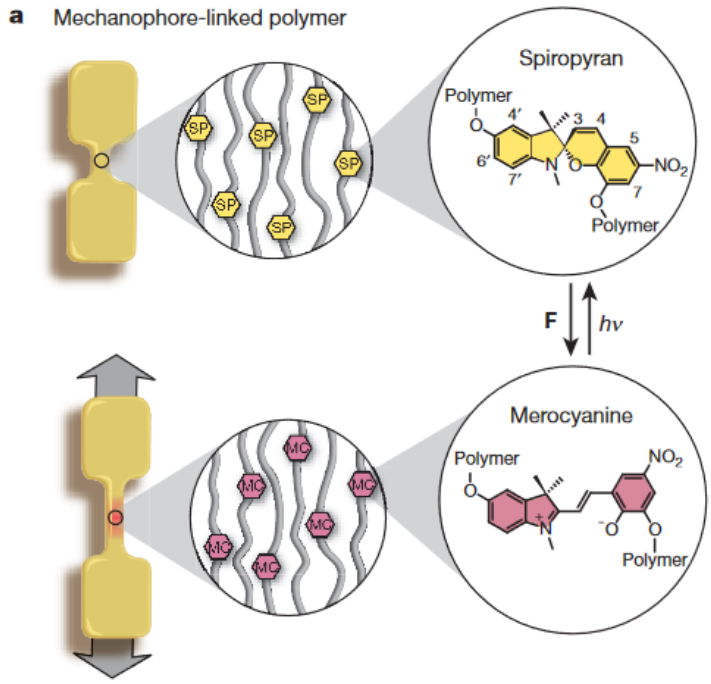

Figure 1. Selective covalent bond scission initiated via manual elongation which induces a visible color change.

Block Copolymer Micelles. Self-Assembly**.** Amphiphilic block copolymer micelles are nanoparticles composed of aggregated block copolymer polymer chains. This class of polymers has two or more discrete 'blocks' in each chain, with one block being hydrophilic and the other being hydrophobic.[15] When placed into aqueous media, these compounds spontaneously self-assemble into spheres, cylinders, vesicles, and many other shapes depending on the chemical structure, concentration, solvent, and temperature.[16] The process of self-assembly is primarily

driven by the interactions between hydrophobic blocks.[17] Figure 2 shows the self-assembly of block copolymers to form a micelle.

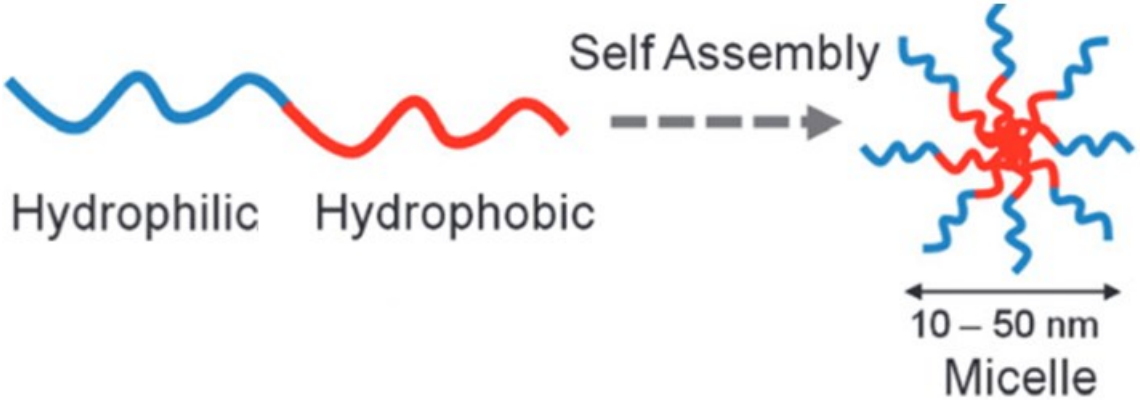

Figure 2. Self-assembly of block copolymer micelles to form spherical micelles in aqueous environment. Figure adapted from work by Johnson *et al*.[17]

An important characteristic of block copolymers is the concentration required to form micelles, known as the critical micelle concentration (CMC). Below this concentration micelles will not readily form as the amphiphilic molecules are more thermodynamically stable in aqueous phase or the water-air or water-organic interface.[17] However, in some systems, once micellar structures are formed they may retain their structure even after diluted below the critical micelle concentration due to their kinetic stability.[18] When the polymer concentration is above CMC, the self-assembly of micelles is in a constant equilibrium where the exchange of polymer chains and fusion and fission of micelles does occur.[19] These are important aspects to be considered when working with block copolymer micelles in very diluted environments.

Small Molecule Uptake**.** The amphiphilic block copolymer micelles facilitate interaction with small molecules. The core of these micelles is hydrophobic, and therefore hydrophobic molecules could be spontaneously incorporated into its interior (Figure 3). This greatly increases the apparent solubility of hydrophobic molecules in aqueous solution and has given rise to a

plethora of interesting applications, such as using block copolymer micelles as nanocontainers, reaction vessels, and drug delivery vesicles.[20,21]

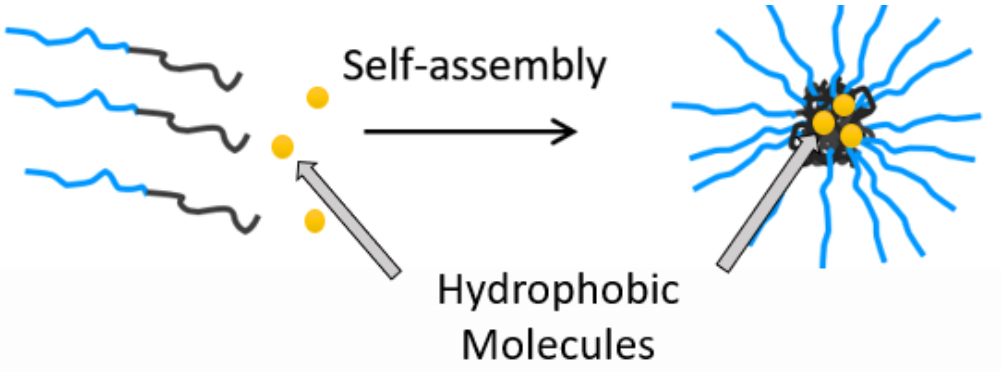

Figure 3. Spontaneous uptake of hydrophobic molecules into the micellar core.

Force Responsiveness. Since block copolymer micelle formation relies on non-covalent interactions, the structure can be reversibly formed or deformed.[19] These micelles can also be disrupted through exposure to mechanical force.[22] One of the most promising areas of research involving stimuli-responsive polymers is drug delivery. Many drugs suffer from poor water solubility or are simply unable to reach their target location in sufficient quantities.[23] This can be circumvented by utilizing amphiphilic micelles which can solubilize the target compound and serve as an intermediary drug carrier in the bloodstream, leading to an increase in solubility and/or half-life.[23] Holme *et al.* have taken advantage of the non-uniform shear force present in the bloodstream to successfully use engineered lipid-based vesicles as drug delivery systems. As shown in Figure 4, lipid-based vesicles were used to selectively deliver drugs to arteries suffering from atherosclerosis, a build-up of fats and carbohydrates in blood vessels.[24] Because of the sudden decrease in cross-sectional area at these locations, shear-stress is increased and forces the release of molecules entrapped in the lipid vesicles.

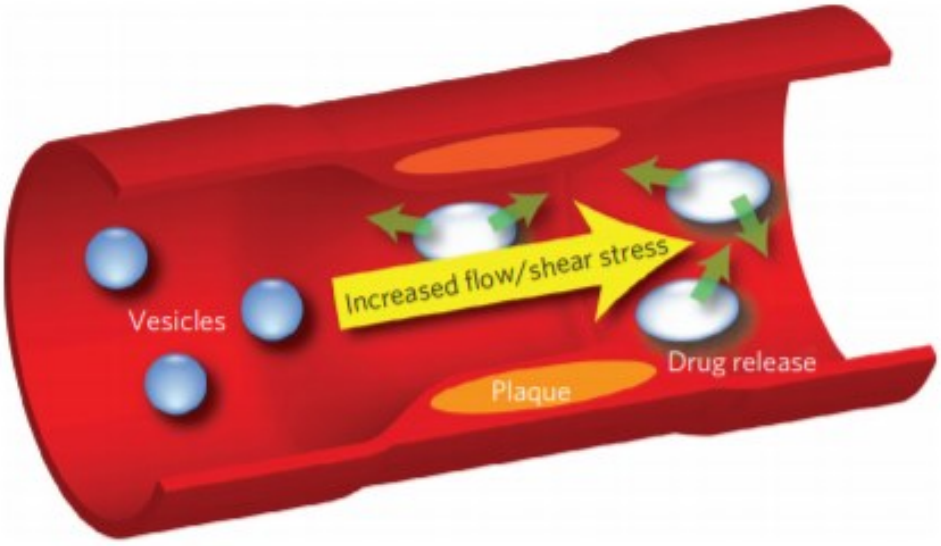

Figure 4. Release of incorporated drugs in an atherosclerotic artery from amphiphilic micelles. Figure adapted from Kucharski *et al*. [24]

Advantages of Block Copolymer Micelles**.** Block copolymer micelles are another class of drug delivery platform and that can be considered a functional analog of the lipid-based systems shown in Figure 4.[24] However, there are several advantages to using block copolymer micelles over low-molecular-weight lipid-based vesicles and small molecule-based micelles. With the advance of polymerization techniques and synthetic routes, polymer chain sizes and compositions can be highly customized.[25] This allows flexibility in polymer design and allows the structural strength to be easily tuned by adjusting chain length. Since hydrophobic interactions are the main driving force for micelle formation, changing the number of monomer units present in the hydrophobic block would increase or decrease the structural strength.[17] Block copolymers also generally have a significantly lower critical micelle concentration than amphiphilic small molecules by 1-2 orders of magnitude, mainly attributed to the significantly larger size and resulting thermodynamic favorability.[19] These inherent advantages have spurred increasing research interest into these systems.

Disruption of Micelle Morphology in Response to Mechanical Stress**.** Even though there is growing interest in these materials, limited research focused on how mechanical stress affects the behavior of block copolymer micelles in application settings. One way to semi-quantitively study the effect of mechanical stress on block copolymer micelles is to examine the changes to structural morphology through imaging techniques such as transmission electron microscopy (TEM). Takeda *et al*. have reported the morphological changes arising from exposure to force. In their work, poly(ethylene glycol)–*b*-poly(*L*-lysine) (PEG-*b*-PLys) block copolymer/DNA polyplex micelles were studied. DNA, which is predominantly negatively charged, was incorporated into the interior to facilitate micelle formation *via* electrostatic interactions with the positively charged lysine residues.[22]

Using a rotational rheometer, the shear forces ranging from 0-100 dyne/cm^2 were applied to the PEG-*b*-PLys/DNA polyplex micelles and the effect of the shear stresses were analyzed by TEM.[22] These values were chosen as it represents the range of shear-stress present in the bloodstreams of humans (1-80 dyne/cm^2) and of mice (100+ dyne/cm^2) which is inversely related to vessel/capillary diameter.[26] Results showed that there is a change in morphology upon exposure to increasing shear-stress. Interestingly, as shown in Figure 5, instead of structural degradation the micelles further aggregate into larger structures. As the applied shear force increases, the block copolymer micelles begin to aggregate into larger cone-like structures.

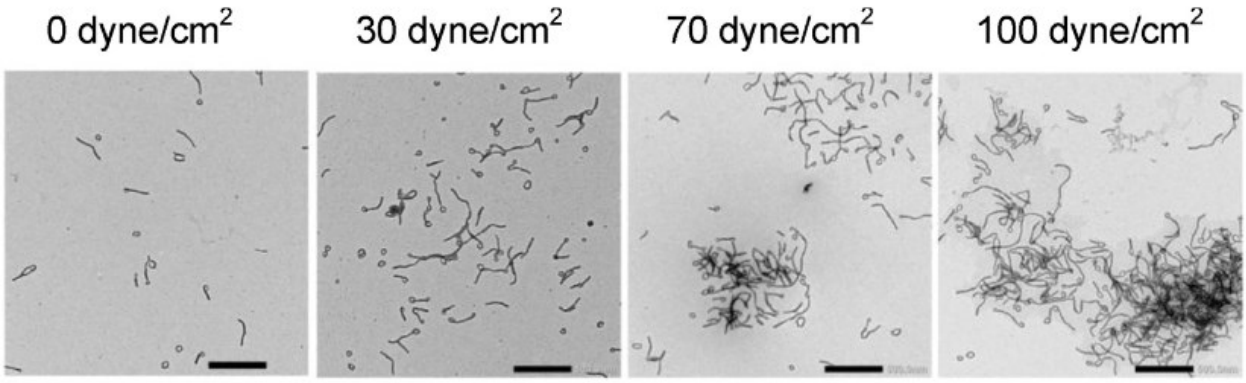

Figure 5. Change in morphology of DNA containing PEG-*b*-PLys block copolymer micelles exposed to shear force. Figure adapted from Cheng *et al*.[26]

Fluorescence and FRET. Fluorescence. Fluorescence is a widely observed phenomenon which involves the photoelectric excitation of an electron to a higher energy level which quickly relaxes back to the ground state, releasing light in the process.[27] It is important to note that the wavelength of light emitted is a longer wavelength than the light absorbed due to a loss of energy. This process is dependent on the molecular structure of the compound. With increased conjugation and planarity lowering the required excitation energy the wavelengths of absorbed/emitted light red-shifts.[27] This is often interpreted spectroscopically by plotting wavelength against relative fluorescence intensity. Since emissions are easily triggered and simple to detect this technique has been utilized in many fields of science. Some of the most common uses are in bioimaging and bioassays.[28,29]

Factors Which Affect Fluorescence. Emissions from fluorescent molecules are affected by a multitude of different factors and are not a perfectly efficient process.[27] When electrons move from an excited state to a ground state and release a photon it is called radiative decay. This is the mechanism responsible for fluorescence emissions, but it is not the only process that can occur during electronic relaxation. The alternative path is non-radiative decay, which is when electrons relax to the ground state without emitting a photon. The ratio of photon input to

photon output in fluorescence systems is termed quantum yield, where perfectly efficient systems provide the value of 1.[30] The loss of quantum yield is primarily caused by the transfer of energy to other molecules or by releasing the energy as heat, though there are other mechanistic routes.[27] More generally, any process that decreases fluorescence output is called quenching. Fluorescence emissions are also affected by their local chemical environment. For example, the wavelengths and relative intensities measured of a given molecule can be drastically different depending on the solvent.[31] This phenomenon is known as the solvatochromic effect.[31] These factors impacting fluorescence emissions have far reaching impacts in research settings. Since small changes in a local environment can alter quantum yield or emission wavelength it is possible to detect minute changes in surroundings spectroscopically.

Fluorescence Resonance Energy Transfer (FRET). Fluorescence resonance energy transfer (FRET) is a non-radiative transfer of energy from a fluorophore in an excited state (donor) to another fluorophore present in close proximity (acceptor).[32] For this phenomenon to occur, there are two main criteria that must be met. First, the emission peak of the donor molecule must significantly overlap with the absorption peak of the acceptor. Second, the two molecules must be in close proximity, approximately within 10 nm of each other.[32] Excitation of the acceptor molecule by FRET from donor molecules results in a decreased emissions by donor molecules and increased emission by acceptor molecules. Alternatively, if the acceptor preferentially releases energy *via* non-radiative processes, a decrease of fluorescence emissions by donor molecules (quenching) would be observed.[33] This phenomenon has been used in quantitative polymerase chain reaction (qPCR), which detects the presence of specific DNA sequences in bulk solution and is commonly used in clinical settings for detection of foreign viral genomes.[34] This technique uses a primer in which a fluorophore is bound in close proximity to a

quencher molecule and upon DNA replication the molecules are separated, leading to an increased fluorescence intensity as the fluorophore is no longer quenched.[34]

Study of Block Copolymer Micelles Using Fluorescence. Fluorescence and FRET have not only proved useful in biological environments, but these have been widely adapted in the analysis of nanomaterials. These tools are especially useful in studying block copolymer micelles. When an appropriate pair of fluorophores is incorporated into the hydrophobic core of micelles, fluorophores could undergo FRET due to the colocalization of both fluorophores in close proximity. Lu *et al.* has reported the use of a donor-acceptor dye pair to study structural stability of copolymer micelles in the presence of several different globular proteins common in the bloodstream. These proteins can disrupt micelle structure which greatly decreases their efficacy as drug carriers.[35] By incorporating a 3,3'-dioctadecyloxacarbocyanine perchlorate (DiO)/ 1,1'-dioctadecyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate (DiI) FRET pair into the core of the micelles, they were able to determine the stability of micelles using fluorescence spectroscopy. While the micelles were relatively stable, in the presence of these globular proteins there was a steady drop in the FRET ratio, indicating the dyes are transferred to proteins and/or the micelles were structurally disrupted.[35] As shown in Figure 6, the drop in acceptor emissions (570 nm) and the subsequent rise of donor emissions (505 nm) correspond to a decrease in FRET activity over time.

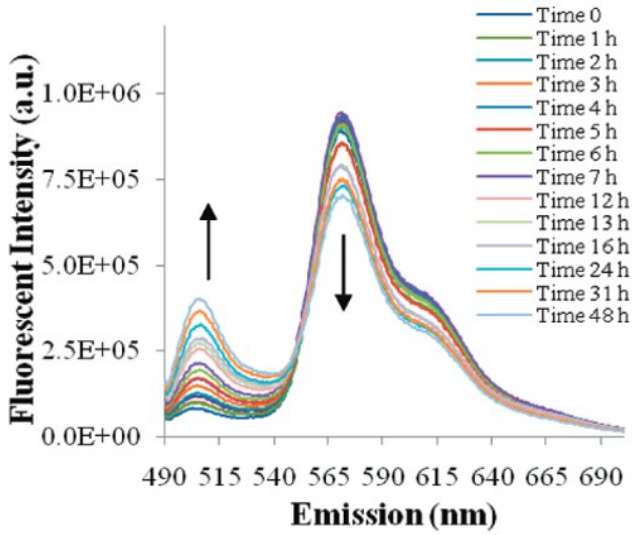

Figure 6. Fluorescence spectrum of block copolymers loaded with a donor-acceptor FRET pair in the presence of globular proteins. Figure adapted from Lu *et al*.[35]

Robin *et al.* have reported the use of fluorescent molecules in block copolymer micelles as a probe to determine whether they are assembled and/or contain a small molecule in the core. They covalently bound the acceptor fluorophore dithiomalamide (DTM) to either the outer corona or inner shell of poly(triethylene glycol acrylate)-*b*-poly(tert-butyl acrylate) (TEGA-*b*-tBA) micelles. They measured the state of the system by measuring fluorescent lifetimes, which is how long it takes a population of fluorophores to relax to the ground state, usually quantified in terms of half-life.[27] The fluorescence lifetime of DTM is sensitive to its surrounding environment and will decrease on average as exposure to intermolecular collisions from solvent increases.[36] It was found that the fluorescence lifetime of core-bound DTM was significantly longer than its corona-bound partner, indicating that the core offers good protection from solvent (Figure 7).[36] They also incorporated a hydrophobic donor fluorophore into the system that will initiate FRET if it is in close enough proximity to DTM molecules bound in the inner core, ultimately providing a secondary confirmation that the micelles are formed.

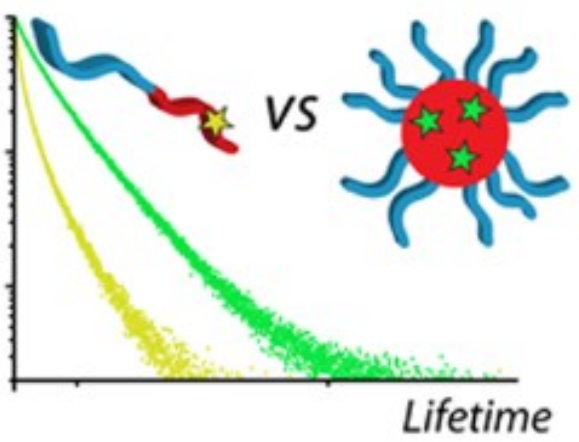

Figure 7. Fluorescence lifetimes of covalently labeled DTM fluorophore when copolymer chains are aggregated into a micelle (green) and when free in aqueous solution (yellow). Figure adapted from Robin *et al*.[36]

Chen *et al.* used dye-containing block copolymer micelles to study how block copolymers and their loaded cargo interact with cellular membranes. This mechanism is analogous to similar systems used in drug delivery. Using a donor-acceptor dye pair incorporated into poly(ethylene glycol)-*b*-poly(lactic acid) (PEG-*b*-PLA) micelles they were able to visualize the rate of dye transfer and deduce several important characteristics of the mechanism. It was found that the dye release was induced by fusion of the PEG corona to the exterior of the lipid membrane, which improves permeability and facilitates dye transfer.[37] As the dyes leave the micelle and enter the lipid membrane the FRET ratio decreases, providing a unique tool to observe partitioning of lipophilic molecules under a microscope.

As described above, FRET has been utilized to study the various behaviors of block copolymer micelles. Since these fluorescent dyes undergo FRET only when the fluorophores are in close proximity, the breakdown of micellar structure and/or repartitioning/transfer of fluorophores can be analyzed by a characteristic change in fluorescence emission spectra (Figure 8). However, there has been little research on applying FRET in studying how flow-induced mechanical stresses affect the stability of block copolymer micelles.

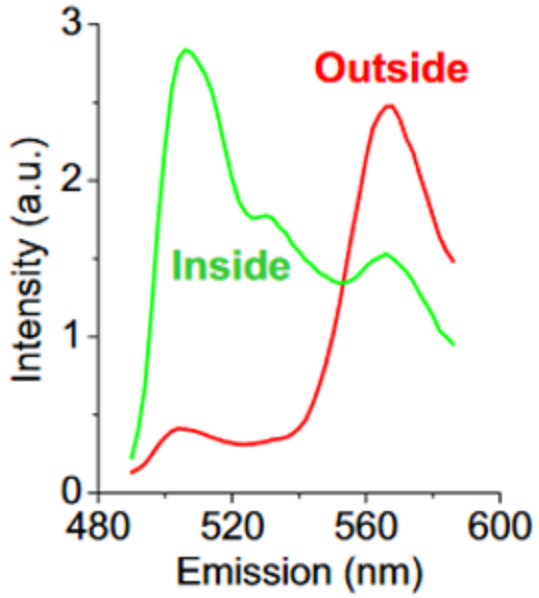

Figure 8. FRET emissions of dye-containing block copolymer micelles in the presence of a membrane when fluorophores are inside the membrane (green) and inside the micelle (red). Figure adapted from Chen *et al* (2008).[37]

Objective

While many mechanical force-responsive materials have been investigated, there has been little research on the development of mechanical-force responsive block copolymer micelles. In this study I focused on the spectroscopic study of the fluorescent dye containing block copolymer micelles to investigate their responses to vortexing-induced mechanical stresses. We chose poly(ethylene glycol)-*b*-poly(lactic acid) (PEG-*b*-PLA) micelles as the model system for their low toxicity and biodegradability, which made them attractive in biomedical applications.[37–39]

My hypothesis was that dye-containing PEG-*b*-PLA micelles can be utilized to detect the mechanical forces in solution. Since the micelles are stabilized solely through non-covalent interactions, structural breakdown or disruption should induce the release of fluorophores into the aqueous phase, resulting in the decrease of FRET efficiency. The specific goal of this study

was to investigate the behavior of dye-containing block copolymers to mechanical stresses using fluorescence spectroscopy. The originality of this work was on the application of dye-containing block copolymer micelles as a force-responsive nanoprobe system to detect mechanical stresses in solution.

Materials

- Polyethylene glycol(5 kDa)-*b*-polylactic acid(5 kDa)
- 3,3′-Dioctadecyloxacarbocyanine perchlorate (DiO) (34215-57-1)
- 1,1′-Di-*n*-octadecyl-3,3,3′,3′-tetramethylindocarbocyanine perchlorate (DiI) (41085-99-8)
- Standard grade cellulose dialysis tubing (molecular weight cut off: 3.5 kD)
- Acetone (67-64-1)
- Methanol (67-56-1)
- Carbon film-coated Cu TEM grids

Methods

Preparation of Block Copolymer Micelles. Stock solutions of the dyes were made by dissolving 1.0 mg of either 3,3'-dioctadecyloxacarbocyanine perchlorate (DiO) or 1,1'-dioctadecyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate (DiI) in 1.0 mL acetone (1.0 mg mL^{-1}) in a 1.5 mL centrifuge tube then stored in a -20 ˚C freezer. The dyes are not fully soluble in acetone and tend to precipitate over time, so absorption measurements were taken before each use to measure effective concentration. It should be mentioned that DiO precipitated out of solution at a significantly faster rate than DiI and the concentration of several stock solutions decreased by more than 50% over a year.

The poly(ethylene glycol) -*b*-poly(lactic acid) (PEG-*b*-PLA) block copolymer micelles were prepared using a procedure adapted from Chen *et al*.[37] The molecular weight of both polymer blocks is ~5 kDa for a total polymer molecular weight of ~10 kDa. A typical protocol is

as follows: 10.0 mg of polymer was dissolved in 1.0 mL of acetone and placed in a 1.5 mL centrifuge tube. The polymer is not immediately soluble, so needs to be shaken until fully dissolved. 0.015 mg and 0.075 mg of DiO and DiI, respectively, were added to the solution. The mixture was then added dropwise over ~3 minutes into a 20 mL borosilicate scintillation vial containing 5.0 mL of deionized water while stirring at 300 rpm. The solution was then stirred for 3 hours to vaporize acetone and to facilitate micelle formation and dye uptake. Once all acetone was removed the solution was transferred to a 5.0 mL centrifuge tube and dialyzed against a 3.5 kDa dialysis membrane for 48 hours to remove any dyes not incorporated into the micelle core. The dialysis water was changed after 24 hours. After dialysis was complete the solution was transferred to a fresh 5 mL polypropylene centrifuge tube and stored at 4 ˚C. All samples were wrapped in aluminum foil to avoid light exposure and any potential photobleaching of the fluorescent dyes.

UV-Vis Measurements. Determining Dye Concentration. As previously mentioned, DiO and DiI tended to precipitate out of acetone over time so it is critical to measure the concentration each time that stock dyes are used. Samples were diluted 1/1000 in acetone and placed into a 1.5 mL quartz cuvette. Absorption was then measured in 1 nm intervals from 400-600 nm using a Cary 60 Spectrophotometer (Agilent Technologies). Molar extinction coefficients used are ε_{484} = 154,000 $M^{-1}cm^{-1}$ and ε_{549} = 148,000 $M^{-1}cm^{-1}$ for DiO and DiI, respectively.[37]

Micelle Absorbance Measurements. DiO/DiI loaded PEG-*b*-PLA micelles were diluted 1/2 in water for a final concentration of 1.0 mg mL^{-1} then scanned from 400-600 nm at 1 nm intervals in the spectrophotometer. In acetone the micelles were diluted 1/3 for a concentration of 0.67 mg mL^{-1} then measured over the same range. 1.5 mL quartz cuvettes with a path length of 1

cm were used for all absorbance measurements. For each run an acetone or water blank was measured and subtracted from the results.

Fluorescence Measurements. All fluorescence measurements were taken on a QuantaMaster spectrofluorometer (Photon Technology International). Generally, the dye-containing block copolymer micelles were diluted 400-fold in water (2 mg mL^{-1} → 0.005 mg mL^{-1}) as the detector becomes saturated at higher micelle concentrations due to the corresponding increase in fluorescent dyes present. Samples were transferred to a 3 mL quartz cuvette then exposed to an excitation wavelength of 425 nm and scanned over 450-650 nm at 1 nm intervals. The instrument was set to have an excitation slit width of 8 nm and an emission slit width of 10 nm.

Measurement of Critical Micelle Concentration. Critical micelle concentration was determined by adapting procedure reported by Yasugi *et al*.[40] An initial stock solution was made by dissolving 5 mg of PEG-*b*-PLA in 2 mL of acetone for a final concentration of 2.5 mg mL^{-1}. This solution was further aliquoted into several 1.5 mL polypropylene centrifuge tubes to obtain concentrations spanning a broad linear range. The final concentrations were: 2.5, 0.625, 0.156, 0.078, 0.039, 0.020 and 0.010 mg mL^{-1}. Pyrene (9.0 mg) was added to a 15 mL polypropylene centrifuge tube and dissolved in 15 mL of acetone (600 μg mL^{-1}) to serve as a stock solution. This stock was then diluted 2000-fold by multiple serial dilutions (0.3 μg mL^{-1}) to yield the final solution used in experimental work. To each prepared PEG-*b*-PLA sample, 11.3 μL of the 0.3 μg mL^{-1} pyrene solution was added. Each mixture was then added dropwise over ~2 minutes into 20 mL borosilicate scintillation vials containing 5 mL of deionized water under vigorous stirring. Solutions were stirred for 5 hours to evaporate any acetone present. The final pyrene

concentration in each solution is 0.67 ng mL^{-1} and after the 5-fold dilution polymer concentrations were 0.5, 0.125, 0.031, 0.0156, 0.0078, 0.0039 and 0.0020 mg mL^{-1}.

Fluorescence excitation spectra of each sample were then measured. Instrumental settings were as follows: 280-380 nm excitation scan, 390 nm fixed emission wavelength, 8 nm excitation slit width, 10 nm emission slit width. Fluorescence intensity at 337 nm was plotted as a function of polymer concentration. Two linear regression lines were created to fit the high and low concentration ranges with the intersection point corresponding to the critical micelle concentration.

Exposure of Micelles to Shear Stress *via* Vortexing. Typical procedure for micelle vortexing is as follows: 2.0 mg mL^{-1} stock micelle samples prepared previously were diluted 400-fold in water (5.0 μg mL^{-1}) to a final volume of 3 mL and placed into 15 mL polypropylene centrifuge tubes. The samples were then placed on a Digital Vortex-Genie 2 vortexer equipped with a six arm, vertical tube holder. Three tubes were always placed on alternating arms with 3 mL of water or micelle solution to equally disperse weight across the instrument. The samples were then vortexed at 3000 rpm for 5 minutes and fluorescence spectra were subsequently measured. While this is the most typical procedure used, depending on the purpose of the experiment the duration of vortexing and rotational speeds varied widely.

Transmission Electron Microscopy Measurements. Transmission electron microscope (TEM) images were acquired using a Tecnai F20 field-emission electron microscope with an acceleration voltage of 200 kV. Several 2.0 mg mL^{-1} polymer solutions were prepared using the standard, previously stated procedure. 3 μL of block copolymer solution was transferred onto carbon film-coated copper TEM grids then air dried and images were subsequently taken. The

size of block copolymer micelles was quantified using ImageJ software. The results were then further analyzed and visualized using Microsoft Excel.

Results & Discussion

Preparation and Optimization of DiO/DiI Containing PEG-*b*-PLA Micelles.

Preparation of Micelles Consisting of Selective Components. To ensure that the fluorescent dyes were encapsulated in the interior core of the PEG-*b*-PLA block copolymer micelles and that FRET is occurring, micelles consisting of all, or part of the components were systematically prepared. Solutions were prepared by an identical procedure except for the exclusion of one or more of components. The results of this experiment are shown in Figure 9. When only polymer and DiO (donor) are present, a large fluorescence peak at 505 nm was observed. Likewise, when polymer and DiI (acceptor) are present a smaller peak at approximately 567 nm was observed. The relatively small fluorescence emission peak observed with DiI containing micelles is primarily due to the use of 425 nm as the excitation wavelength, which is the edge of the absorption region for DiI. On the other hand, when both DiO and DiI are present, a large peak at 567 nm and a drastically decreased fluorescence signal at 505 nm was observed. This indicates the occurrence of FRET, indicating that both dyes are successfully incorporated into the interior core of the micelles.[32] Interestingly, when the polymer is excluded, the dyes do not exhibit any appreciable fluorescence signal, possibly due to a lack of solubility in aqueous solution.

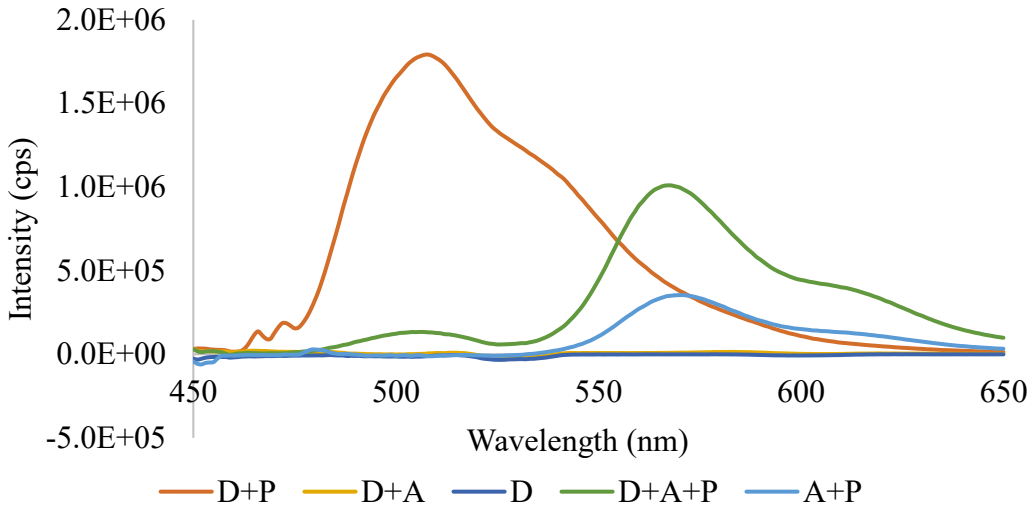

Figure 9. Effect of removing select components from our dye-containing block copolymer micelles on the fluorescence spectra. Donor + polymer (red), donor + acceptor (yellow), donor only (dark blue), donor + acceptor + polymer (green), acceptor + polymer (light blue).

Optimization of DiO/DiI Ratio. A similar system was reported by Chen *et al.* that uses 0.75% w/w DiO and DiI relative to polymer mass.[37] Since we are ultimately interested in monitoring effects of mechanical stress on fluorescence emissions of our dye containing block copolymer micelles it is optimal to have as high of a FRET efficiency as a possible ($I_{acceptor} \gg I_{donor}$). This case seeks to minimize donor emissions and increase acceptor emissions, ensuring that fluorescence spectra of free dyes and encapsulated dyes are different and easily distinguishable from each other. To test this, we prepared block copolymer micelle solutions using standard procedure but decreased the amount of DiO (donor) relative to DiI (acceptor). The resulting fluorescence spectra are shown in Figure 10, with the percentages below the graph corresponding to the concentration of DiO relative to DiI. FRET efficiency generally increased as we decreased concentration of the donor dye. This indicates that at as the ratio of DiO to DiI approaches 1:1 the acceptor becomes saturated and FRET emissions begin to approach a maximum. Table 1 shows the ratio of acceptor peak intensity to donor peak intensity at varying concentrations of DiO. DiO concentrations of 50% (relative to DiI concentration) and lower had

roughly comparable donor/acceptor peak ratios and similar maximum emissions at 567 nm (acceptor/FRET peak). We chose to use 20% DiO concentration for the remaining work as it ensures the acceptor is not fully saturated. The data suggested 33% or 50% would likely have good ratios as well.

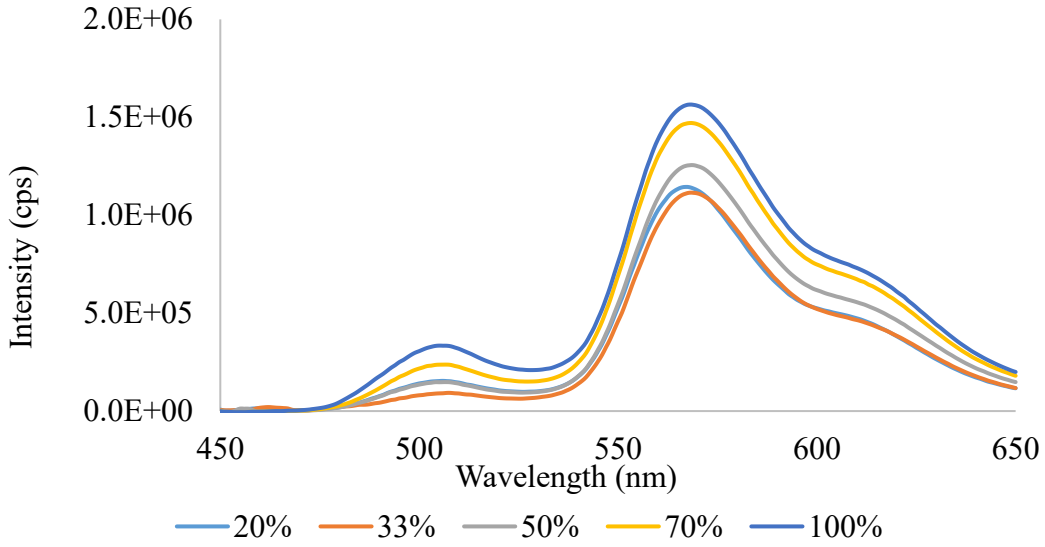

Figure 10. Fluorescence spectrum of varying DiO concentrations as a relative percentage to DiI concentration.

Table 1. Ratio of acceptor to donor peaks at varying concentrations of DiO expressed as a relative percentage of DiI.

DiO concentration relative to DiI (in relative % to DiI, w/w)	20	33	50	70	100
Ratio of acceptor peak to donor peak	7.42	12.3	8.49	6.20	4.68

Optimization of Dye/Polymer Ratio. To investigate the relationship between polymer concentration and fluorescence emissions, we prepared samples with a fixed DiO/DiI ratio (1:5, w/w) and altered the ratio the total dyes to PEG-*b*-PLA block copolymer. In order to keep the theoretical peak intensity constant, the total amount of dye was kept the same in all measurements, while the total amount of the polymer in each sample were varied. As seen in

Figure 11, changing the polymer concentration has a significant impact on the fluorescence emission spectra. As the dye/polymer ratio increases, an increase in donor peak emissions was observed. A decrease in acceptor peak emissions was also observed in most cases. We attributed this trend to the result of dyes being further spaced out as the increased polymer concentration leads to the formation of more micelles and/or larger micelles. Importantly, the results from earlier control experiments showed that in the absence of polymer, DiO and DiI do not fluoresce. This means that the observed change in fluorescence signal must be caused by the physical isolation of dyes. For the rest of the work, we employed the dye/polymer ratio of 0.75 mg DiI and 0.15 mg DiO per mg of PEG-*b*-PLA block copolymer.

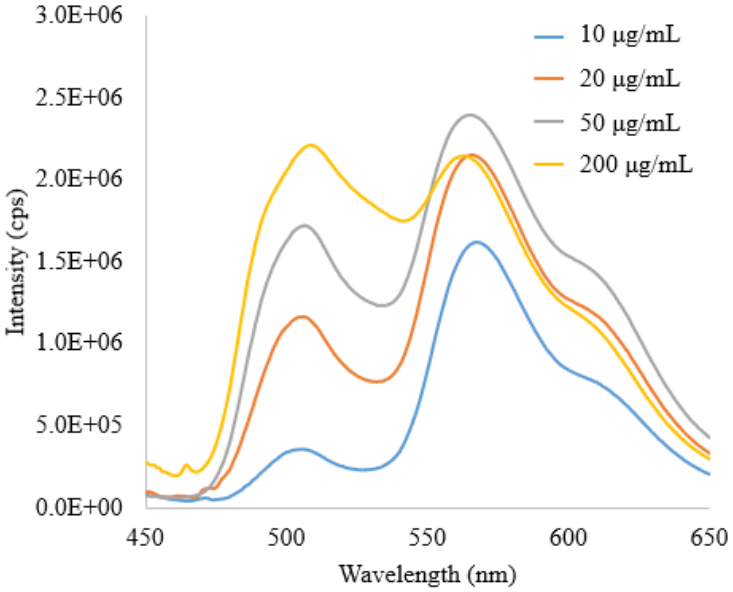

Figure 11. Fluorescence spectra of altered dye/polymer ratio.

Critical Micelle Concentration. Before starting experiments testing force-responsiveness of the micelles, it was important to establish the critical micelle concentration (CMC). Above the CMC, the micelle formation occurs spontaneously and is thermodynamically favored.[17] In this study it is preferred that any morphological or chemical changes upon exposure

to mechanical forces cause irreversible changes. In order to ensure these criteria, samples must be prepared above CMC and then diluted below the CMC where they are only kinetically stable and will not readily reform micelles.[19] To determine CMC, we used an established technique which utilizes pyrene as a fluorescent probe. Pyrene undergoes a significant solvatochromic shift depending on the polarity of the surrounding environment. This phenomenon can be used to measure the CMC of block copolymers since pyrene will preferentially be incorporated into the hydrophobic block copolymer micelle core.[40] Fluorescence emissions at 390 nm were plotted as a function of concentration and two linear regression trendlines were assigned to high and low concentration ranges (Figure 12). The inflection point between the two trendlines indicates that the CMC of our polymer was 19 μg mL^{-1}, which agrees with previously reported results of PEG-*b*-PLA composed of similar chain lengths.[41] To ensure micelle formation, we prepared our block copolymer micelles at a concentration of 2 mg mL^{-1}, then diluted the solution to 5 μg mL^{-1} for tests and fluorescence measurements for the remainder of the work.

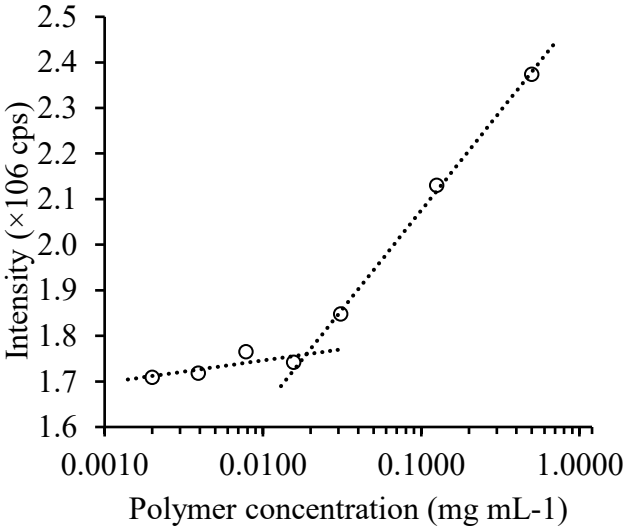

Figure 12. Fluorescence intensity at 390 nm plotted against concentration of PEG-*b*-PLA, used to determine critical micelle concentration.

Effect of Vortexing on Fluorescence Properties. Micelle Stability and the Effect of Vortexing. To investigate the stability of the dye-containing block copolymer micelles, the fluorescence emission spectrum of micelle samples that were diluted below the CMC were measured. A stock solution containing 5 μg mL^{-1} of DiO/DiI containing block copolymer was aliquoted into 15 mL polypropylene centrifuge tubes and left under static conditions. Fluorescence spectra were then measured at intervals of 1 hour over a 5-hour period, then samples were vortexed at 3000 rpm for 5 mins. After the vortexing, the samples were further incubated statically, and the spectrum was measured at 1-hour intervals for 3 hours. As shown in Figure 13, little to no change was observed in fluorescence emission spectra when samples were statically incubated, indicating that the micelles are kinetically stable even below CMC and do not decompose over the course of, at least, several hours. On the other hand, as soon as the solutions are vortexed, a roughly 50% decrease in fluorescence emission intensity was observed, indicating that exposure to mechanical stress significantly affects the optical properties of the micelles. Interestingly, the emissions remain at this lower level and did not significantly change during the subsequent 3 hours of incubation under static condition. This suggests that the process is irreversible, and the dyes do not spontaneously reincorporate into the micelle core.

Vortex Time Dependence. To investigate the effect of vortexing on the optical properties of the block copolymer micelles, we subjected the micelle samples to vortexing for varying lengths of time. It was found that the correlation between the vortexing time and fluorescence intensity is not linear. Within the first 30 seconds of vortexing, fluorescence emissions at 567 nm decreased by ~25% then the rate of change begins to slow down, with a total drop of ~50% after 5 minutes (Figure 14). It is noteworthy that not only does the acceptor peak decrease, but the donor peak decreases after vortexing as well. This result is somewhat unexpected because we

anticipated a reduction in FRET efficiency as dyes become physically separated. One plausible explanation for this is that the dyes are being expelled from the micelle core to form aggregates in aqueous phase and underwent self-quenching. Since the dyes do not necessarily have to be expelled at the same rate, this phenomenon could alternatively be explained by the preferential expulsion and of DiO from the core as it is primary excitation source for DiI. Regardless of the precise mechanism, these explanations agree with the results of a control experiment, where dyes added into aqueous solution in absence of polymer did not exhibit fluoresce (Figure 12).

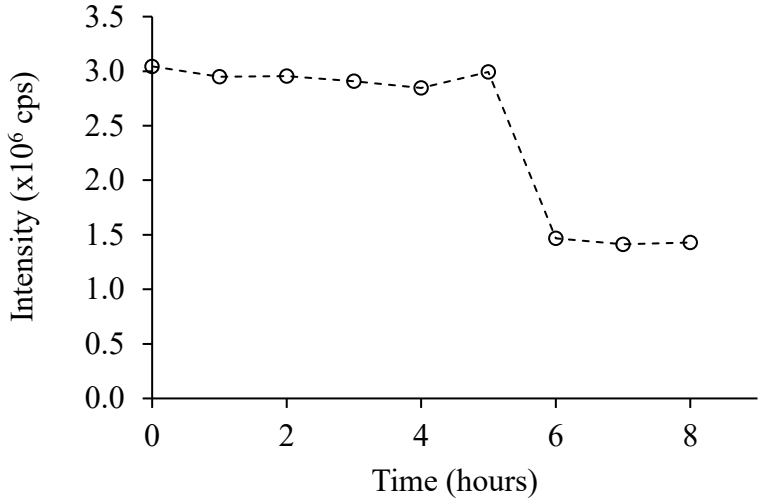

Figure 13. Fluorescence intensity at 567 nm of our dye-containing block copolymer micelles left under static conditions (hours 0-5), vortexed (hour 5), then left again under static conditions (hours 6-8).

We next performed a similar experiment where our dye-containing block copolymer solutions were vortexed for up to 10 minutes at two rotational speeds, 1500 or 3000 rpm. Figure 15 shows the fluorescence emissions at 567 nm plotted against time of vortexing for both rotational speeds. It was observed that samples vortexed at 3000 rpm had a significant decrease in fluorescence emissions within 30 seconds and continued to drop through 10 minutes of vortexing. A similar effect is seen to previous results, where the rate of change begins to slow

and seemingly approach a lower plateau as vortex time is extended. The results for 1500 rpm are similar, except the drop in fluorescence emissions proceeds at a much slower rate than 3000 rpm and does not decrease as much over the same period. A plausible explanation is a fraction of the micelles are more robust than others, and therefore are more resistant to the effect of mechanical stress. At the excitation wavelength used (425 nm) DiI exhibits a small, but not insignificant amount of fluorescence activity. This secondary contribution is likely contributing to the 'plateau' effect as well since these background emissions occur independently of FRET. Overall, this provides evidence that exposing the dye-containing micelles to mechanical stress *via* vortexing induces a decrease in fluorescence emissions.

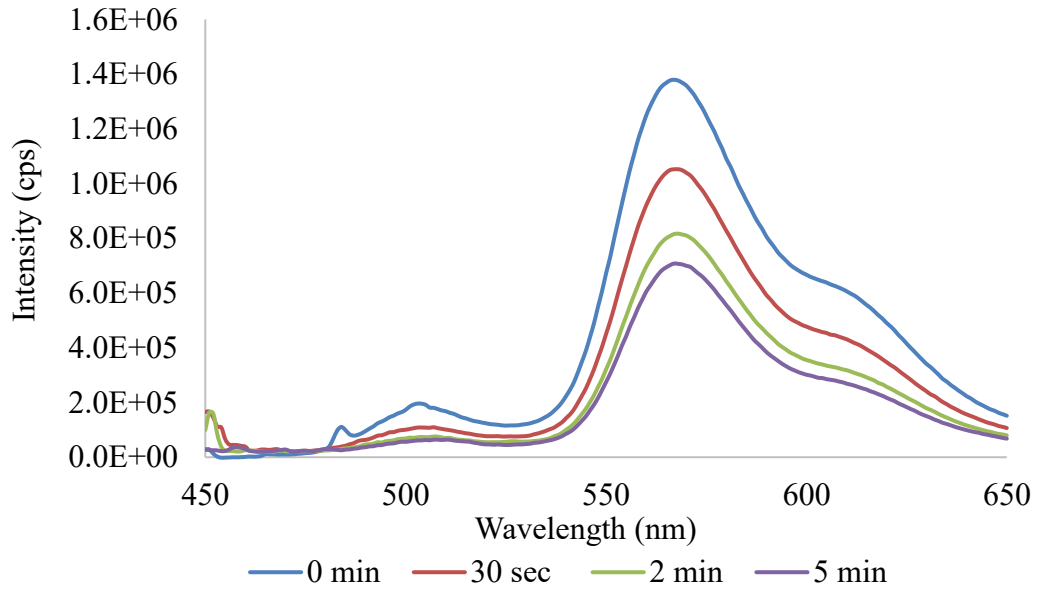

Figure 14. Fluorescence spectrum of DiO/DiI containing PEG-*b*-PLA block copolymer micelles vortexed at 3000 rpm from 0-5 minutes.

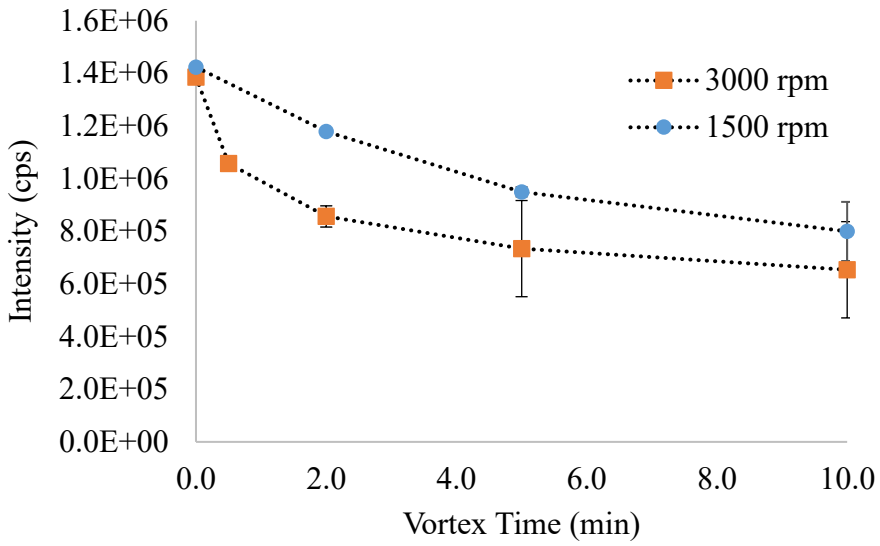

Figure 15. Fluorescence emissions at 567 nm of DiO/DiI containing PEG-*b*-PLA block copolymer micelles vortexed for varying times, up to 10 minutes total.

Rotational Speed Dependence. To study the relationship between rotational speed and fluorescence activity in a more detailed manner, dye-containing block copolymer micelles were exposed to a broad range of rotational speeds. In this experiment samples were vortexed for 5 minutes at speeds ranging from 0-3000 rpm at intervals of 250 rpm. The fluorescence intensity at 567 nm (acceptor peak) is shown in Figure 16 as a function of rotational speed. The correlation between rate of rotation and fluorescence emissions is non-linear and tends to follow a sigmoidal trend. At lower rotational speeds (<1000 rpm) there is very little change in fluorescence emissions. However, between 1250-1750 rpm is a significant decrease in emissions. When the samples were vortexed at or above 1750 rpm, fluorescence emissions were similar to each other regardless of the rotational speeds. These results suggest that the fluorescence emissions are affected only when the micelles are exposed to mechanical stress above a minimum threshold whereas the decrease in fluorescence intensity begins to plateau when the vortexing speed attains a certain level.

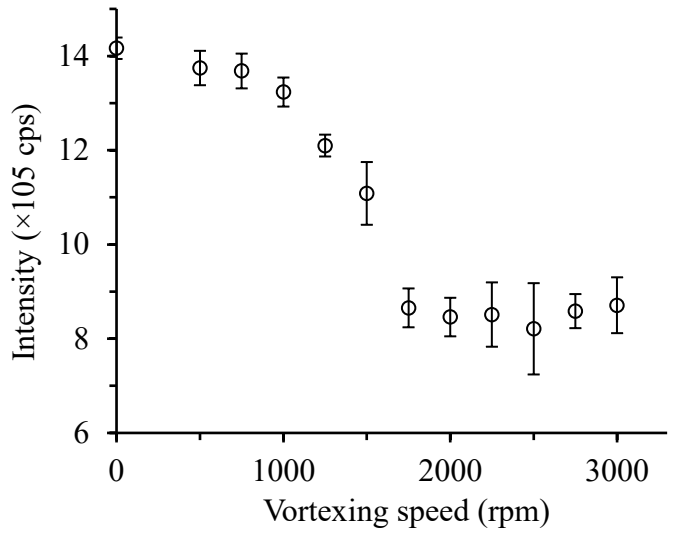

Figure 16. Fluorescence emissions at 567 nm plotted against applied rotational speed.

Transmission Electron Microscopy Imaging. Fluorescent spectroscopic data showed that vortexing-induced mechanical stress affects the fluorescent properties of dye-containing block copolymer micelles. To gain further insight into the effect of vortexing transmission electron microscopy (TEM) analysis was performed (full images shown in Appendix A). Our initial hypothesis was that the application of mechanical stress causes structural deterioration and micelle dissociation, especially since the system is below the critical micelle concentration. We assumed this is the mechanism releasing dyes into the aqueous phase and subsequently decreasing fluorescence emissions. However, as shown in Figure 17, results from TEM imaging contradicted our theories as vortexing actually caused an increase in average micelle diameter. Micelles in the unvortexed sample had an average diameter of 17.5 nm with a standard deviation of 4.0 nm (n = 91). After vortexing the average micelle diameter was 22.7 nm with a standard deviation of 10.7 nm (n = 180). Not only is the average size different between the two samples, but size distribution follows a different pattern as well. The sample left under static conditions follows a typical monomodal distribution while the vortexed solution has a bimodal distribution

with skew towards larger diameters. Because of the resolution limits of TEM, we were unable to definitively measure micelles below 9 nm as they are indistinguishable from instrumental noise. Because of this, we cannot eliminate the possibility that smaller structures are forming, or that complete degradation is occurring.

Overall, this work shows that the vortexing-induced mechanical stress applied to the PEG-*b*-PLA block copolymer micelles was sufficient to trigger morphological changes, with the micelles becoming larger on average after exposure to vortexing. This phenomenon is likely being caused by aggregation, fusion, chain exchange, or through some combination of these mechanisms. This appears to be consistent with previously observed results since this level of mechanical stress can cause the expulsion of dyes from the interior micelle core which subsequently aggregate and/or self-quench, leading to a loss in fluorescence signal.

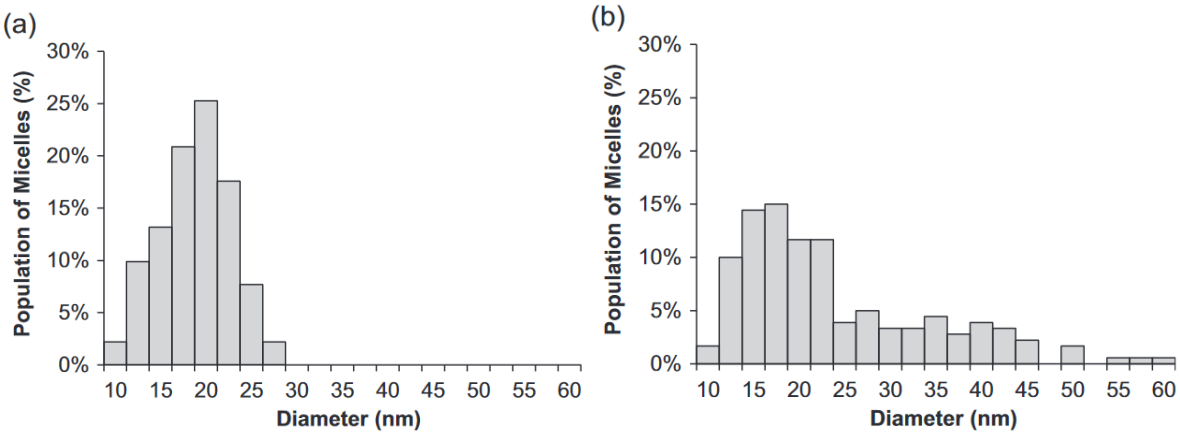

Figure 17. Distribution of micelle diameter based on TEM data. PEG-*b*-PLA micelles before vortexing (top) and after vortexing (bottom).

References

(1) Wei, M.; Gao, Y.; Li, X.; Serpe, M. J. Stimuli-Responsive Polymers and Their Applications. *Polym. Chem.* **2017**, *8*, 127–143.

(2) Dai, S.; Ravi, P.; Tam, K. C. PH-Responsive Polymers: Synthesis, Properties and Applications. *Soft Matter* **2008**, *4*, 435.

(3) Chiantore, O.; Guaita, M.; Trossarelli, L. Solution Properties of Poly(N-Isopropylacrylamide). *Makromol. Chem.* **1979**, *180*, 969–973.

(4) Kucharski, T. J.; Boulatov, R. The Physical Chemistry of Mechanoresponsive Polymers. *J. Mater. Chem.* **2011**, *21*, 8237.

(5) Tanaka, T.; Nishio, I.; Sun, S.-T.; Ueno-Nishio, S. Collapse of Gels in an Electric Field. *Science* **1982**, *218*, 467–469.

(6) Roy, D.; Cambre, J. N.; Sumerlin, B. S. Future Perspectives and Recent Advances in Stimuli-Responsive Materials. *Prog. Polym. Sci.* **2010**, *35*, 278–301.

(7) Wang, S.; Urban, M. W. Self-Healing Polymers. *Nat. Rev. Mater.* **2020**, *5*, 562–583.

(8) Wang, Y.; Han, P.; Xu, H.; Wang, Z.; Zhang, X.; Kabanov, A. V. Photocontrolled Self-Assembly and Disassembly of Block Ionomer Complex Vesicles: A Facile Approach toward Supramolecular Polymer Nanocontainers. *Langmuir* **2010**, *26*, 709–715.

(9) Burnworth, M.; Tang, L.; Kumpfer, J. R.; Duncan, A. J.; Beyer, F. L.; Fiore, G. L.; Rowan, S. J.; Weder, C. Optically Healable Supramolecular Polymers. *Nature* **2011**, *472*, 334–337.

(10) Colson, Y. L.; Grinstaff, M. W. Biologically Responsive Polymeric Nanoparticles for Drug Delivery. *Adv. Mater.* **2012**, *24*, 3878–3886.

(11) Wiggins, K. M.; Brantley, J. N.; Bielawski, C. W. Methods for Activating and Characterizing Mechanically Responsive Polymers. *Chem. Soc. Rev.* **2013**, *42*, 7130.

(12) Davis, D. A.; Hamilton, A.; Yang, J.; Cremar, L. D.; Van Gough, D.; Potisek, S. L.; Ong, M. T.; Braun, P. V.; Martínez, T. J.; White, S. R.; Moore, J. S.; Sottos, N. R. Force-Induced Activation of Covalent Bonds in Mechanoresponsive Polymeric Materials. *Nature* **2009**, *459*, 68–72.

(13) Carrington, S. P.; Odell, J. A. How Do Polymers Stretch in Stagnation Point Extensional Flow-Fields? *J. Non-Newton. Fluid Mech.* **1996**, *67*, 269–283.

(14) Caruso, M. M.; Davis, D. A.; Shen, Q.; Odom, S. A.; Sottos, N. R.; White, S. R.; Moore, J. S. Mechanically-Induced Chemical Changes in Polymeric Materials. *Chem. Rev.* **2009**, *109*, 5755–5798.

(15) Gohy, J.-F. Block Copolymer Micelles. In *Block Copolymers II*; Abetz, V., Ed.; Advances in Polymer Science; Springer-Verlag: Berlin/Heidelberg, **2005**; *190*, 65–136.

(16) Wang, C.-W.; Sinton, D.; Moffitt, M. G. Flow-Directed Block Copolymer Micelle Morphologies via Microfluidic Self-Assembly. *J. Am. Chem. Soc.* **2011**, *133*, 18853–18864.

(17) Johnson, B. K.; Prud'homme, R. K. Mechanism for Rapid Self-Assembly of Block Copolymer Nanoparticles. *Phys. Rev. Lett.* **2003**, *91*, 118302.

(19) Lu, Y.; Zhang, E.; Yang, J.; Cao, Z. Strategies to Improve Micelle Stability for Drug Delivery. *Nano Res.* **2018**, *11*, 4985–4998.

(18) Gibson, M. I.; O'Reilly, R. K. To Aggregate, or Not to Aggregate? Considerations in the Design and Application of Polymeric Thermally-Responsive Nanoparticles. *Chem Soc Rev* **2013**, *42*, 7204–7213.

(20) Kane, R. S.; Cohen, R. E.; Silbey, R. Synthesis of PbS Nanoclusters within Block Copolymer Nanoreactors. *Chem. Mater.* **1996**, *8*, 1919–1924.

(21) Kataoka, K.; Harada, A.; Nagasaki, Y. Block Copolymer Micelles for Drug Delivery: Design, Characterization and Biological Significance. *Adv. Drug Deliv. Rev.* **2012**, *64*, 37–48.

(22) Takeda, K. M.; Yamasaki, Y.; Dirisala, A.; Ikeda, S.; Tockary, T. A.; Toh, K.; Osada, K.; Kataoka, K. Effect of Shear Stress on Structure and Function of Polyplex Micelles from Poly(Ethylene Glycol)-Poly(l -Lysine) Block Copolymers as Systemic Gene Delivery Carrier. *Biomaterials* **2017**, *126*, 31–38.

(23) Langer, R. New Methods of Drug Delivery. *Science* **1990**, *249*, 1527–1533.

(24) Holme, M. N.; Fedotenko, I. A.; Abegg, D.; Althaus, J.; Babel, L.; Favarger, F.; Reiter, R.; Tanasescu, R.; Zaffalon, P.-L.; Ziegler, A.; Müller, B.; Saxer, T.; Zumbuehl, A. Shear-Stress Sensitive Lenticular Vesicles for Targeted Drug Delivery. *Nat. Nanotechnol.* **2012**, *7*, 536–543.

(25) O'Reilly, R. K.; Hawker, C. J.; Wooley, K. L. Cross-Linked Block Copolymer Micelles: Functional Nanostructures of Great Potential and Versatility. *Chem. Soc. Rev.* **2006**, *35*, 1068.

(26) Cheng, C.; Helderman, F.; Tempel, D.; Segers, D.; Hierck, B.; Poelmann, R.; van Tol, A.; Duncker, D. J.; Robbers-Visser, D.; Ursem, N. T. C.; van Haperen, R.; Wentzel, J. J.; Gijsen, F.; van der Steen, A. F. W.; de Crom, R.; Krams, R. Large Variations in Absolute Wall Shear Stress Levels within One Species and between Species. *Atherosclerosis* **2007**, *195*, 225–235.

(27) Valeur, B. Molecular Fluorescence Principles and Applications. *Mol. Fluoresc.* **2001**, 399.

(28) Albani, J. R. Principles and Applications of Fluorescence Spectroscopy, 1st ed.; Wiley, 2007.

(29) Lakowicz, J. R.; Masters, B. R. Principles of Fluorescence Spectroscopy, Third Edition. *J. Biomed. Opt.* **2008**, *13*, 029901.

(30) Weber, G.; Teale, F. W. J. Determination of the Absolute Quantum Yield of Fluorescent Solutions. *Trans. Faraday Soc.* **1957**, *53*, 646.

(31) Kamlet, M. J.; Abboud, J. L.; Taft, R. W. The Solvatochromic Comparison Method. The Pi* Scale of Solvent Polarities. *J. Am. Chem. Soc.* **1977**, *99*, 6027–6038.

(32) Jares-Erijman, E. A.; Jovin, T. M. FRET Imaging. *Nat. Biotechnol.* **2003**, *21*, 1387–1395.

(33) Yuan, L.; Lin, W.; Zheng, K.; Zhu, S. FRET-Based Small-Molecule Fluorescent Probes: Rational Design and Bioimaging Applications. *Acc. Chem. Res.* **2013**, *46*, 1462–1473.

(34) Zeng, N.; Liu, L.; McCabe, M. G.; Jones, D. T. W.; Ichimura, K.; Collins, V. P. Real-Time Quantitative Polymerase Chain Reaction (QPCR) Analysis with Fluorescence Resonance Energy Transfer (FRET) Probes Reveals Differential Expression of the Four *ERBB4* Juxtamembrane Region Variants between Medulloblastoma and Pilocytic Astrocytoma. *Neuropathol. Appl. Neurobiol.* **2009**, *35*, 353–366.

(35) Lu, J.; Owen, S. C.; Shoichet, M. S. Stability of Self-Assembled Polymeric Micelles in Serum. *Macromolecules* **2011**, *44*, 6002–6008.

(36) Robin, M. P.; Osborne, S. A. M.; Pikramenou, Z.; Raymond, J. E.; O'Reilly, R. K. Fluorescent Block Copolymer Micelles That Can Self-Report on Their Assembly and Small Molecule Encapsulation. *Macromolecules* **2016**, *49*, 653–662.

(37) Chen, H.; Kim, S.; Li, L.; Wang, S.; Park, K.; Cheng, J.-X. Release of Hydrophobic Molecules from Polymer Micelles into Cell Membranes Revealed by Forster Resonance Energy Transfer Imaging. *Proc. Natl. Acad. Sci.* **2008**, *105*, 6596–6601.

(38) Zeng, Z. Recent Advances in PEG-PLA Block Copolymer Nanoparticles. *Int. J. Nanomedicine* **2010**, 1057.

(39) Wang, J.; Li, S.; Han, Y.; Guan, J.; Chung, S.; Wang, C.; Li, D. Poly(Ethylene Glycol)–Polylactide Micelles for Cancer Therapy. *Front. Pharmacol.* **2018**, *9*, 202.

(40) Yasugi, K.; Nagasaki, Y.; Kato, M.; Kataoka, K. Preparation and Characterization of Polymer Micelles from Poly(Ethylene Glycol)-Poly(d,l-Lactide) Block Copolymers as Potential Drug Carrier. *J. Controlled Release* **1999**, *62*, 89–100.

(41) Yamamoto, Y.; Yasugi, K.; Harada, A.; Nagasaki, Y.; Kataoka, K. Temperature-Related Change in the Properties Relevant to Drug Delivery of Poly(Ethylene Glycol)–Poly(d,l-Lactide) Block Copolymer Micelles in Aqueous Milieu. *J. Controlled Release* **2002**, *82*, 359–371.

COMPUTATIONALLY ASSISTED DESIGN OF A STAPLED PEPTIDE BUNDLE

Introduction

Coiled-Coil Peptides. Non-covalent interactions are ubiquitous in chemical systems and serve as the primary driving force for protein folding and behavior.[42–44] The complex nature of these interactions leads to highly varied protein structures, with the most common secondary structure being alpha helices.[45] While the high prevalence of alpha helices suggests great stability, often these forms are unstable in aqueous environments if isolated and require interaction from nearby chains to be energetically favorable.[46,47] To counteract these stability problems, certain alpha helical peptides of two or more subunits will coil around each other; these structures are termed coiled-coils.[48] This process is driven by non-covalent interactions which greatly enhance stability of helical folding in comparison to the monomeric counterpart of these structures.[49] In some cases coiled-coils consist of identical chains self-assembling, while other times it is driven by peptides with different sequences.[50,51] It should be noted that this aggregation is not random as these structures are highly ordered and form discrete arrangements.[52] Coiled-coils are highly widespread in biological systems and are predicted to be present in ~10% of all eukaryotic proteins.[53,54] They can be seen as small motifs in sections of a larger protein or as isolated strands consisting only of a coiled-coil, for example, keratin and tropomyosin.[46] Another common example of coiled-coils is the 'leucine-zipper', where the interior consists of alternating leucine residues. This is a structure commonly observed in transcription factors as a binding module for DNA (Figure 18).[55,56]

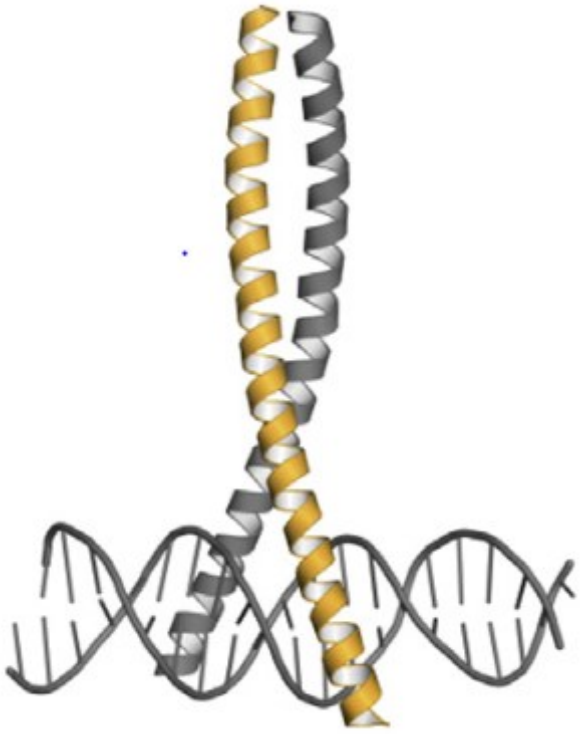

Figure 18. Schematic representation of a coiled-coil protein binding to DNA and acting as a transcription factor. Adapted from Landshultz *et al*.[55]

Self-Assembly of Coiled-Coils. Not all alpha helical sequences will spontaneously form coiled-coils, there are several rules which are commonly observed. Generally, the amino acid sequence follows a seven-residue (heptad) repeat of HPPHPPP, where H is a hydrophobic residue and P is a polar residue.[57] Relative positions in the sequence are denoted by letters ***a-g*** (Figure 19). The heptad segments do not have to be identical in sequence, they only need to have similar properties at the same relative positions. The primary driving force for coiled-coil formation is hydrophobic interactions between nonpolar side chains.[57] A common property of these is that for every heptad repeat there are two hydrophobic amino acids (***a*** and ***d*** positions) for an average of one nonpolar residue per 3.5 residues. Since there are 3.6 residues in an alpha helical turn, a non-linear pitch of hydrophobic amino acids is created across the peptide.[57] The pitch of hydrophobic residues across the peptide face provides steric conditions appropriate for

coiling while simultaneously increasing stability through desolvation of water from hydrophobic residues.[58] Coiled-coils have a wide variety of oligomerization states, which is primarily dependent on the pitch of hydrophobic residues.[59] In natural systems dimers, trimers and tetramers dominate, but bundles of up to 10 monomers have been discovered.[48,60,61]

Rational Design of Coiled-Coils. Protein design and structure prediction is an ever-growing field of research because of the multitude of applications in biochemistry, biophysics, and medicine.[62,63] While predictive methods have significantly increased in recent years, notably because of the advance of computational techniques, protein design and structure prediction are still an incredibly difficult challenge.[64] However, coiled-coils are one of the most well understood systems.[59] This can be attributed to their simplistic and highly repetitive nature. The ability to reasonably predict structural behavior from several fundamental principles has made these materials strong candidates for *de novo* protein design.[60] Fundamental investigation into properties of these relatively simple complexes paves the way for understanding and design of higher-level protein structures.

As previously mentioned, coiled-coil proteins follow a repetitive heptad sequence which contains hydrophobic residues every 3.5 amino acids on average (*a* and *d* positions). This hydrophobic interaction is the primary driving force for coil formation.[56] However, it is not only ***a*** and ***d*** residues that are involved in helix stabilization. All other residues have stabilizing roles as well, even though they are less critical for structural integrity. Residues at the *e* and *g* positions are usually charged or highly polar and are involved in ionic or dipole-dipole interactions between chains which further stabilizes the coil.[59] Amino acids at the ***b***, ***c***, and ***f*** positions are more varied than other locations since they do not closely interact with adjacent chains and mainly interface with bulk solvent.[59] While there is more flexibility in amino acids at

these three locations, normally residues are polar or simply facilitate alpha helix formation, such as alanine. It should be noted, glycine and proline are not viable at any location since they have detrimental impacts on alpha helical structure.[57] Cysteine is another residue that is sometimes avoided since it readily forms disulfide linkages or is oxidized. However, this unique behavior has led to several specific applications such as crosslinking polypeptide chains.[65]

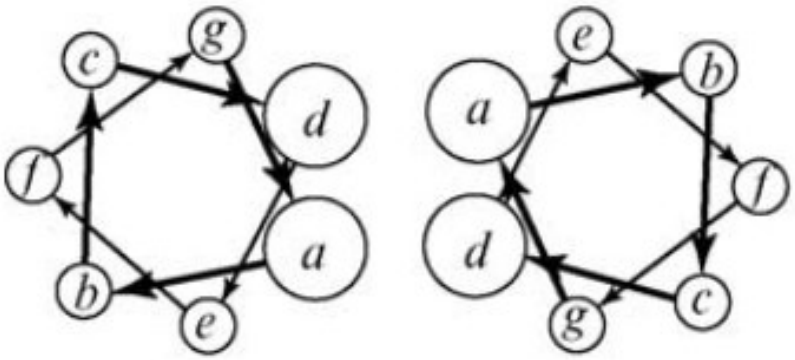

Figure 19. Schematic representation two peptide coiled-coil. The 7-residue amino acid repeat is denoted ***a*-*g***, with the ***a*** and ***d*** positions containing hydrophobic residues. Figure adapted from Woolson *et al.* (2005).[57]

Designed Coiled-Coils Derived from Natural Systems. Some of the earliest attempts to design coiled-coils were done by modifying well-studied natural systems. An example of this is the parallel, dimeric coiled-coil GCN4 that acts as a transcription factor in yeast.[66] This complex is a highly stable dimer that follows typical heptad repeats with the letters ***a*-*g*** representing relative positions in the sequence. Like most coiled-coils, ***a*** and ***d*** are hydrophobic residues. Since this complex is normally a dimer, a simple design approach is to make small modifications that 'destabilize' the structure.[67] The ultimate goal is to form a higher oligomerization state that is more favored. Once these sequences are known then behavior can be rationalized, and the system can be modified further.

One of the first examples of a successfully designed and characterized trimeric coiled-coil trimer derived from GCN4 was reported by DeGrado and coworkers.[67] They created a small library of GCN4 derivatives and examined the resulting structures, with a primary goal to prove that modifications of this natural system could lead to the formation of stable trimers. In early work there was difficulty in forming stable trimeric states, as several of the peptides tested were in equilibrium between dimers, trimers, and tetramers.[67,68] However, after structural modifications they were able to isolate a stable trimeric state. In the native structure each hydrophobic position contains leucine, in the modified structure each ***a*** position is changed to valine while ***d*** remains leucine. It was also found that including a single serine residue at the ***f*** position caused formation of an all-parallel trimer, likely due to electrostatic repulsions.[67] The full amino acid sequence and corresponding residue locations are show in Figure 20.

Another notable finding from this work is that charged residues are major contributing factors to coiled-coil stability and oligomerization state. While the hydrophobic interactions are the primary forces inducing self-assembly, depending on the location and interaction of charged species the complex can be further stabilized or heavily destabilized.[68] As would be expected, destabilization occurs when residues holding the same charge are in proximity, while the same case with opposite charges will increase stability. A top-down visualization of the orientations and interactions between neighboring amino acids is depicted in Figure 21.

g		a	b	c	d	e	f	g		a	b	c	d	e	f	g		a	b	c	d	e	f	g		a	b	c	d	e	f	g
E		V	E	A	L	E	K	K		V	A	A	L	E	S	K		V	Q	A	L	E	K	K		V	E	A	L	E	H	G

Figure 20. Sequence of a rationally designed trimeric peptide bundle derived from the naturally occurring GCN4 transcription factor.[67]

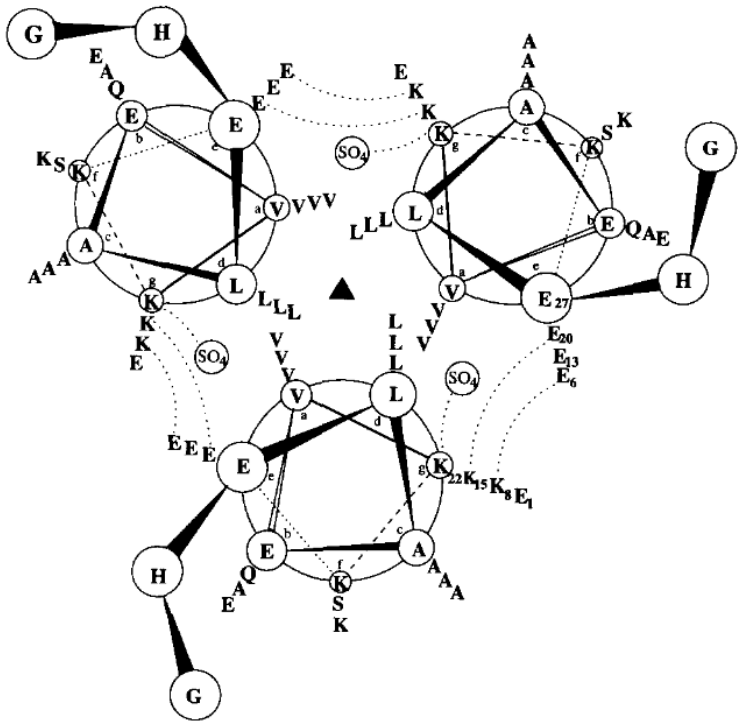

Figure 21. Top-down depiction of non-covalent interactions which stabilize the designed trimer coiled-coil reported by DeGrado and coworkers.[67]

Flexible Coiled-Coil Design Templates. There is a growing variety of modern approaches for designing *de novo* coiled-coil peptides, from large scale computational screening to utilizing basic patterns derived from natural systems.[51,69] Building upon previous research, a simple approach is to take a model heptad sequence and modify select components of it. One example is a scheme reported by Bromley *et al*. where a EΦAAΦKX sequence is superimposed onto a ***gabcdef*** heptad repeat.[69] Φ represents either leucine or isoleucine and the X position is generally either glutamine or lysine to facilitate water solubility. It should be noted that Φ is always a hydrophobic residue and falls on an ***a*** or ***d*** position, following patterns found in natural systems. Depending on the variation at these three positions, this template will facilitate formation of dimers, trimers, or tetramers.[69] Interestingly, the difference in sequence between these three oligomerization states is very small. Simply changing whether Φ is leucine or isoleucine is often

enough to alter the favored number of oligomers. This is mainly attributed to steric effects that change the angle of the hydrophobic pitch across the peptide face.[59] There are some other small modifications to sequence that will favor certain conformations as well. For example, placing a single asparagine residue at the ***a*** position will heavily favor dimerization. Similarly, inserting an asparagine residue at the ***d*** position will prefer a trimer state.[69] This simple template serves as a flexible scaffold for coiled-coil structure design.

Woolfson Peptide Bundle. In 2014 Dek Woolfson and coworkers reported the computational design of a small library of *de novo* designed peptide bundles that forms tetramers, pentamers, hexamers, and heptamers.[51] Prior to this work, almost all coiled-coil peptides were designed by modifying sequences found in natural systems. The peptides by Woolfson and coworkers are not designed based off peptides existing in nature. However, those peptides were designed by taking inspirations from the sequence patterns commonly found in the known peptide bundles. Heptad repeats were used with hydrophobic residues at the ***a*** and ***d*** positions with polar or charged residues occupying the ***e*** and ***g*** locations. The computational screening process allowed for the evaluation of many more candidate structures, leading to successful design of coiled-coils that are at higher oligomerization states than typically observed in nature (5 monomers or higher).[51] For smaller coiled-coils such as dimers or trimers residues at ***b***, ***c***, and ***f*** positions do not largely interact with adjacent chains. However, as higher oligomer states are reached the ***b*** and ***c*** residues become involved in chain stabilization since the angle between adjacent residues becomes larger and interactions near the exterior peptide face are increased. A visual depiction of this is shown in Figure 22. Because of this behavior, the system is better modeled as the supercoiling of two separate dimers. By evaluating the stability of two separate 'seams' (***deab*** and ***cdga***) with adjacent chains a huge number of candidate structures can

be quickly analyzed. To narrow down the search, only amino acids with relatively high prevalence in coiled-coil structures were used; including A, E, I, K, L, N, Q, R, S, and V.[70]

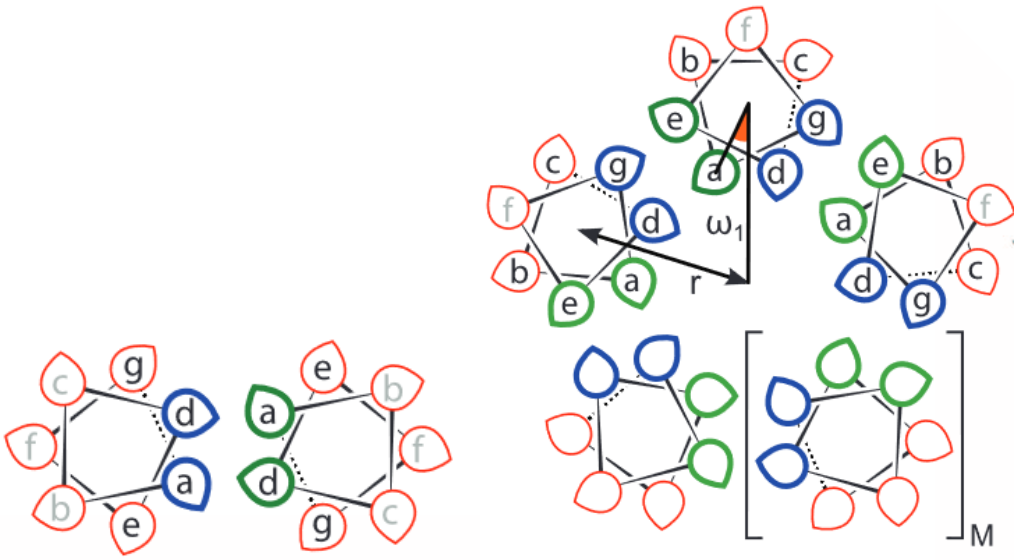

Figure 22. Orientation of heptad repeats in dimer (left) and pentamer (right) in coiled-coils. Figure adapted from Thomas *et al*. (2014).[51]

By varying the amino acid present in the heptad sequence, over 10^6 possible structures were produced. Most of these systems formed commonly seen tetramers or trimers, so polar residues at the ***e*** or ***g*** positions were substituted for hydrophobic residues to facilitate formation of higher oligomerization states. Once initial coiled-coils were screened, molecular dynamics simulations were performed on the final candidates and top structures were synthesized and characterized. A stable pentamer, hexamer, and heptamer were produced from this work. Their structures are shown in Figure 23.

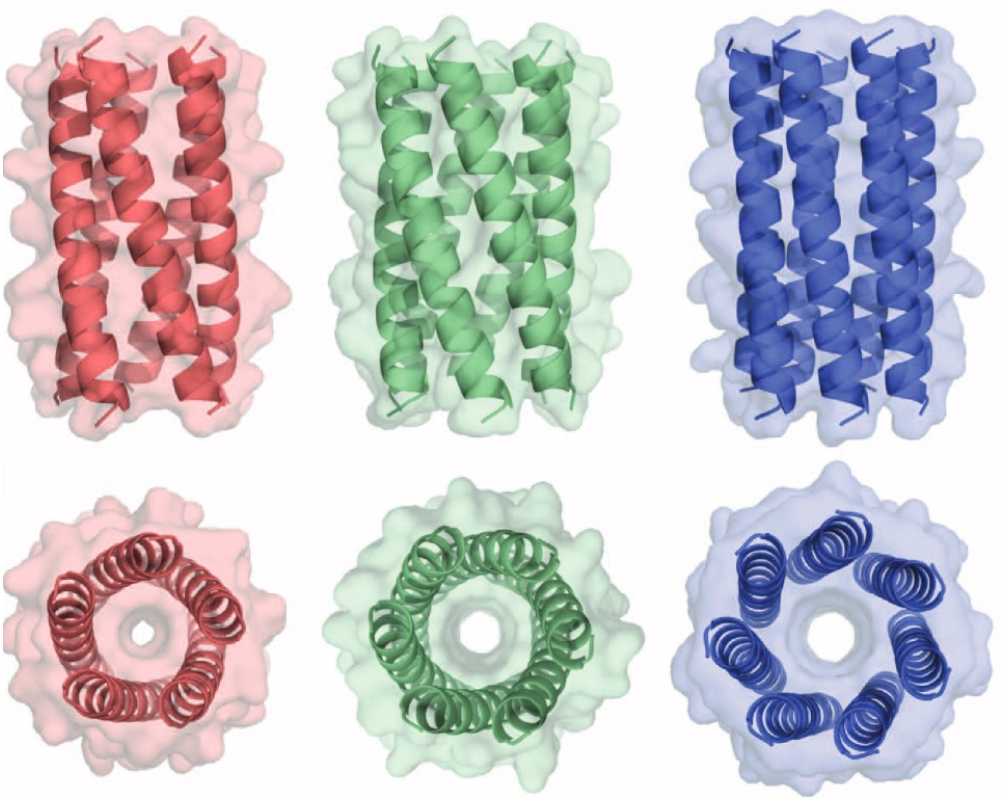

Figure 23. Coiled-coil bundles designed computationally. Pentamer (red), hexamer (green) and heptamer (blue). Figure adapted from Thomas *et al*. (2014).[51]

One interesting feature of these peptide bundles is that they contain an open internal pore, compared to a solid hydrophobic core seen in coiled-coils consisting of four monomers or less.[71] To follow up the computational work, the same group reported these structures and tested the ability to selectively incorporate small molecules into the internal pore (depicted in Figure 24). It was found that these peptide bundles are able to spontaneously take up small, hydrophobic molecules. These affinities are not the same across oligomer states and vary greatly. Since the core of each of these coiled-coils is hydrophobic and contains similar residues, this process is primarily limited by steric repulsion with the most important factor being pore diameter.[71]

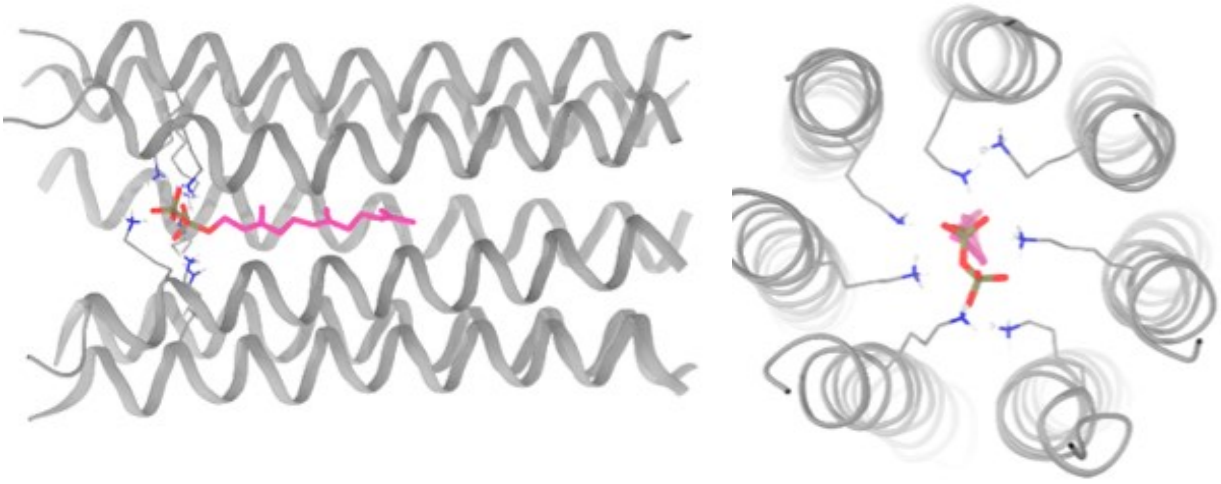

Figure 24. *De novo* designed hexamer coiled-coil that can incorporate appropriate small molecules into the core, in this case retinol phosphate. Figure adapted from Thomas *et al*. (2018).[71]

While most research on coiled-coils mainly focused on fundamental aspects the system to study protein design and structure prediction, there are several unique application areas where coiled-coils are being employed.[58] One advantage to these systems is that they are non-toxic to biological systems since they consist solely of amino acids. The development of coiled-coils that contain a relatively large internal pore allows for semi-selective binding of small molecules, with potential uses as a nanocontainer or toxin neutralization tool.[71] Outside of simple molecular binding, it has also been shown that these can be used as ion channels in cellular membranes.[72] With this said, the introduction of foreign protein sequences into humans often triggers a significant immune response so applications in medicine may be limited unless this problem can be circumvented.[73] As the fundamental knowledge and design strategies for these systems increases, the computational design of coiled-coils seems to be promising for increasingly advanced applications in coming years.

Photoregulation of Peptide Function Using Azobenzene. Properties of Azobenzene.

Azobenzene is a small molecule consisting of two phenyl rings bound to a central azo group (Figure 25). It is highly conjugated and readily absorbs light in the ultraviolet to visible region.[74] Azobenzene isomerizes between its *trans* and *cis* conformation upon UV irradiation. The *trans* conformation is the thermally favored isomer with 99.9% prevalence under standard conditions.[74] However, when formed the *cis* conformation is kinetically-stable and decays slowly back to the *trans* state in the dark. This decay can be accelerated through exposure to appropriate wavelengths of light. The mechanism for isomerization is under debate, and likely can proceed through multiple pathways (rotation or inversion).[75]

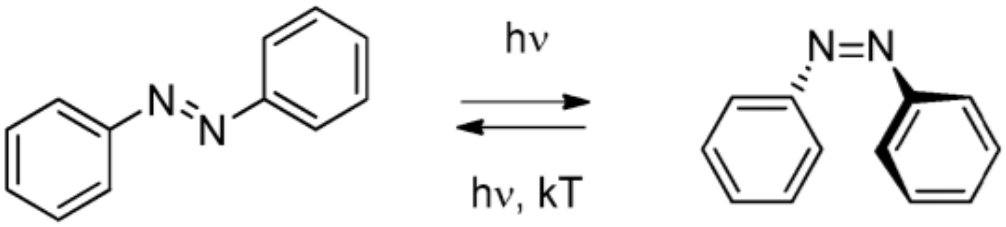

Figure 25. Reversible photoisomerization of azobenzene upon irradiation of ultraviolet light. Figure adapted from Bandara *et al*. (2012).[74]

The most simple azobenzene moiety, shown in Figure 25, will undergo a *trans* → *cis* conversion upon irradiation of ~320 nm light and will undergo a *cis* → *trans* switch at ~440 nm.[76] These wavelengths correspond to the lowest electronic transition for either case, pi-pi*.[76] However, depending on the functional groups bound the absorption wavelength can drastically change. For example, the addition of amine groups at *ortho* or *para* positions causes the absorption band to significantly redshift.[75] Introduction of substitution groups to the phenyl ring that contributes resonance will have a similar effect. Since ultraviolet light can be harmful to

biological systems, an azobenzene derivative has been developed by Woolley *et al*. that absorbs light at 520 nm which is in the visible light region.[77] It contains methoxy groups at the *ortho* position and an acetylated amine group at the *para position* on both phenyl rings (Figure 26). The large resonance contribution from these added functional groups is responsible for the ~200 nm redshift relative to the native structure.[77]

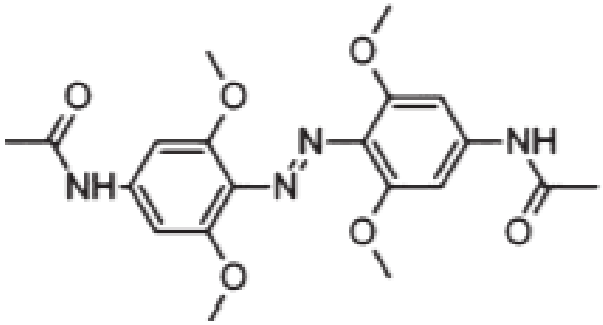

Figure 26. Azobenzene derivative that will undergo a *trans* → *cis* conformational change at 520 nm, outside of the ultraviolet spectrum. Figure adapted from Beharry *et al*. (2011).[77]

Azobenzene as a Peptide Linker. The ability to control *trans* and *cis* conformation of azobenzene has led to many applications in developing chemical systems such as molecular switches, photoresponsive micelles, and light induced phase changes.[78,79] Among those, one of the most notable applications is as a photosensitive peptide linker.[75] There is a growing library of azobenzene derivatives that have been reported to date, several of which have been designed to react with the thiol groups present in the cysteine residues on proteins.[75,80] Coupling an appropriate azobenzene derivative to two locations on the peptide allows reversible photocontrol over structural stability. When a conformational switch is triggered using UV light the native peptide structure is significantly disrupted (Figure 27). Typically, this has been used to reversibly inhibit or generate protein function.[76] There are other interesting cases, such as changing an alpha helix to a beta sheet upon azobenzene isomerization.[81]

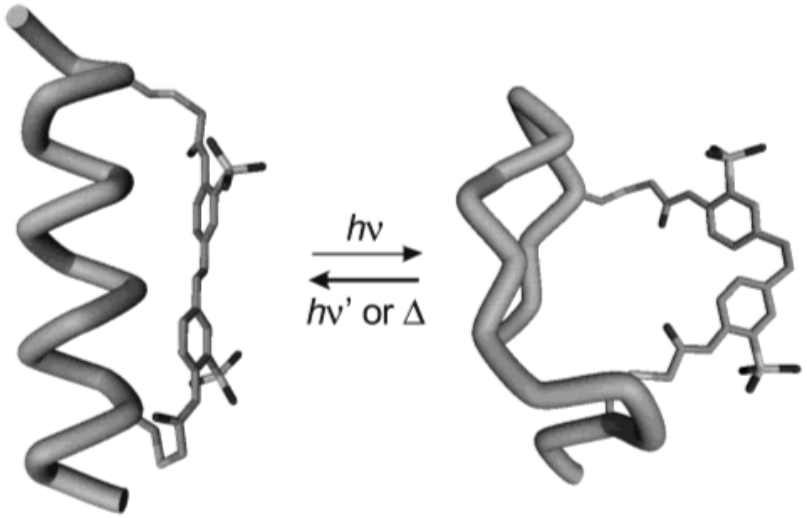

Figure 27. An azobenzene derivative coupled to two cysteine residues on an alpha helical peptide. The structure is shown before and after *trans* → *cis* isomerization.[75]

One of the most used peptide linkers is 3,3'-bis(sulfonato)-4,4'-bis(chloroacetamido) azobenzene (BSBCA; the structure is shown in Figure 28). Notably, the structure contains a chlorinated acetyl group and sulfate group to improve water solubility. The thiol group present in cysteine residues will readily undergo an S_N2 reaction with this chlorinated moiety to bind with the linker molecule (Figure 29). While the experimental procedure for the coupling is relatively simple, there are several factors to consider when trying to design such a system. First of all, to successfully bind BSBCA to two independent locations distance and steric hinderance must be taken into consideration. For the *trans* isomer to react, cysteine residues need to be 11 amino acids apart, denoted as *i*+11.[80] The *cis* isomer can react with residues at *i*+4 or *i*+7 positions.[80] These spacing concerns will change depending on the azobenzene derivative in question. The cysteine residues also need to be readily accessible for binding to occur. For example, if the groups are buried inside a globular protein, even if spaced correctly, binding cannot occur.

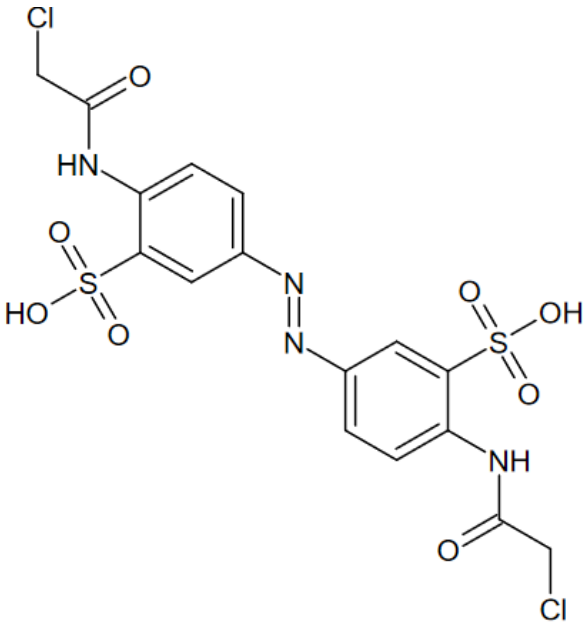

Figure 28. Chemical structure of 3,3'-bis(sulfonato)-4,4'-bis(chloroacetamido) azobenzene (BSBCA), a popularly used photoresponsive peptide linker.

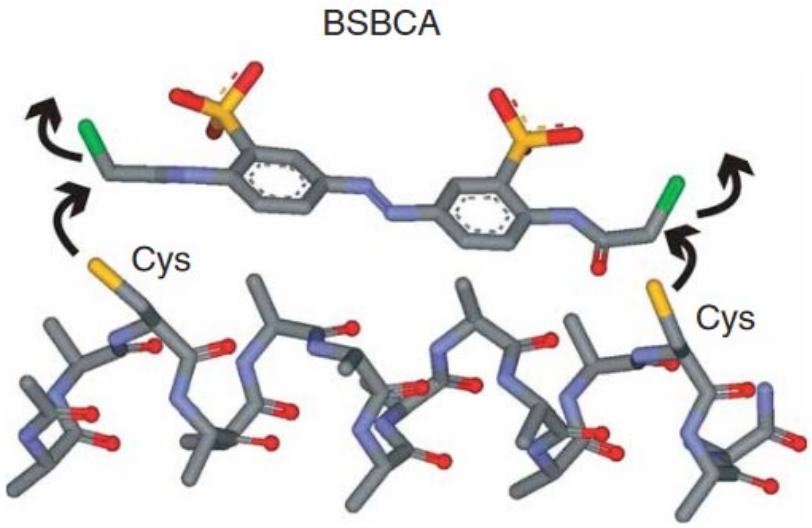

Figure 29. Binding mechanism of BSBCA to appropriately spaced cysteine residues on an alpha helical peptide. Figure adapted from Burns *et al*. (2007).[80]

Photocontrol of Coiled-Coil Activity Using Azobenzene. Woolley and coworkers have recently reported utilizing an azobenzene derivative to establish photocontrol of a coiled-coil peptide.[50] The coiled-coil is AP-1 protein and part of the bZIP protein family. These proteins

bind to DNA and function as transcription factors. The previously discussed peptide linker, BSBCA, was coupled to cysteine residues present on the exterior coiled-coil face. To determine the effects of azobenzene isomerization on activity of the coiled-coil, this system was tested in human HEK293T cells. Plasmids containing a green fluorescent protein (GFP) reporter and AP-1 binding region were transfected into these cells. The azobenzene-linked coiled-coils were then incorporated into the cells to observe whether they would readily inhibit GFP synthesis with the linked azobenzene moiety in a *trans* or *cis* state. Under typical conditions, AP-1 will bind to its associated promoter region and promote RNA synthesis, leading to synthesis of the protein of interest, in this case GFP. Since GFP is inherently fluorescent, its production and relative conformation can be monitored using fluorescence imaging.[82] Results show that azobenzene in a *cis* conformation did not affect activity of AP-1, and GFP synthesis was heavily inhibited. However, in the *trans* conformation activity of AP-1 was decreased drastically, due to the structural changes induced by azobenzene isomerization. While in this work researchers used a *cis* → *trans* conformational change to decrease peptide activity, often a *trans* → *cis* change will have the same effect. This difference is mainly due to the initial conformational state when binding occurs. For example, if azobenzene is bound in the *trans* form then a conformational change to the *cis* state is likely to cause structural change.[80] The opposite case is also true if it is bound in the *cis* form.

Rosetta/PyRosetta. Computational Protein Folding Prediction. With the ever-growing advance of computing power, computational methods are being increasingly utilized to study chemical and biological systems. These techniques are especially useful with complex biological systems such as proteins, lipid membranes, and crowded cellular environments.[83–85] One of the most interesting challenges has been the prediction of protein folding based only on the amino

acid sequence, known as the protein folding problem.[43,64] There has, however, been significant advances in this area over the last several years. Structures submitted to the Critical Assessment of Protein Structure Prediction competition have been predicted with up to 90% accuracy relative to unreleased crystallographic data .[86] One of the consistently best performing computational suites over the last decade is Rosetta, pioneered by David Baker and coworkers at the University of Washington. This software suite utilizes knowledge-based energy calculations along with probabilistic algorithms to attempt to predict a native structure.[84] In recent years AlphaFold, created by Google's subsidiary company Deep Mind, has shown the superior performance in this area by utilizing machine learning.[87] While it has performed exceedingly well, it is not available for public or academic use so will not be discussed in depth.

The Rosetta Computational Suite. Rosetta computational suite is best known for its success in protein folding prediction, protein design, and protein-ligand modeling.[88] However, the software is written in C++. In contrast to many current computational packages, it is probabilistic and not deterministic. In other words, it does not attempt to find the native structure of a protein on a single attempt, but instead it calculates protein structures over a large sample size to find the protein structure with the most stable energy level. Rosetta algorithms operate based on one simple observation; virtually all conformations of a protein can be found by varying torsion angles in the protein backbone and side chains.[84] However, proteins tend to only allow certain torsion angles due to steric and free energy reasons which narrows down possible states. By semi-randomly varying torsion angles over many iterations and measuring the resulting energy, a local minimum can be found. To address this limitation, Rosetta utilize a Monte Carlo protocol, where structural changes that lower energy are kept and changes that increase energy are discarded or accepted based on energy-dependent probabilistic criteria.[85] It is

important to note that higher energy structures are not always discarded, for every iteration there is some probability that it is retained. This is to allow protein structures to escape local minimums (Figure 30). Protein design protocol in Rosetta uses essentially the same steps being used in protein folding procedure. However, protein design takes an already known structure and predicts which sequences are most likely to fold into the specified shape.[84] Meanwhile the steps for determining the folding that gives the lowest energy state for each sequence is generally the same.

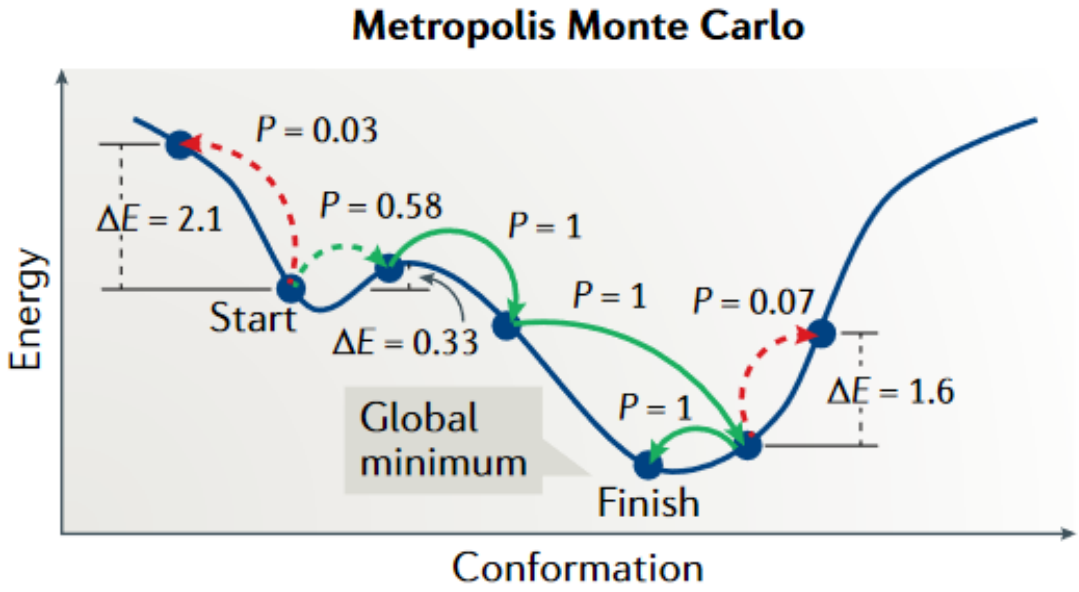

Figure 30. Probabilistic Monte Carlo search to find a global minimum. P represents the probability that the random change will be accepted.[89]

PyRosetta. PyRosetta is a Python-based front-end adaptation of Rosetta, developed at the University of Johns Hopkins.[90] Since the Rosetta suite doesn't have a user interface, Python, a higher-level programming language, allows for a more intuitive programming. Most Rosetta functions are fully available for use in PyRosetta. Since Python is an interpreted programming language, it is slower than a compiled language such as C++. However, most of this computational penalty is circumvented since the energy calculations are still performed using the back-end C++ code so it is about 95% as efficient as vanilla Rosetta.[89,90]

The Rosetta Online Server for Everyone (ROSIE). While most Rosetta functionality can only be run using in-house scripts, an online server has been developed that runs preestablished protocol.[91,92] Rosetta Online Server for Everyone (ROSIE) was built with the intention of lowering the barrier of entry for computational work by providing easy to use functionality. It has a variety of functions that users can use by inputting a protein data bank file (.pdb as file type) of the compound of interest and adjust the associated parameters. This acts as a useful tool to perform basic computational research as these calculations are well established and have proven to produce reliable data.[92] The major downside is that the programs offer little room for customization.

Symmetric docking is a ROSIE operation that is designed for studying symmetric protein assemblies.[93] This often includes coiled-coils, as they are highly symmetrical in nature. From the user side, a single chain of a protein is input, and the number of subunits is selected. The backbone torsion angles are not moved, but side chains are minimized and the orientation of chains in three-dimensional space is changed. Like almost all Rosetta protocols, this is a probabilistic search and often 10,000+ structures are simulated. The population with the highest number of related structures is the most likely to be observed in natural systems (Figure 31). This is typically depicted graphically as free energy plotted against root mean squared deviation relative to the most energetically favored structure or, if the crystal structure has been previously solved, the native protein (Figure 32). Often once the structure begins to approach the conformation of the native protein there is a rapid decrease in free energy and structures begin to converge. Symmetrical multi-unit proteins have been predicted with high accuracy using this method.[93]

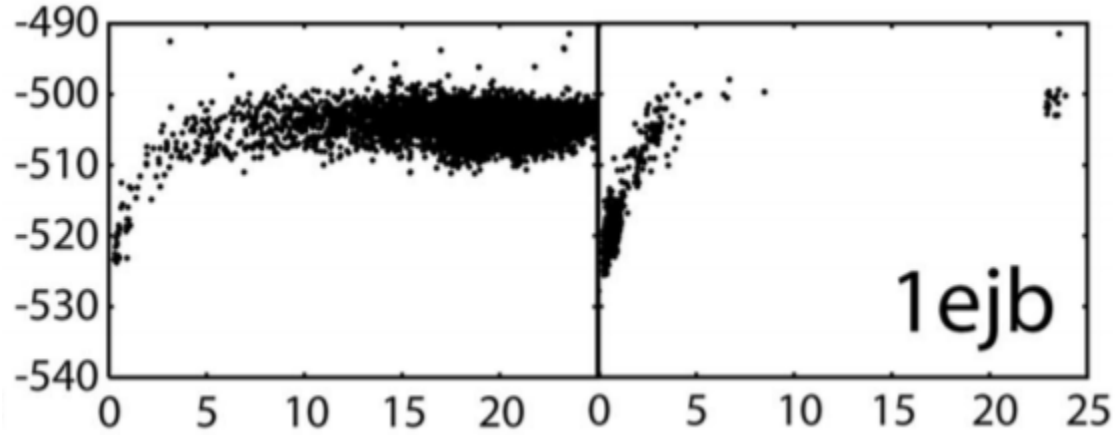

Figure 31. Low-resolution (left) and high-resolution (right) results structure prediction of symmetric proteins using Rosetta. Free energy (arbitrary units) is plotted against root mean squared deviation relative to the native protein. Protein data bank code: 1ejb. Figure adapted from Andre *et al*. (2007).[93]

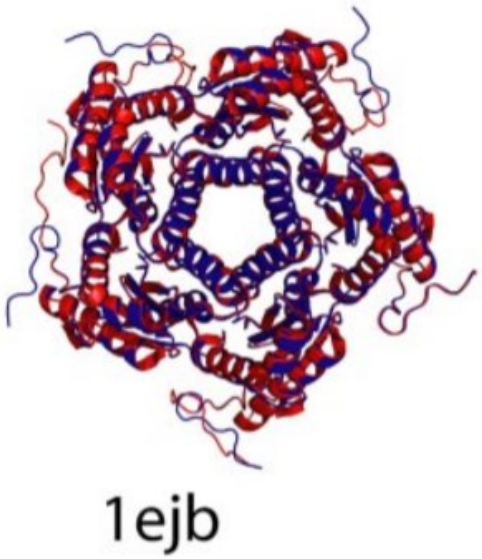

Figure 32. Crystal structure of native protein (red) overlayed with the computationally predicted symmetric oligomer (blue). Protein data bank code: 1ejb. Figure adapted from Andre *et al*. (2007).[93]

Objective

With the growing knowledge of fundamental protein folding principles, as well as the advance of computational techniques, research in *de novo* protein design is gaining momentum.[59,83,85] One of the more recent and advanced coiled-coil systems is the *de novo*

designed trimer to heptamer peptide bundles reported by Woolfson *et al*.[51,71] These peptide bundles contain an internal pore that will readily incorporate small hydrophobic molecules.

The objective of this work is to computationally design a hexamer peptide bundle by engineering previously reported hex_cc peptide hexamer bundle (4PN9) and staple it with a photo-responsive linker, 3,3'-bis(sulfonato)-4,4'-bis(chloroacetamido) azobenzene (BSBCA). This coiled-coil structure contains an open internal pore and will spontaneously incorporate appropriate hydrophobic molecules into the core. Irradiation with UV light will induce isomerization of BSBCA from the *trans*→*cis* state (or vice versa), the switch in isomer state should cause structural degradation and release internalized molecules. BSBCA must be coupled to cysteine residues which are not present in the native structure, so to rapidly screen mutant structures we have applied computational methods. The long-term goal of the project is to develop a photosensitive peptide bundle that is capable of reversible 'catch and release' of small hydrophobic molecules.

Materials

- 2-amino-5-nitrobenzenesulfonic acid (96-75-3)
- 2-acetamido-5-nitrobenzenesulfonic acid
- 2-acetamido-5-aminobenzenesulfonic acid
- Acetic anhydride (108-24-7)
- Glacial acetic acid (64-19-7)
- Palladium on activated carbon (Pd/C, 10% w/w) (12135-22-7)
- Methanol (67-56-1)
- Acetone (67-64-1)
- DMSO-d_6 (2206-27-1)

Methods

Computational Studies. Hardware/Software. All computational works were performed on an Acer Aspire Slim 5 laptop with an AMD Ryzen 5 3500U processor (2100 MHz, 4 cores, 8 logical cores) and 8 gigabytes of RAM. The Rosetta molecular modeling suite with PyRosetta4

interface for Python 3.8 was used in all peptide modeling work. This software was only compatible with Linux-based operating systems. Therefore, Ubuntu 18.04 LTS (64 bits) terminal was installed under Windows Subsystem for Linux (WSL 2) on Windows 10 (version 20H2) and used to run the scripts. Meanwhile, all files were managed by using Windows 10 and Visual Studio Code was used as the code editor.

Factors Taken into Consideration for Peptide Bundle Design. The immediate goal of this computational work was to search for a pair of residues that can be substituted with cysteine in order to conjugate with photo-responsive linker. To this end, I focused on two main factors. First, mutations must not alter the energetically favored oligomeric state of hex_cc (4PN9) hexameric peptide bundle. To maintain the hexameric form the residue substitution should not be significantly energetically unfavored. Alternatively, if residues are highly energetically favored the 'overstabilization' of the monomeric coil could hinder the formation of peptide bundles.[68] As with the simplification of any chemical system, coiled-coils with less chains are more entropically favored. For higher level coiled-coils to form, the entropic penalty must be offset by highly favorable thermodynamic interactions between chains and structures containing less subunits must be relatively unfavored. Often, to promote the formation of larger oligomerization states residues are inserted that do not adhere to the typical behavior seen in these heptad repeats. For example, it has been reported that insertion of an asparagine at the ***a*** position, a position that typically contains nonpolar residues, will tend to form larger coiled-coils.[57,59] This is likely because the free energy penalty caused by the interface of polar and nonpolar residues is avoided by forming a larger structure.

The second factor being considered was that residues must be able to bind the photo-responsive linker, BSBCA. Not all residues are viable substitution points. Since our azobenzene

derivative must be physically bound to the external face of the peptide, it is unlikely internal residues can be coupled to BSBCA because of the steric limitations. Amino acids at the ***b***, ***c***, and ***f*** position are strong candidates for mutation since they are on outside of the peptide bundle and mainly interact with bulk solvent and not adjacent chains.[57,60] The distance between residues also needs to be viable for BSBCA binding. If BSBCA is in the *trans* conformation, this distance should follow the *i*, *i*+11 rule where the amino acids of interest are 11 residues apart. In the *cis* conformation the azobenzene derivative can be coupled to following the *i, i*+4 or *i, i*+7 rules.[80] It should be noted that the *trans* conformation is easier to use in experimental settings because it is the thermally favored isomer and under typical conditions accounts for 99%+ of the isomer states observed in a given sample.[76,78]

Python Codes Used in Computational Work. PyRosetta4 interface can only process protein data bank (.pdb) files, which are text files that contain coordinates of each individual atom in a given protein. Using a Python code developed in this work, amino acid residues in cc_hex2 peptide bundles were sequentially substituted with a single amino acid on each chain at a time with cysteine. The energy function of the mutated peptide bundles was evaluated after each cycle. The general structure of Python code used in almost all scripts in this study is as follows: The first step is to load the pdb file of interest as a "**pose**" (built in PyRosetta function), which loads the protein into memory then prepares it for modification and evaluation. An object-oriented approach was taken where our peptide bundle is loaded as the 'object' and stored in computer memory. This methodology improves code flexibility, as the program can also be applied to any symmetrical coiled coil. In this example, there are four parameters: chain length, the number of chains, residue number to mutate, and amino acid to mutate the residue with.

```
class SingleMutantHexamer:
```

```
    def __init__(self, chain_length, num_of_chain, mut_num1, mut_aa1):
        self.chain_length = chain_length
        self.num_of_chain = num_of_chain
        self.mut_num1 = mut_num1
        self.mut_aa1 = mut_aa1
```

A list of residues to mutate is made and stored in "**res_to_mutate**". The below example could simply be substituted with a for-loop since this script will cycle through every residue in the protein but is left in this more complex state to allow for easy modification. For example, if you wanted to substitute only the ***b***, ***c***, and ***f*** residues this change could easily be made by removing extra code.

```
res_to_mutate = []
res_in_chain = 29
for r in range(0, 5):
    for j in range(0, 5):
        c = (1 + 7 * j + res_in_chain * r)
        d = (2 + 7 * j + res_in_chain * r)
        e = (3 + 7 * j + res_in_chain * r)
        f = (4 + 7 * j + res_in_chain * r)
        g = (5 + 7 * j + res_in_chain * r)
        a = (6 + 7 * j + res_in_chain * r)
        b = (7 + 7 * j + res_in_chain * r)
        if c < res_in_chain:
            res_to_mutate.append(c)
        if d < res_in_chain:
            res_to_mutate.append(d)
        if e < res_in_chain:
            res_to_mutate.append(e)
        if f < res_in_chain:
            res_to_mutate.append(f)
        if g < res_in_chain:
            res_to_mutate.append(g)
        if a < res_in_chain:
            res_to_mutate.append(a)
        if b < res_in_chain:
            res_to_mutate.append(b)
```

The core part of the code is a for-loop that cycles "**res_to_mutate**" and runs several functions during each iteration. Each cycle of the for-loop execute stores the mutation information, mutates the residue of interest, applies gradient-based minimization, and then resets the structure to its native conformation. Notice that out of the four parameters applicable to the "**SingleMutantMonomer**" object the only parameter that changes in each cycle of loop is the mutated residue. Several functions are then applied to the object in cycle of the loop. In cases where it needs to mutate two residues simultaneously an extra for-loop and extra variable were added to the class.

```
for k in res_to_mutate:
    Hex2_smm = SingleMutantMonomer(29, 1, k, 'C')
    initial_mut_num1 = Hex2_smm.mut_num1
    Hex2_smm.mutation_info()
    Hex2_smm.mutate_pose()
    Hex2_smm.packing()
    Hex2_smm.calc_energy()
    pose = pose_from_pdb('file_thread.pdb')
```

Below is the full code of the functions described above. Although it is not explicitly shown below, all of these functions are incorporated within the "**SingleMutantMonomer"** class. Minimization protocol ("**packing"** function) in this case is only applied to the residue that is mutated and the two residues adjacent to it. It should be noted that the if-else statements in the "**packing"** function are for exception handling, namely, to avoid attempting to modify residues that are outside the available residues on the protein.

```
def mutation_info(self):
        if self.mut_num1 <= self.chain_length:
            mutation_name1.append(pose.residue(initial_mut_num1).name() + '_'
            + str(self.mut_num1) + '_' + self.mut_aa1)
```

```
            mut_res_num1.append(self.mut_num1)
        else:
            pass
    def mutate_pose(self):
        mutate_residue(pose, self.mut_num1, self.mut_aa1)
    def packing(self):
        mm1 = MinMover()
        # assign scorefunction to MinMover
        mm1.score_function(scorefxn)
        # generate movemap object/range and assign it to MinMover
        mm1_xx = MoveMap()
        if self.mut_num1 == 1:
            mm1_xx.set_bb_true_range(self.mut_num1, (self.mut_num1 + 1))
        elif self.mut_num1 == 28:
            mm1_xx.set_bb_true_range(self.mut_num1 - 1, self.mut_num1)
        else:
            mm1_xx.set_bb_true_range((self.mut_num1 - 1), (self.mut_num1 + 1)
)
        mm1.movemap(mm1_xx)
        # apply minimization protocol to mutated pose
        mm1.apply(pose)
    def calc_energy(self):
        scores.append(scorefxn(pose))
        energy_change.append(scorefxn(pose) - init_score)
```

During the execution of this code, several lists are being appended for each iteration of the main for-loop. They contain important information for analysis such as the residues mutated during each iteration and the calculated free energy of the mutant protein. These lists are then exported as a comma delimited (.csv) file and processed using Microsoft Excel. The strings inserted into the list simply allow for easier visual analysis once the file is exported. Full source code for several programs is in Appendix B**.**

```
mut_res_num1 = ['Residue Number']
mutation_name1 = ['Mutation Info']
scores = ['Scores’]
energy_change_pack = ['Energy Change Relative to Native Structure’]
def csv_export():
    with open('file_thread.csv', 'w', newline='') as e:
        write = csv.writer(e)
        write.writerow(['Original score =', init_score])
```

```
write.writerow(mutation_name1)
write.writerow(energy_change)
write.writerow(scores)
write.writerow(mut_res_num1)
```

Synthesis of 3,3'-bis(sulfonato)-4,4'-bis(chloroacetamido) azobenzene (BSBCA). 2-acetamido-5-nitrobenzenesulphonic acid (1). The procedure that was previously reported by Derda and coworkers was adopted and used with some modifications for the synthesis of 2-acetamido-5-nitrobenzenesulphonic acid (Figure 33).[94] 2-amino-5-nitrobenzensulphonic acid (2.055 grams, 8.57 mmol) and 40 mL of glacial acetic acid was added to a 100 mL round bottom flask. The round bottom flask was sealed with a rubber septum and acetic anhydride (2.43 mL, 25.7 mmol) was added using a 5 mL syringe over 3 mins. During this addition, a secondary syringe was added to the septum to assure adequate venting. This solution was then transferred to a 115˚C silicon oil bath. The rubber septum was removed from the flask and attached to a Liebig condenser. The solution was constantly stirred at 300 rpm and the reaction was run for 3 hours. The starting reagent was initially insoluble but after ~30 mins all components in the mixture were fully dissolved. To recover 2-acetamido-5-nitrobenzene sulphonic acid from solution phase the round bottom flask was placed into an ice bath. After about 10 minutes the product recrystallized onto the round bottom flask. The mother liquor was poured into a separate beaker and the same recrystallization procedure was performed to remove any remaining product from solution. All recovered solids were scraped into a Hirsch funnel and dried *in vacuo* for 3 hours. It was then transferred to an anhydrous oil pump vacuum chamber and dried overnight to remove any residual acetic acid. The final product is a white, powdery solid.

2-acetamido-5-aminobenzenesulphonic acid (2). The procedure that was previously reported by Derda and coworkers was adopted and used with some modifications for the

synthesis of 2-acetamido-5-aminobenzenesulphonic acid.[94] 2-acetamido-5-nitrobenzenesulphonic acid (1.002 grams, 3.84 mmol) and 55 mL of methanol was added into a 100 mL round bottom flask. The flask was then sealed with a rubber septum and the mixture was purged with argon gas for 10 minutes. A separate 100 mL round bottom flask was sealed with a rubber septum then purged with a steady stream of argon for 10 minutes and 0.3851 grams (0.361 mmol Pb atoms) of Pd/C 10% was added. A secondary needle was placed into both flasks for venting purposes. These purging steps are highly important due to the volatile and flammable nature of the catalyst.

It should be noted, the argon flow rate needs to be relatively low in the flask containing Pd/C to not disturb the fine particles. Should they be suspended into air they could spontaneously combust, especially in the presence of gaseous solvent. The septum was then removed from both round bottom flasks, and both reagents and catalysts were combined into a larger glass container and attached to a Parr high pressure apparatus. The glass container was purged with H_2 gas in the pressure apparatus at 1 atm, and the solution was shaken at room temperature. After 3 hours, the solution was filtered through celite *in vacuo* for 30 minutes to remove the Pd/C catalyst. The recovered filtrate was then concentrated by using a rotary evaporator to remove the solvent. The final product was a light brown solid. The rotary evaporator is not sufficient to remove all solvent, so the sample was additionally dried under high vacuum for 12 hours to remove all residual methanol.

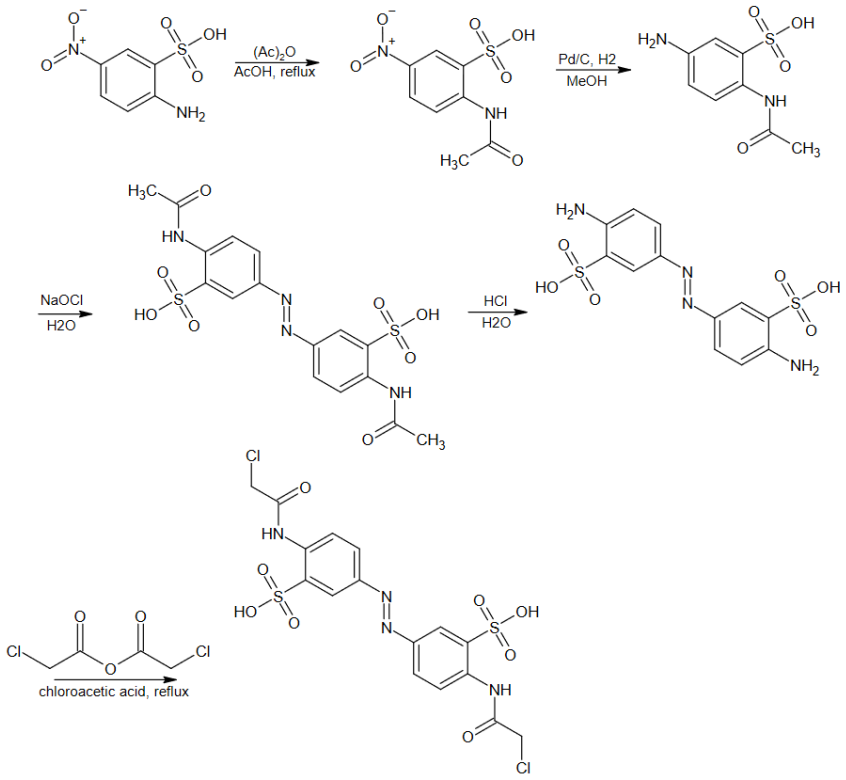

Figure 33. Proposed scheme for synthesis of 3,3'-bis(sulfonato)-4,4'-bis(chloroacetamido) azobenzene (BSBCA) using modified schemes reported by Woolley *et al.* and Derda *et al.*[80,94]

Results & Discussion

Computational Studies. Single Mutant Stability Tests. One of the first computational experiments we performed was examining the effect that mutating individual residues with cysteine has on the energy function of the hex2_cc hexameric peptide bundle. This calculation was performed to evaluate how a single mutation would affect the energy function. As mentioned in an earlier section, we anticipated that residues at, the ***b***, ***c***, and ***f*** positions are the most likely candidates to be substituted by cysteines for covalently linking with BSBCA without destabilization of the hexameric bundle. Meanwhile, in order to gain more insight on the effect of mutations, every single position on the peptide was sequentially mutated. The code follows a straightforward protocol shown in Figure 34. It should be noted that the calculations by Rosetta

yield arbitrary units (Rosetta energy units, REU) and should only be used in a relativistic manner. The results of this are shown in Figure 35. Full source code can be found in Appendix B.

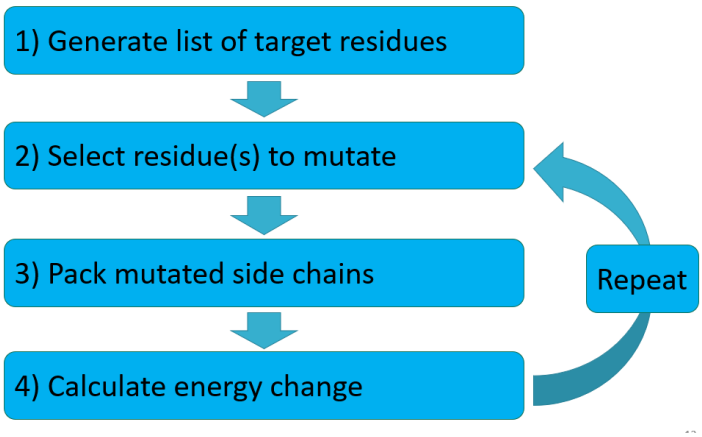

Figure 34. General flow for the PyRosetta mutation program. This cycle is repeated for each residue, resetting to the native structure after each iteration.

The result of the calculation showed several residues that should not be substituted with cysteine (Figure 35). For example, substituting alanine at any position seems to provide a harsh energy penalty. Since alanine is known to promote alpha helix formation, the substitution of alanine with cysteine likely causes sterically unfavored change.[60] In contrast, the mutation of lysine and serine residues are highly energetically favored at all positions. Unsurprisingly, mutations of hydrophobic residues in the core destabilized the helix bundle, resulting in the large increase in the energy function.

1	*c*	GLU	5.7
2	*d*	ILE	0.5
3	*e*	ALA	17.2
4	*f*	LYS	-6.6
5	*g*	SER	-13.3
6	*a*	LEU	23.2
7	*b*	LYS	-36.6
8	*c*	GLU	-15.7
9	*d*	ILE	9.4
10	*e*	ALA	45.7
11	*f*	LYS	-13.8
12	*g*	SER	-20.5
13	*a*	LEU	15.7
14	*b*	LYS	-16.4
15	*c*	GLU	11.7
16	*d*	ILE	21.2
17	*e*	ALA	59.0
18	*f*	TRP	-6.8
19	*g*	SER	-14.4
20	*a*	LEU	20.6
21	*b*	LYS	-11.5
22	*c*	GLU	6.7
23	*d*	ILE	9.5
24	*e*	ALA	32.0
25	*f*	LYS	-15.3
26	*g*	SER	-6.8
27	*a*	LEU	12.5
28	*b*	LYS	-2.1

Figure 35. Energy change from residue mutation for each position on the hexamer bundle. Shown on a green → red gradient corresponding to more negative or more positive values, respectively.

Double Mutant Stability Tests. The single mutant data allowed for gaining insights on the contribution of the residues to the energy function of hexamer bundle. We next calculated the effect of double mutations. This code follows the mostly identical steps as the code for single mutants except that two residues are simultaneously mutated and minimized instead of one. One

major difference with this calculation is on the much larger sample size. While the calculation of single-mutations generated 28 mutant structures, the calculation of double-mutations led to over 300 possible mutants. Given the large nature of this data set, the results were processed into a heatmap using an in-house Python script, with the green→red color gradient corresponding to more negative or more positive values, respectively (Figure 36). Once again, the results indicated that substitution of alanine by cysteine is likely to destabilize the hexamer bundle. Similarly, mutations at residues present in the interior (isoleucine, leucine) of the bundle result in an energy penalty. The substitution of lysine and serine residues with cysteine are typically energetically favored. This result allows for gaining further insights on the general trends on the residues that can be potentially mutated without destabilizing the hexameric bundle.

Free Energy of Bundling. When Rosetta calculates free energy, several parameters are calculated such as Van der Waals forces, ionic interactions, solvent interactions, and hydrogen bonding.[84,89] in order to shorten the calculation time Rosetta also utilizes semi-empirically precalculated values such as table values for amino acid free energy.[89] For this reason, substitution of a residue with a relatively small side chain, such as alanine, with a bulky residue, such as tryptophan, generally has impact on the calculated energy function. Therefore, the energy function does not necessarily reflect the propensity of a peptide to form hexameric bundles. To circumvent this issue, we have estimated the gain in the energy function that comes from the formation of the hexamer bundle from its monomeric state. For this purpose, we used a simple assumption that the gain in the energy function can be estimated by subtracting the energy functions of monomers from the energy function of the hexamer bundle. This approach also allows for of offsetting any differences that are caused solely by residue substitutions. To this end, the information on the shape of hexamer bundles is programmatically truncated to generate

the PDB file that only contains the information of a single peptide chain (Figure 37). The energy function of these monomers was then loaded into PyRosetta4 for the energy calculation (Figure 38, full source code in Appendix B). It is important to note that when this monomer is generated, no minimization is performed. Therefore, the PDB file only contains the relative atomic coordinates of a single chain from the hexamer.

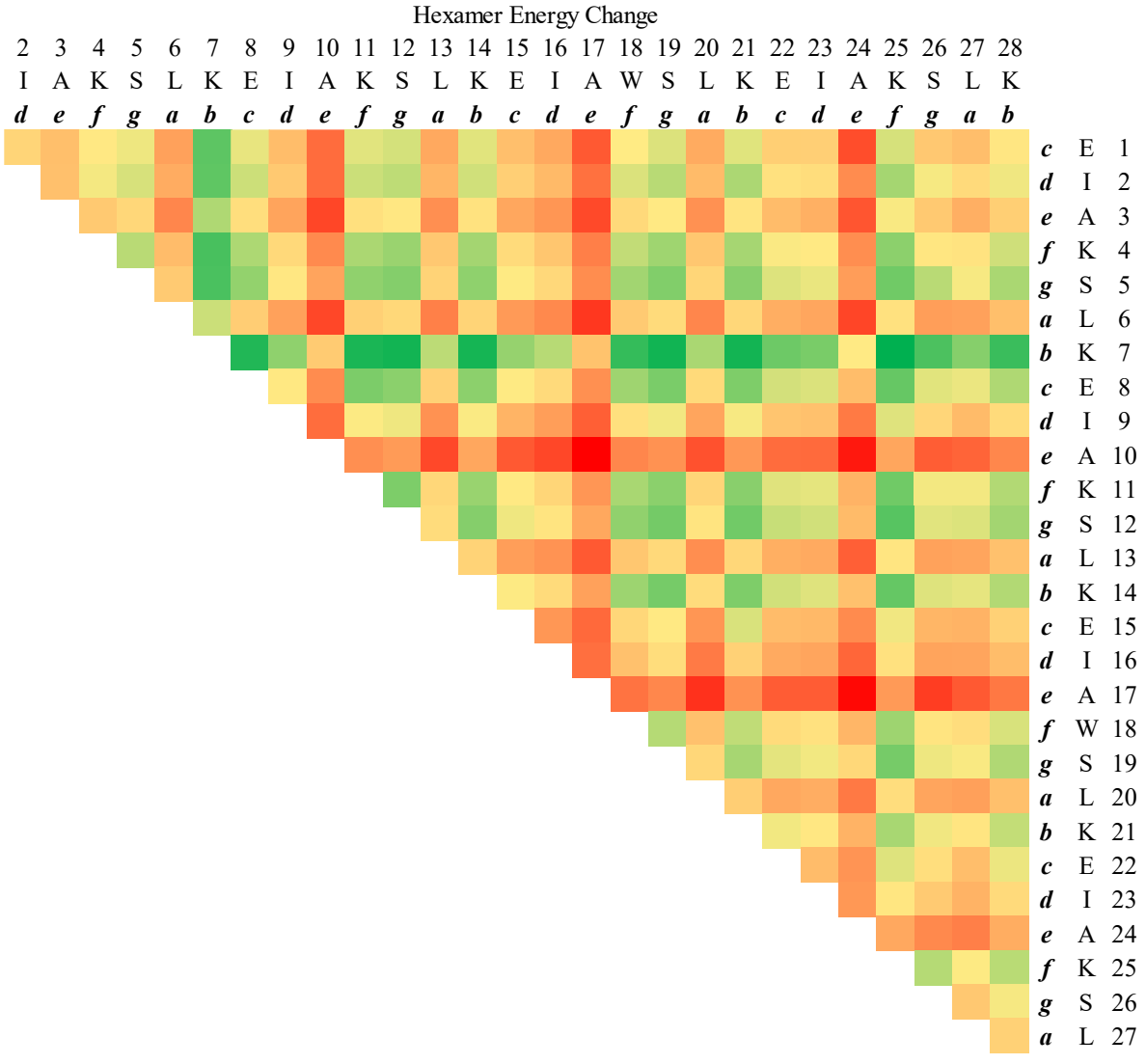

Figure 36. Energy change of the hexamer coiled-coil bundle when mutated with cysteine in two locations. Residue number, one letter amino acid code, and relative position in the heptad repeat (***a***-***g***) are given. Shown on a green → red gradient corresponding to more negative or more positive values, respectively.

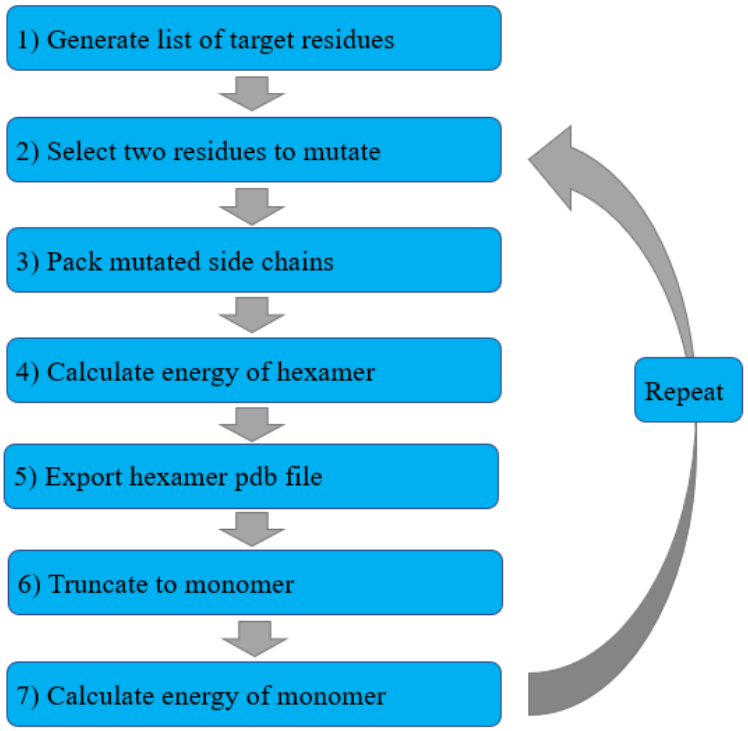

Figure 37. Code methodology to generate and calculate energy of a single monomer from mutated hexamer bundle.

Monomer energy calculations show almost all introduction of cysteine by mutations tend to destabilize the stability of hexamer bundles (Figure 38). This was not particularly surprising considering the system has been optimized for its hexameric form. To calculate the free energy of bundling, which will be referred to as $\Delta G_{bundling}$, Equation 1 is used.

$$\Delta G_{bundling} = G_{hexamer} - (G_{monomer} \times 6) \qquad \text{Eq. 1}$$

By multiplying the monomer times six then subtracting it from the calculated free energy of the mutant hexamer, residual free energy values or extraneous precalculated values that could potentially skew data are offset. In the calculation, absolute values of energy are used instead

relative values. Results show similarity to data from the calculation of double mutants; however, the trends are more pronounced and consistent across rows/columns (Figure 39).

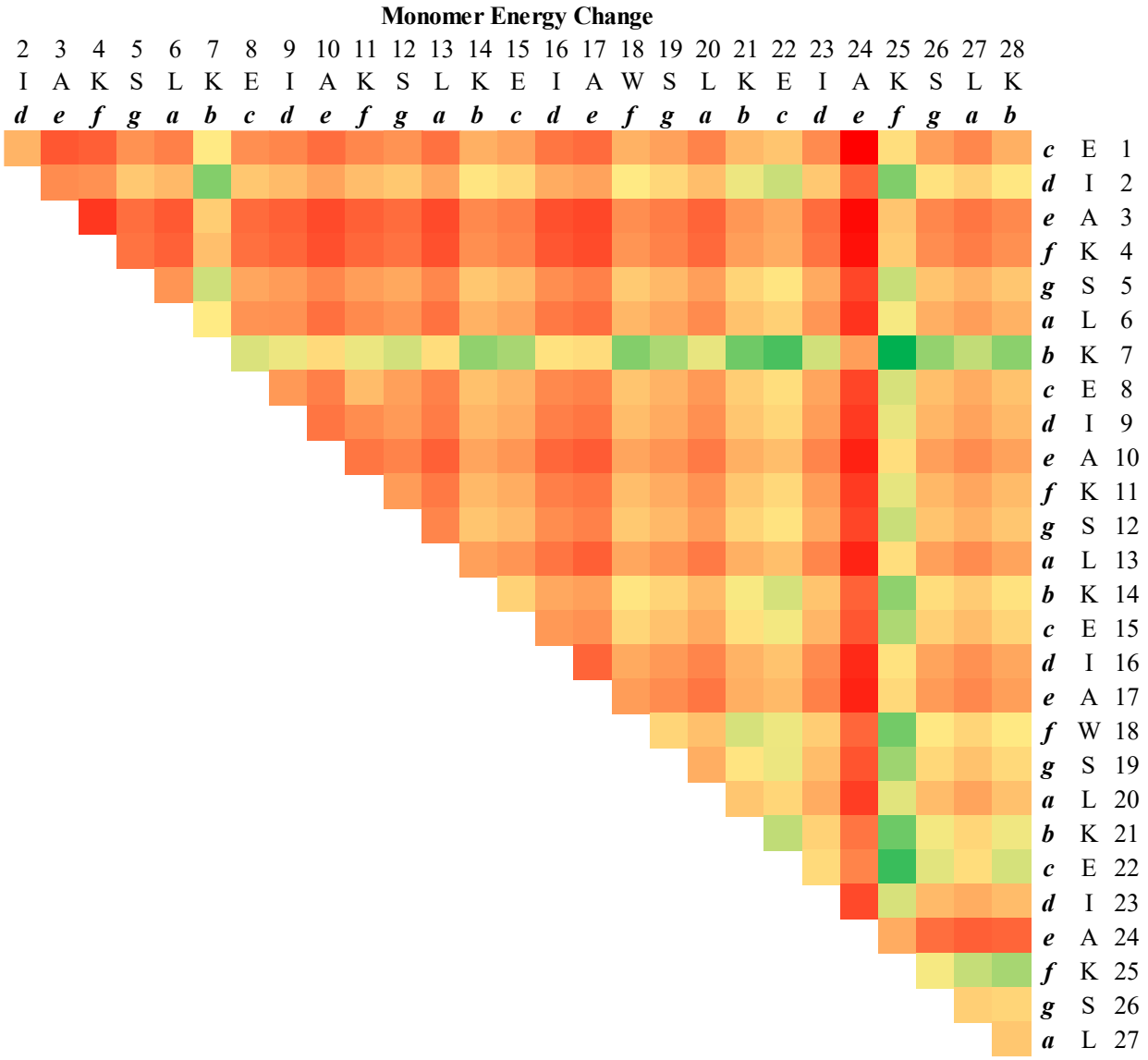

Figure 38. Energy changes of monomeric mutants generated through truncation of the hexamer bundle, compared to the native monomer structure. Shown on a green → red gradient corresponding to more negative or more positive values, respectively.

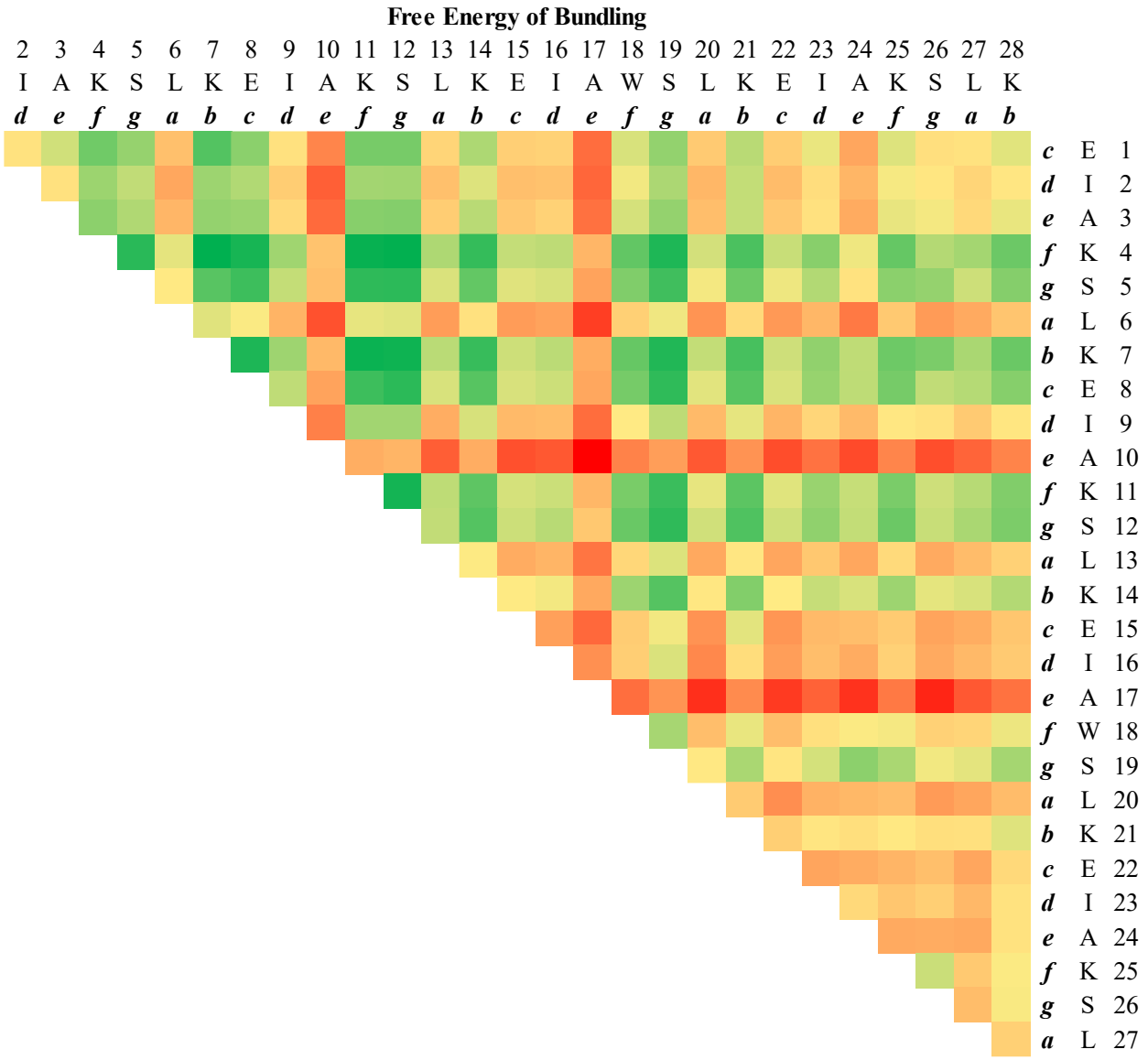

Figure 39. Free energy of bundling for each residue mutation calculated using Equation 1. Shown on a green → red gradient corresponding to larger → smaller changes in free energy.

Protein folding theory explains that native protein conformations are close to optimal.[43,95] It is plausible to assume that substitution at interior residues are already optimized and that substitution will lead to a decrease in stability. Therefore, the exterior of the peptide would be the most favorable location to mutate and have the least impact on bundling and/or tertiary/quaternary structure. In the case of the hexamer bundle, this includes residues at the ***b***, ***c***,

and ***f*** positions. Residues at the ***e*** and ***g*** positions could be viable as well but as the number of oligomers increases these residues become increasingly involved with interchain interactions.[59] The 10 mutants with the largest $\Delta G_{bundling}$ are shown in Table 2 (n =378). Every candidate in this list is at either a ***b***, ***c***, ***f***, or ***g*** position, agreeing closely to intuitive prediction. It is likely that the serine at position ***g*** is more favored than other amino acids at the same position would be since it is structurally analogous to cysteine. Its mirrored counterpart at position ***e*** (alanine) has been shown to be energetically disfavored in virtually all calculations. Overall, this provides a promising set of data from which the final candidates can be narrowed down.

Table 2. Residue information of the top 10 structures based on $\Delta G_{bundling}$ score (arbitrary units). Columns 'Residue 1' and 'Residue 2' contain the single letter amino acid code, residue number, and position in the heptad repeat (***a***-***g***), respectively.

Rank	Residue 1	Residue 2	$\Delta G_{bundling}$
1	K-7-***b***	K-4-***f***	-172.886
2	S-12-***g***	K-4-***f***	-172.506
3	K-11-***f***	K-4-***f***	-171.602
4	K-11-***f***	K-7-***b***	-171.149
5	S-12-***g***	K-7-***b***	-170.46
6	S-12-***g***	K-11-***f***	-169.428
7	E-8-***c***	K-4-***f***	-168.969
8	E-8-***c***	K-7-***b***	-168.004
9	S-19-***g***	K-4-***f***	-167.856
10	S-19-***g***	K-7-***b***	-167.273

Top Candidates for BSBCA Binding and Future Work. Based on the computational work and rational design principles, several top structures for synthesis have been identified. Out of the initial search of 378 structures, candidates can be narrowed down by applying three criteria to our data set. First, the $\Delta G_{bundling}$ must be large relative to other mutants since this is likely to correlate with bundle stability. Second, the two mutated positions must be 11 residues apart, following the i, i+11 rule as BSBCA will readily couple this distance in the *trans* isomer. It

should be noted, in the *cis* state BSBCA can couple at *i*, *i*+4 and *i*, *i*+7 positions but it is more difficult to work with experimentally because *trans* is the thermally favored isomer (99%+ at room temperature). Third, the residues should be at the ***b***, ***c***, or ***f*** position to facilitate binding, though this parameter may be flexible. There is nothing inherently wrong with coupling BSBCA at a different position, but other positions are generally more involved with interchain interactions.[59,65] For example, serine at the ***g*** position is indicated to be a favorable mutation based off computational data, but steric issues are more likely to arise so for purposes of simplification those cases have been excluded.

With these criteria considered, three candidates have been identified, shown in Table 3. All three match the *i*, *i*+11 rule, are mutated at ***b***, ***c***, or ***f*** positions, and are in the top 1/3 of all mutant structures sorted by $\Delta G_{bundling}$. These are all strong candidates for synthesis, though multiple factors need to be weighed outside of free energy of bundling when attempting to compare such systems. ***Candidate #1*** (mutations at residues 7 and 18) seems to have promising properties over the other mutants. However, when the covalent attachment of BSBCA to the peptide is considered, one potential issue is the bulky tryptophan residue at position 18. If this amino acid is in between the two cysteine residues that are being coupled, it may sterically inhibit binding. Since tryptophan is substituted with cysteine in this structure the issue is circumvented. This mutant also has best $\Delta G_{bundling}$ score which ranks in the top 10% of all mutants; these factors combined make it the front runner for synthesis.

Candidate #3 has the second best $\Delta G_{bundling}$ rank out of these three structures, but there are several issues that could affect viability. First, both residues to be mutated are lysine which exhibit a positive charge. While energy calculations are seemingly unphased by this change, intuitively is seems that drastically changing the charge relative to the native structure could

destabilize the hexamer bundle. Additionally, as previously discussed, tryptophan at residue 18 could sterically hinder the covalent linkage of BSBCA with the peptide. This mutant has cysteine residues on either side (positions 14 and 25), which may lead to problems. ***Candidate #2*** has the lowest $\Delta G_{bundling}$ rank out of these three structures but avoids the issues with tryptophan and does not affect net charge since positively and negatively charged species are removed simultaneously (lysine, aspartic acid).

Table 3. Final candidates for BSBCA binding and experimental characterization, tentatively ranked in terms of theoretical viability. Columns 'Residue 1' and 'Residue 2' contain the single letter amino acid code, residue number, and position in the heptad repeat (***a***-***g***), respectively.

	ΔG_b Rank	ΔG_b Rank%	Residue 1	Residue 2
Candidate #1	37	9.8%	K-7-***b***	W-18-***f***
Candidate #2	120	31.7%	K-4-***f***	E-15-***c***
Candidate #3	78	20.6%	K-14-***b***	K-25-***f***

Analysis of Top Candidates Using the ROSIE Server. While energy gain of bundling seems to be a reasonable metric to predict hexamer bundle stability, this methodology does not encompass every variable. One potential issue is that there may be an oligomer state that is preferred more than a hexamer. To study this possibility, Symmetric Docking protocol available on the online ROSIE server was used to investigate the top three candidates.[91,93] This computational tool is used to predict the structure of symmetrical protein assemblies. On the user interface, a single chain is input, and the oligomer state is specified. The program then attempts to find the most stable conformation of side chains as well as the orientation in three-dimensional space without moving backbone torsion angles. It performs these calculations by sampling a

large sample size of possible conformations (n = 10,000) in a gradient based, Monte Carlo minimization.[93]

The relative stability of each oligomer state can be approximated by taking the score of the top predicted structure then dividing it by the number of chains. This 'energy per chain' metric allows structures with different number of chains to be evaluated against each other with a comparable calculation. Each of the three candidates were tested by scanning the oligomer states from 2-10, then the energy functions were evaluated. It was found that for all three candidates, the six-monomer state was highly favored (Figure 40). For ***Candidate #1***, the seven-monomer state seems to be also favored, which could be a potential limitation to generation of a stable oligomer state with this specific sequence. Overall, these results support our in-house computational predictions and provide promising results for the viability of our candidate structures.

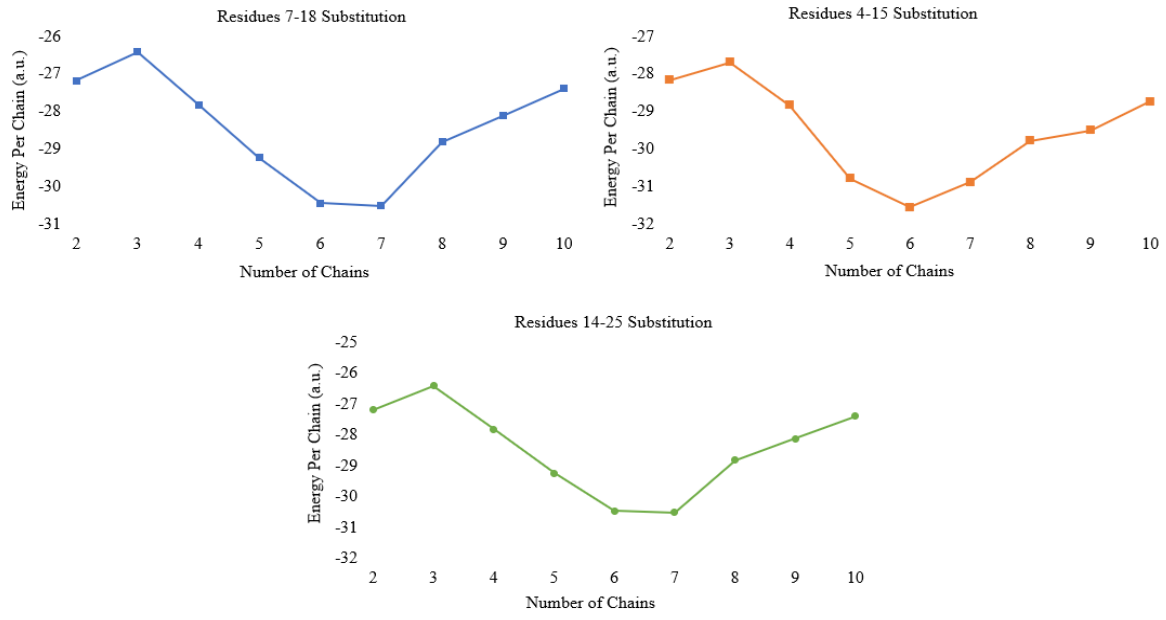

Figure 40. Results from Symmetric Docking protocol run on the ROSIE server.

Toward the Synthesis of 3,3'-bis(sulfonato)-4,4'-bis(chloroacetamido) azobenzene (BSBCA). 2-acetamido-5-nitrobenzenesulfonic acid (1). Initially, we attempted to synthesize BSBCA by following the procedure reported by Woolley and coworkers.[80] The first step of their procedure involved acetylation of 2,5-aminobenzenesulfonic acid. However, there were problems with unselective acetylation which gave multiple products. To circumvent this problem, we have decided to adapt a different procedure that were reported from Derda and coworkers and used a nitro and amino disubstituted compound as the starting material.[94]

After the reaction of 2-amino-5-nitrobenzenesulfonic acid with acetic anhydride under reflux conditions for 3 hours, it was found that the product become fully soluble in the solvent. After recrystallization of the compound, the final product, 2-acetamido-5-nitrobenzenesulfonic acid, could be obtained as a white, powdery solid (2.231 g, 91.1% yield). ^{1}H NMR verifies our synthesis was successful (Figure 41). The newly acetylated amino group (**1B**) shows significant deshielding due to the electron withdrawing effect of the nearby oxygen. The three aromatic hydrogens (**1E, 1D, 1C**) appear close together at 8.1-8.5 ppm. The three methyl hydrogens (**A**) are at 1.95 ppm. Residual acetic acid is present on the spectrum as well, even after multiple drying attempts. Solvent peak and water are observed as well, most likely attributed to wet deuterated solvent. Fortunately, there is no reason to believe this will have detrimental impacts on future reactions. ^{1}H NMR: (400 MHz, DMSO-d_6) δ 11.92, 10.64 (s, 1 H, N**H**), 8.54 (d, 1 H, C**H**), 8.42 (d, 1 H, C**H**), 8.21 (d, 1 H, C**H**), 3.33 (DMSO), 2.47 (H_2O), 2.21, 1.92 (s, 3 H, C**H**$_3$) ppm

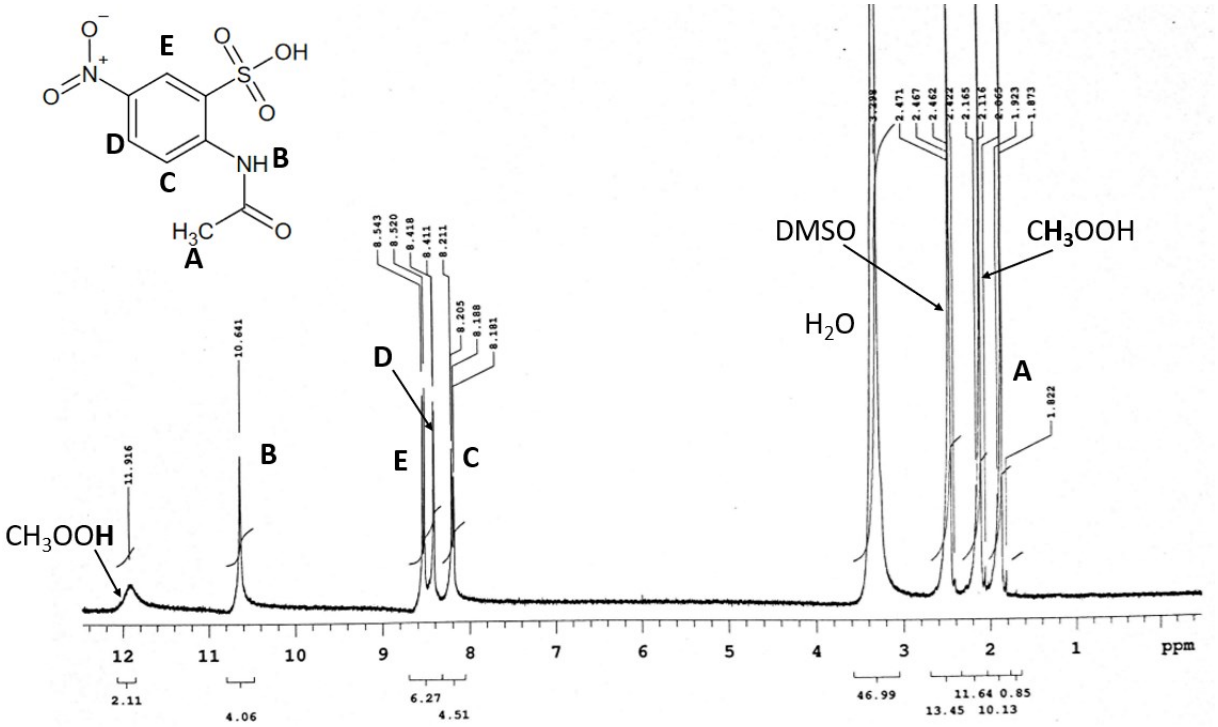

Figure 41. Labeled ^{1}H NMR spectrum of 2-acetamido-5-nitrobenzenesulfonic acid.

2-acetamido-5-aminobenzenesulfonic acid (2). In the second step of this synthetic scheme, the nitro group is reduced to an amine to prepare the next intermediate. The hydrogenation reaction catalyzed by Pd/C 10% yielded the product, 2-acetamido-5-aminobenzenesulfonic acid, as a light brown, powdery solid (0.688 g, 77.8% yield).

^{1}H NMR results confirm the reaction was successful (Figure 42). A peak at 4.88 ppm is observed, corresponding to the newly generated amino group (**2F**). Aromatic peaks are more spread out and shifted downstream, indicating that they are more shielded (**2E**, **2C**, **2E**). This is likely due to the removal of the highly electron withdrawing nitro group. The acetamido peak and methyl group are at 9.96 ppm and 1.93 ppm, respectively (**2B**, **2A**). Solvent and water peaks are seen again, further confirming wet NMR solvent. ^{1}H NMR: (400 MHz, DMSO-d_6): δ 9.96(s, 1 H, N**H**), 7.86(d, 1 H, C**H**), 6.93(d, 1 H, C**H**), 6.44(d, 1 H, C**H**), 4.88(s, 2H, N**H**$_2$), 3.33(DMSO), 2.47(H_2O), 1.93(s, 3 H, C**H**$_3$) ppm.

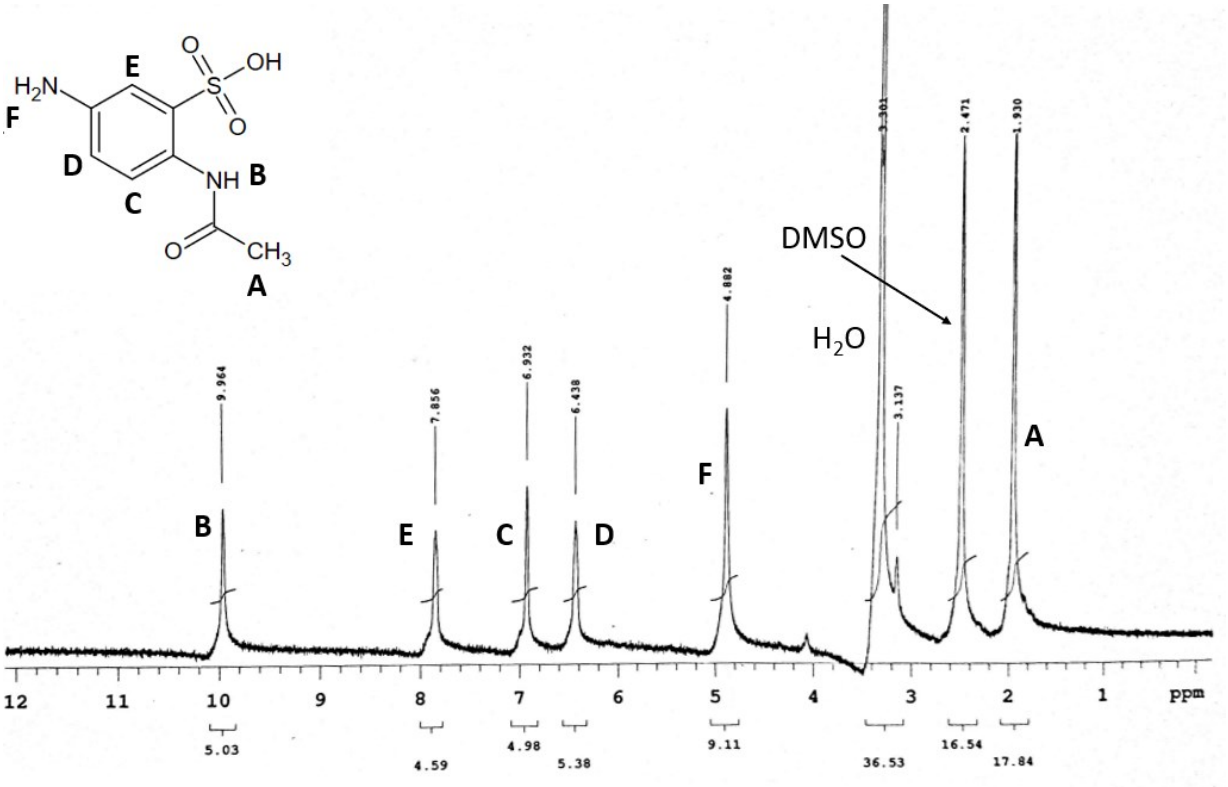

Figure 42. Labeled ^{1}H NMR spectrum of 2-acetamido-5-aminobenzenesulfonic acid.

Future Work to Be Done. To further advance this project, there are several tasks to work towards next. First, synthesis of BSBCA needs to be completed. The first two steps of our 5-step synthetic scheme have been completed with high yield. The third step is the primary bottleneck in the synthetic route, with previously reported literature yields being below 15%. Not only are the yields low, but compound separation is relatively difficult since both the starting reactant and product are charged species and highly polar. The last two steps have been reported to proceed with relatively high yield.

With computational characterization complete and the three top candidate structures identified, the next step is to synthesize the three sequences using solid phase peptide synthesis. Before any concrete progress can be made on binding the linker, the structures need to be further

characterized. To confirm the synthesis was successful, mass spectrometry should be performed on each sample to check molecular weight. In theory these structures should spontaneously self-assemble into stable hexameric forms in aqueous phase, however several tests need to be done to confirm this before BSBCA binding can be considered. Circular dichroism should be performed to ensure that alpha helices are forming, and oligomer state can be quantitatively measured using analytical ultracentrifugation.[71] If results from these characterization steps indicate the formation of a stable hexamer, then it is appropriate to proceed to BSBCA binding. If coupling is successful, the final objective is to characterize the photoresponsive nature of the bundle. A variety of approaches could be viable. Circular dichroism could be used to evaluate whether a change in helix conformation is occurring. Since the structure does contain an internal pore there is also a possibility to monitor bundle formation spectroscopically through the incorporation of hydrophobic fluorophores.

References

(42) Dobson, C. M. Protein Folding and Misfolding. **2003**, *426*, 7.

(43) Onuchic, J. N.; Wolynes, P. G. Theory of Protein Folding. *Curr. Opin. Struct. Biol.* **2004**, *14*, 70–75.

(44) Dill, K. A. Dominant Forces in Protein Folding. *Biochemistry*. **1990**, *29*, 7133-7155.

(45) Nick Pace, C.; Martin Scholtz, J. A Helix Propensity Scale Based on Experimental Studies of Peptides and Proteins. *Biophys. J.* **1998**, *75*, 422–427.

(46) Cohen, C.; Parry, D. A. D. α-Helical Coiled Coils — a Widespread Motif in Proteins. *Trends Biochem. Sci.* **1986**, *11*, 245–248.

(47) Scholtz, J. M.; Baldwin, R. L. The Mechanism of Alpha-Helix Formation by Peptides. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 95–118.

(48) Lupas, A. N.; Bassler, J. Coiled Coils – A Model System for the 21st Century. *Trends Biochem. Sci.* **2017**, *42*, 130–140.

(49) Indelicato, G.; Wahome, N.; Ringler, P.; Müller, S. A.; Nieh, M.-P.; Burkhard, P.; Twarock, R. Principles Governing the Self-Assembly of Coiled-Coil Protein Nanoparticles. *Biophys. J.* **2016**, *110*, 646–660.

(50) Zhang, F.; Timm, K. A.; Arndt, K. M.; Woolley, G. A. Photocontrol of Coiled-Coil Proteins in Living Cells. *Angew. Chem.* **2010**, *122*, 4035–4038.

(51) Thomson, A. R.; Wood, C. W.; Burton, A. J.; Bartlett, G. J.; Sessions, R. B.; Brady, R. L.; Woolfson, D. N. Computational Design of Water-Soluble α-Helical Barrels. *Science* **2014**, *346*, 485–488.

(52) Yu, Y. B. Coiled-Coils: Stability, Specificity, and Drug Delivery Potential. *Adv. Drug Deliv. Rev.* **2002**, *54*, 1113–1129.

(53) Liu, J.; Rost, B. Comparing Function and Structure between Entire Proteomes. *Protein Sci.* **2001**, *10*, 1970–1979.

(54) Rose, A.; Schraegle, S. J.; Stahlberg, E. A.; Meier, I. Coiled-Coil Protein Composition of 22 Proteomes – Differences and Common Themes in Subcellular Infrastructure and Traffic Control. *BMC Evol. Biol.* **2005**, *5*, 66.

(55) Landschulz, W.; Johnson, P.; McKnight, S. The Leucine Zipper: A Hypothetical Structure Common to a New Class of DNA Binding Proteins. *Science* **1988**, *240*, 1759–1764.

(56) O'Shea, E. K.; Lumb, K. J.; Kim, P. S. Peptide 'Velcro': Design of a Heterodimeric Coiled Coil. *Curr. Biol.* **1993**, *3*, 658–667.

(57) Woolfson, D. N. The Design of Coiled-Coil Structures and Assemblies. *Advances in Protein Chemistry*. **2005**, *70*, 79–112.

(58) Lupas, A. Coiled Coils: New Structures and New Functions. *Trends Biochem. Sci.* **1996**, *21*, 375-82.

(59) Woolfson, D. N. Coiled-Coil Design: Updated and Upgraded. In *Fibrous Proteins: Structures and Mechanisms*. **2017**, *82*, 35–61.

(60) Woolfson, D. N.; Bartlett, G. J.; Bruning, M.; Thomson, A. R. New Currency for Old Rope: From Coiled-Coil Assemblies to α-Helical Barrels. *Curr. Opin. Struct. Biol.* **2012**, *22*, 432–441.

(61) Sun, L.; Young, L. N.; Zhang, X.; Boudko, S. P.; Fokine, A.; Zbornik, E.; Roznowski, A. P.; Molineux, I. J.; Rossmann, M. G.; Fane, B. A. Icosahedral Bacteriophage ΦX174 Forms a Tail for DNA Transport during Infection. *Nature* **2014**, *505*, 432–435.

(62) Yon, J. M. Protein Folding: A Perspective for Biology, Medicine and Biotechnology. *Braz. J. Med. Biol. Res.* **2001**, *34*, 419–435.

(63) Khoury, G. A.; Smadbeck, J.; Kieslich, C. A.; Floudas, C. A. Protein Folding and de Novo Protein Design for Biotechnological Applications. *Trends Biotechnol.* **2014**, *32*, 99–109.

(64) Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. The Protein Folding Problem. *Annu. Rev. Biophys.* **2008**, *37*, 289–316.

(65) Fairlie, D. P.; Dantas de Araujo, A. Stapling Peptides Using Cysteine Crosslinking: Stapling Peptides. *Biopolymers* **2016**, *106*, 843–852.

(66) O'Shea, E. K.; Klemm, J. D.; KiIM, P. S.; Alber, T. X-Ray Structure of the GCN4 Leucine Zipper, a Two-Stranded, Parallel Coiled Coil. **2021**, *254*, 7.

(67) Ogihara, N. L.; Weiss, M. S.; Eisenberg, D.; Degrado, W. F. The Crystal Structure of the Designed Trimeric Coiled Coil Coil-V $_a$ L $_d$: Implications for Engineering Crystals and Supramolecular Assemblies: Structure of a Designed Parallel Three-Helix Coiled Coil Protein. *Protein Sci.* **1997**, *6*, 80–88.

(68) Lombardi, Angela; Degrado, William. De Novo Design of Heterotrimeric Coiled Coils. **1996**. *40*. 495-504.

(69) Bromley, E. H. C.; Sessions, R. B.; Thomson, A. R.; Woolfson, D. N. Designed α-Helical Tectons for Constructing Multicomponent Synthetic Biological Systems. *J. Am. Chem. Soc.* **2009**, *131*, 928–930.

(70) Testa, O. D.; Moutevelis, E.; Woolfson, D. N. CC+: A Relational Database of Coiled-Coil Structures. *Nucleic Acids Res.* **2009**, *37*, 315–322.

(71) Thomas, F.; Dawson, W. M.; Lang, E. J. M.; Burton, A. J.; Bartlett, G. J.; Rhys, G. G.; Mulholland, A. J.; Woolfson, D. N. *De Novo* -Designed α-Helical Barrels as Receptors for Small Molecules. *ACS Synth. Biol.* **2018**, *7*, 1808–1816.

(72) Scott, A. J.; Niitsu, A.; Kratochvil, H. T.; Lang, E. J. M.; Sengel, J. T.; Dawson, W. M.; Mahendran, K. R.; Mravic, M.; Thomson, A. R.; Brady, R. L.; Liu, L.; Mulholland, A. J.; Bayley, H.; DeGrado, W. F.; Wallace, M. I.; Woolfson, D. N. Constructing Ion Channels from Water-Soluble α-Helical Barrels. *Nat. Chem.* **2021**.

(73) Sauerborn, M. Immunological Aspects of Antibody Formation against Recombinant Human Therapeutics, 2010.

(74) Bandara, H. M. D.; Burdette, S. C. Photoisomerization in Different Classes of Azobenzene. *Chem Soc Rev* **2012**, *41*, 1809–1825.

(75) Beharry, A. A.; Woolley, G. A. Azobenzene Photoswitches for Biomolecules. *Chem. Soc. Rev.* **2011**, *40*, 4422.

(76) Samanta, S.; Woolley, G. A. Bis-Azobenzene Crosslinkers for Photocontrol of Peptide Structure. *ChemBioChem* **2011**, *12*, 1712–1723.

(77) Beharry, A. A.; Sadovski, O.; Woolley, G. A. Azobenzene Photoswitching without Ultraviolet Light. *J. Am. Chem. Soc.* **2011**, *133*, 19684–19687.

(78) Mahimwalla, Z.; Yager, K. G.; Mamiya, J.; Shishido, A.; Priimagi, A.; Barrett, C. J. Azobenzene Photomechanics: Prospects and Potential Applications. *Polym. Bull.* **2012**, *69*, 967–1006.

(79) Aoki, K.; Nakagawa, M.; Ichimura, K. Self-Assembly of Amphoteric Azopyridine Carboxylic Acids: Organized Structures and Macroscopic Organized Morphology Influenced by Heat, PH Change, and Light. *J. Am. Chem. Soc.* **2000**, *122*, 10997–11004.

(80) Burns, D. C.; Zhang, F.; Woolley, G. A. Synthesis of 3,3′-Bis(Sulfonato)-4,4′-Bis(Chloroacetamido)Azobenzene and Cysteine Cross-Linking for Photo-Control of Protein Conformation and Activity. *Nat. Protoc.* **2007**, *2*, 251–258..

(81) Fissi, A.; Pieroni, O.; Ciardelli, F. Photoresponsive Polymers: Azobenzene-Containing Poly(L-Lysine). *Biopolymers* **1987**, *26*, 1993–2007.

(82) Zimmer, M. Green Fluorescent Protein (GFP): Applications, Structure, and Related Photophysical Behavior. *Chem. Rev.* **2002**, *102*, 759–782.

(83) Oprea, T. I.; May, E. E.; Leitão, A.; Tropsha, A. Computational Systems Chemical Biology. In *Chemoinformatics and Computational Chemical Biology*; Bajorath, J., Ed.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2011; Vol. 672, pp 459–488.

(84) Rohl, C. A.; Strauss, C. E. M.; Misura, K. M. S.; Baker, D. Protein Structure Prediction Using Rosetta. In *Methods in Enzymology*; Elsevier, 2004; Vol. 383, pp 66–93.

(85) Kuhlman, B.; Bradley, P. Advances in Protein Structure Prediction and Design. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 681–697.

(86) Kryshtafovych, A.; Schwede, T.; Topf, M.; Fidelis, K.; Moult, J. Critical Assessment of Methods of Protein Structure Prediction (CASP)—Round XIII. *Proteins Struct. Funct. Bioinforma.* **2019**, *87*, 1011–1020.

(87) Senior, A. W.; Evans, R.; Jumper, J.; Kirkpatrick, J.; Sifre, L.; Green, T.; Qin, C.; Žídek, A.; Nelson, A. W. R.; Bridgland, A.; Penedones, H.; Petersen, S.; Simonyan, K.; Crossan, S.; Kohli, P.; Jones, D. T.; Silver, D.; Kavukcuoglu, K.; Hassabis, D. Protein Structure Prediction Using Multiple Deep Neural Networks in the 13th Critical Assessment of Protein Structure Prediction (CASP13). *Proteins Struct. Funct. Bioinforma.* **2019**, *87*, 1141–1148.

(88) Kaufmann, K. W.; Lemmon, G. H.; DeLuca, S. L.; Sheehan, J. H.; Meiler, J. Practically Useful: What the Rosetta Protein Modeling Suite Can Do for You. *Biochemistry* **2010**, *49*, 2987–2998.

(89) Alford, R. F.; Leaver-Fay, A.; Jeliazkov, J. R.; O'Meara, M. J.; DiMaio, F. P.; Park, H.; Shapovalov, M. V.; Renfrew, P. D.; Mulligan, V. K.; Kappel, K.; Labonte, J. W.; Pacella, M. S.; Bonneau, R.; Bradley, P.; Dunbrack, R. L.; Das, R.; Baker, D.; Kuhlman, B.; Kortemme, T.; Gray, J. J. The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. *J. Chem. Theory Comput.* **2017**, *13*, 3031–3048.

(90) Chaudhury, S.; Lyskov, S.; Gray, J. J. PyRosetta: A Script-Based Interface for Implementing Molecular Modeling Algorithms Using Rosetta. *Bioinformatics* **2010**, *26*, 689–691.

(91) Lyskov, S.; Chou, F.-C.; Conchúir, S. Ó.; Der, B. S.; Drew, K.; Kuroda, D.; Xu, J.; Weitzner, B. D.; Renfrew, P. D.; Sripakdeevong, P.; Borgo, B.; Havranek, J. J.; Kuhlman, B.; Kortemme, T.; Bonneau, R.; Gray, J. J.; Das, R. Serverification of Molecular Modeling Applications: The Rosetta Online Server That Includes Everyone (ROSIE). *PLoS ONE* **2013**, *8*, e63906.

(92) Moretti, R.; Lyskov, S.; Das, R.; Meiler, J.; Gray, J. J. Web-Accessible Molecular Modeling with Rosetta: The Rosetta Online Server That Includes Everyone (ROSIE): ROSIE: Web-Accessible Modeling with Rosetta. *Protein Sci.* **2018**, *27*, 259–268.

(93) Andre, I.; Bradley, P.; Wang, C.; Baker, D. Prediction of the Structure of Symmetrical Protein Assemblies. *Proc. Natl. Acad. Sci.* **2007**, *104*, 17656–17661.

(94) Jafari, M. R.; Deng, L.; Kitov, P. I.; Ng, S.; Matochko, W. L.; Tjhung, K. F.; Zeberoff, A.; Elias, A.; Klassen, J. S.; Derda, R. Discovery of Light-Responsive Ligands through Screening of a Light-Responsive Genetically Encoded Library. *ACS Chem. Biol.* **2014**, *9*, 443–450.

(95) Kuhlman, B.; Baker, D. Native Protein Sequences Are Close to Optimal for Their Structures. *Proc. Natl. Acad. Sci.* **2000**, *97*, 10383–10388.

SUMMARY

The overarching goal for my Master thesis work was to develop two smart materials that are capable of encapsulating small molecules and will readily release the 'cargo' in response to external stimuli. In the first chapter, I described the study on the development and characterization of DiO/DiI containing PEG-*b*-PLA block copolymer micelles that respond to vortexed-induced mechanical-stresses. Preliminary studies were performed to optimize the micellar system. The ratio between DiO/DiI and the dye/polymer ratio were optimized to increase FRET efficiency. It was found that using DiO/DiI in a 1:5 ratio led to a high FRET yield. To investigate the effects of vortex-induced mechanical stress on fluorescence emissions, dye-containing block copolymer micelle solutions were subjected to vortexing followed by the measurement of fluorescence spectra. It was found that exposure to vortexing-induced mechanical-stress causes a decrease in fluorescence emissions. This phenomenon showed a dependence on both rotational speed and the time of vortexing. When solutions were vortexed at speeds under 1000 rpm there is little change in fluorescence behavior. However, when this threshold is crossed, there is a significant decrease in emissions. To further investigate the effects of vortexing-induced mechanical-stresses on micelle morphology, transmission electron microscopy (TEM) analysis was performed. It was found that vortexing increased the average size of the micelles, indicating that aggregation, fusion, and/or chain exchange is occurring. These processes could trigger the release of dyes from the micelle core, leading to aggregation and self-quenching. Ultimately, this work adds valuable fundamental knowledge of block copolymer micelle systems and establishes the potential utility of the dye-containing block copolymer micelles as a nanoprobe for the optical detection of flow-induced mechanical-stresses.

In the second chapter I described the work toward designing a photo-responsive hexameric peptide bundle which is capable of internalizing small molecules and reversibly forming/deforming a hexameric state in response to light. As the photo-responsive linker, BSBCA, is not available commercially, I have set out to synthesis this compound and the first two steps of the synthesis have been completed. In parallel , I worked on identifying the two amino acids residues that can be substituted by cysteine residues for the covalent attachment of BSBCA linker. In order to aid the screening of appropriate locations for introducing cysteine residues, computational methods were employed. Using the PyRosetta4 computational suite and in-house scripts we estimated the energy function of single- and double-mutants and the energy gain by the hexamer bundle formation by double-mutants. Computational results suggested that mutations at internal hydrophobic residues and alanine destabilize the peptide bundle, while substitution on the exterior of the peptide face are more energetically favored. Out of an initial search of 387 structures, three potential candidates have been identified. Since the synthesis of the photo-responsive linker has not be completed, further experiments are required before the synthesis of the photo-responsive peptide bundles. However, the current work has allowed for paving the path toward accomplishing the end goal of synthesizing and characterizing the properties of the photo-responsive linker staple hexameric peptide bundles that can reversibly ‘catch and release’ small molecular ligand in response to the irradiation of ultraviolet light.

APPENDICES

Appendix A

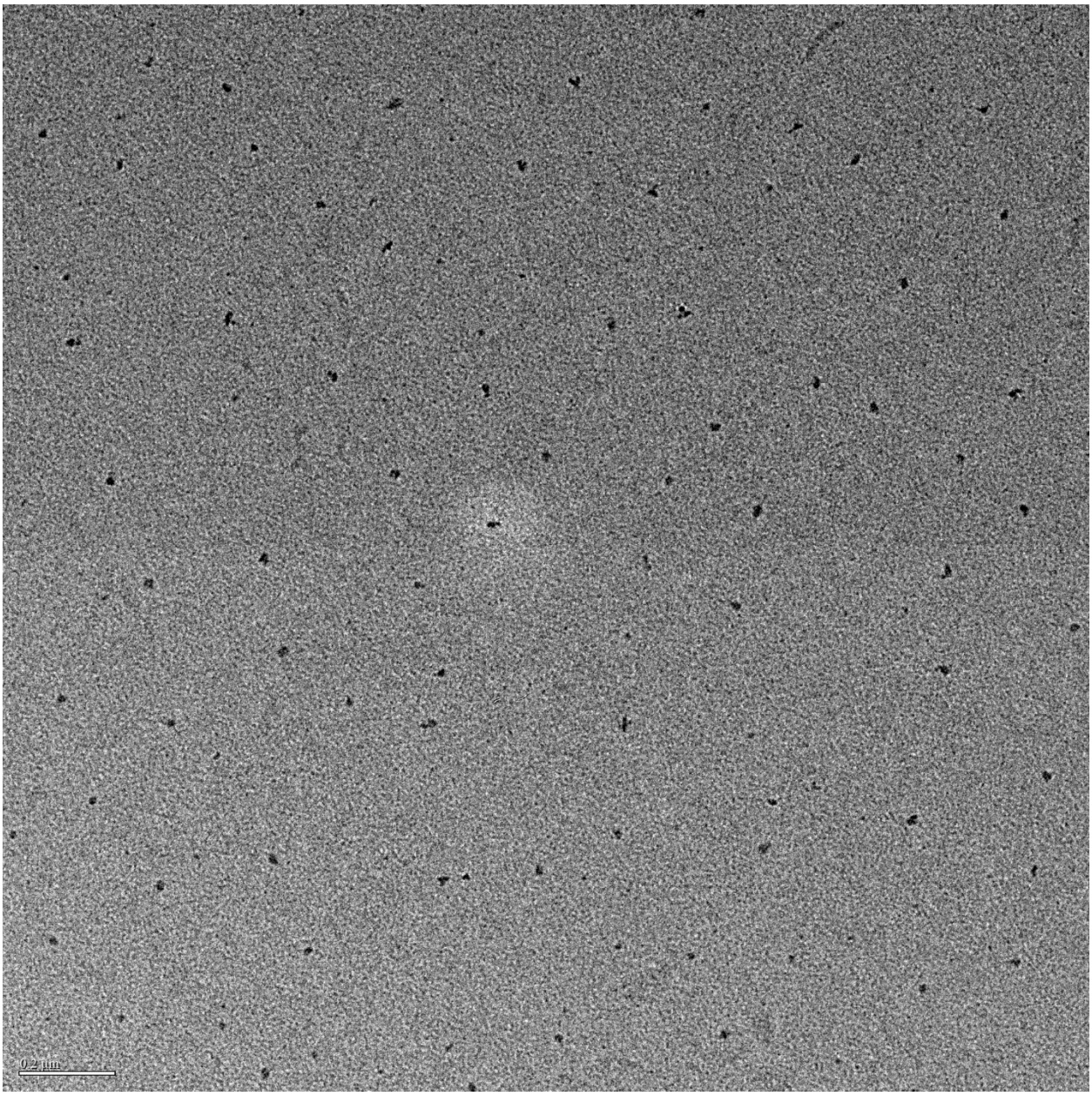

Figure A.1. TEM Image of PEG-*b*-PLA block copolymer micelles before exposure to vortexing. Scale bar is equal to 200 nm.

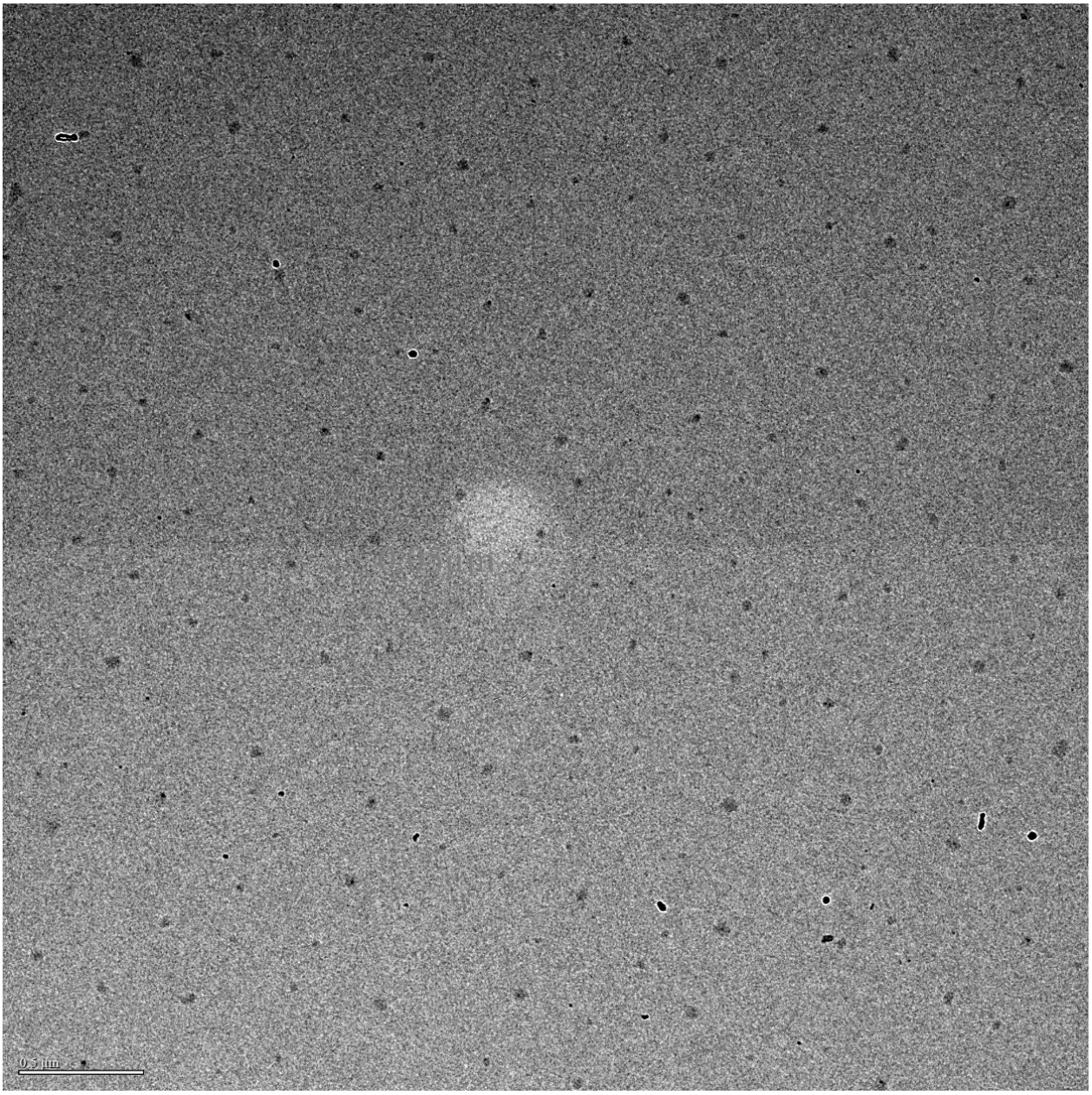

Figure A.2. TEM image of PEG-*b*-PLA block copolymer micelles after exposure to 5 minutes of vortexing at 3,000 rpm. Scale bar is equal to 500 nm.

Appendix B

Single Mutant Python Script.

```
# initialize pyrosetta
from pyrosetta import *
init()

# imports
from pyrosetta.toolbox import mutate_residue
from pyrosetta.teaching import *
from rosetta.protocols.simple_moves import *
import csv
scorefxn = get_fa_scorefxn()

# assignment of pose
pose =
pose_from_pdb('/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/pdb_files/CC_h
ex2.clean.edited_monomer.pdb')
original_pose =
pose_from_pdb('/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/pdb_files/CC_h
ex2.clean.edited_monomer.pdb')
init_score = scorefxn(original_pose)

# object oriented mutation of b, c, f residues with cysteine

res_to_mutate = []
res_in_chain = 29

for r in range(1, 29):
    res_to_mutate.append(r)

# mut_num is residue that is being mutated
# mut_aa is amino acid that residue is being switched too

mut_res_num1 = ['Residue Number']
mutation_name1 = ['Mutation Info']
scores = ['Scores With Minimization']
energy_change_pack = ['Energy Change With Minimization']


class SingleMutantMonomer:
    def __init__(self, chain_length, num_of_chain, mut_num1, mut_aa1):
        self.chain_length = chain_length
        self.num_of_chain = num_of_chain
        self.mut_num1 = mut_num1
        self.mut_aa1 = mut_aa1

    def mutation_info(self):
```

```
        if self.mut_num1 <= self.chain_length:
            mutation_name1.append(pose.residue(initial_mut_num1).name() + '_'
+ str(self.mut_num1) + '_' + self.mut_aa1)
            mut_res_num1.append(self.mut_num1)
        else:
            pass

    def mutate_pose(self):
        mutate_residue(pose, self.mut_num1, self.mut_aa1)

    def packing(self):
        mm1 = MinMover()
        # assign scorefunction to MinMover
        mm1.score_function(scorefxn)
        # generate movemap object/range and assign it to MinMover
        mm1_xx = MoveMap()

        if self.mut_num1 == 1:
            mm1_xx.set_bb_true_range(self.mut_num1, (self.mut_num1 + 1))
        elif self.mut_num1 == 28:
            mm1_xx.set_bb_true_range(self.mut_num1 - 1, self.mut_num1)
        else:
            mm1_xx.set_bb_true_range((self.mut_num1 - 1), (self.mut_num1 +
1))

        mm1.movemap(mm1_xx)
        # apply minimization protocol to mutated pose
        mm1.apply(pose)

    def calc_energy(self):
        scores.append(scorefxn(pose))
        energy_change_pack.append(scorefxn(pose) - init_score)


for k in res_to_mutate:
    Hex2_smm = SingleMutantMonomer(29, 1, k, 'C')

    initial_mut_num1 = Hex2_smm.mut_num1

    Hex2_smm.mutation_info()
    Hex2_smm.mutate_pose()
    Hex2_smm.packing()
    Hex2_smm.calc_energy()
    pose = pose_from_pdb('file_thread.pdb')


def csv_export():
    with
open('/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/PyRosetta_results/cyste
ine_sub_data_monomer_single_3min.csv', 'w', newline='') as e:
```

```
        write = csv.writer(e)

        write.writerow(['Original score =', init_score])
        write.writerow(mutation_name1)
        write.writerow(energy_change)
        write.writerow(scores)
        write.writerow(mut_res_num1)
print("Write <csv_export()> to export results.")
```

Double Mutant Python Script.

```
# initialize pyrosetta
from pyrosetta import *
init()

# imports
from pyrosetta.toolbox import mutate_residue
from pyrosetta.teaching import *
from rosetta.protocols.simple_moves import *
import csv
import re
import os
scorefxn = get_fa_scorefxn()

# assignment of pose
pose =
pose_from_pdb('/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/pdb_files/CC_h
ex2.clean.edited2.pdb')
original_pose =
pose_from_pdb('/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/pdb_files/CC_h
ex2.clean.edited2.pdb')
init_score_hex = scorefxn(original_pose)
init_score_mon =
scorefxn(pose_from_pdb('/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/pdb_f
iles/CC_hex2.clean.edited_monomer.pdb'))

# object oriented mutation of b, c, f residues with cysteine

res_to_mutate = []
res_in_chain = 29

for r in range(0, 5):
    for j in range(0, 5):
        c = (1 + 7 * j + res_in_chain * r)
        d = (2 + 7 * j + res_in_chain * r)
        e = (3 + 7 * j + res_in_chain * r)
        f = (4 + 7 * j + res_in_chain * r)
        g = (5 + 7 * j + res_in_chain * r)
```

```
        a = (6 + 7 * j + res_in_chain * r)
        b = (7 + 7 * j + res_in_chain * r)
        if c < res_in_chain:
            res_to_mutate.append(c)
        if d < res_in_chain:
            res_to_mutate.append(d)
        if e < res_in_chain:
            res_to_mutate.append(e)
        if f < res_in_chain:
            res_to_mutate.append(f)
        if g < res_in_chain:
            res_to_mutate.append(g)
        if a < res_in_chain:
            res_to_mutate.append(a)
        if b < res_in_chain:
            res_to_mutate.append(b)

print(res_to_mutate)

# mut_num is residue that is being mutated
# mut_aa is amino acid that residue is being switched too

mut_res_num1 = ['Residue Number']
mutation_name1 = ['Mutation Info']
scores_hex = ['Scores Hexamer']
energy_change_hex = ['Energy Change Hexamer']
scores_mon =['Scores Monomer']
energy_change_mon = ['Energy Change Monomer']

pdb_hex_text = []
pdb_mon_text = []



class SingleMutantHexamer:
    def __init__(self, chain_length, num_of_chain, mut_num1, mut_aa1):
        self.chain_length = chain_length
        self.num_of_chain = num_of_chain
        self.mut_num1 = mut_num1
        self.mut_aa1 = mut_aa1

    def mutation_info(self):
        if self.mut_num1 <= self.chain_length:
            mutation_name1.append(pose.residue(initial_mut_num1).name() + '_'
+ str(self.mut_num1) + '_' + self.mut_aa1)
            mut_res_num1.append(self.mut_num1)
        else:
            pass

    def mutate_pose(self):
```

```
        mutate_residue(pose, self.mut_num1, self.mut_aa1)

    def packing(self):
        mm1 = MinMover()
        # assign scorefunction to MinMover
        mm1.score_function(scorefxn)
        # generate movemap object/range and assign it to MinMover
        mm1_xx = MoveMap()

        if self.mut_num1 == 1:
            mm1_xx.set_bb_true_range(self.mut_num1, (self.mut_num1 + 1))
        elif self.mut_num1 == 29:
            mm1_xx.set_bb_true_range((self.mut_num1 - 1), self.mut_num1)
        else:
            mm1_xx.set_bb_true_range((self.mut_num1 - 1), (self.mut_num1 +
1))

        mm1.movemap(mm1_xx)
        # apply minimization protocol to mutated pose
        mm1.apply(pose)

    def calc_hex_energy(self):
        scores_hex.append(scorefxn(pose))
        energy_change_hex.append(scorefxn(pose) - init_score_hex)

    def move_to_next_chain(self):
        self.mut_num1 = self.mut_num1 + self.chain_length
        if self.mut_num1 <= 173:
            print(self.mut_num1, pose.residue(self.mut_num1).name())
        else:
            pass

    def apply_to_all_chains(self):
        for x in range(0, self.num_of_chain):
            self.mutate_pose()
            self.packing()
            self.move_to_next_chain()

    def export_pdb(self):

pose.dump_pdb('/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/pdb_files/CC_h
ex2.hex_mutant.pdb')

os.rename(r'/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/pdb_files/CC_hex2
.hex_mutant.pdb',

r'/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/pdb_files/CC_hex2.hex_mutan
t.txt')

    def pdb_to_text(self):
```

```
        with
open('/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/pdb_files/CC_hex2.hex_m
utant.txt', 'r') as m:
            for line in m:
                pdb_hex_text.append(line.strip())

    def loop_pdb_text(self):
        for i in pdb_hex_text:
            match = re.search('TER', i)
            if match:
                pdb_mon_text.append(i)
                break
            else:
                pdb_mon_text.append(i)

    def write_mon_file(self):
        with
open('/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/pdb_files/CC_hex2.mon_m
utant.txt', 'w') as w:
            for i in pdb_mon_text:
                w.write(i + '\n')

    def calc_mon_energy(self):
        pose =
pose_from_pdb('/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/pdb_files/CC_h
ex2.mon_mutant.txt')
        scores_mon.append(scorefxn(pose))
        energy_change_mon.append(scorefxn(pose) - init_score_mon)




for k in res_to_mutate:
    Hex2_smh = SingleMutantHexamer(29, 6, k, 'C')

    initial_mut_num1 = Hex2_smh.mut_num1

    print(initial_mut_num1, pose.residue(initial_mut_num1).name())
    Hex2_smh.mutation_info()
    Hex2_smh.apply_to_all_chains()
    Hex2_smh.calc_hex_energy()
    Hex2_smh.export_pdb()

    # functions below are for monomer generation and evaluation
    Hex2_smh.pdb_to_text()
    Hex2_smh.loop_pdb_text()
    Hex2_smh.write_mon_file()
    Hex2_smh.calc_mon_energy()

    pdb_hex_text.clear()
```

```
    pdb_mon_text.clear()
    pose =
pose_from_pdb('/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/pdb_files/CC_h
ex2.clean.edited2.pdb')


def csv_export():
    with
open('/mnt/c/Users/Tyler/Dropbox/2018_Tyler_Odom/Code/PyRosetta_results/cyste
ine_sub_data_hex_to_mon.csv', 'w', newline='') as e:
        write = csv.writer(e)

        write.writerow(['Mutating all residues in Hex2 bundle then converting
to monomer. Adjacent residue minimization'])
        write.writerow(['Init Hex Score =', init_score_hex])
        write.writerow(['Init Mon Score =', init_score_mon])
        write.writerow(mutation_name1)
        write.writerow(mut_res_num1)
        write.writerow(scores_hex)
        write.writerow(energy_change_hex)
        write.writerow(scores_mon)
        write.writerow(energy_change_mon)



print("Write <csv_export()> to export results.")
```

Heat Map Generation Script.

```
import openpyxl
from openpyxl.utils import get_column_letter

wb =
openpyxl.load_workbook(r'C:\Users\tyler\Dropbox\2018_Tyler_Odom\Code\PyRosett
a_results\00_compiled_csv_data.xlsx')
sheet = wb['hex_to_mon_double']


raw_res1_list = []
heat_map_column_list = []
raw_res2_list = []
heat_map_row_list = []
scores_list = []

res_scores = []

def generate_lists():
    for raw_res1 in sheet['B7':'NO7']:
```

```
        for cell1 in raw_res1:
            raw_res1_list.append(cell1.coordinate)

    for raw_res2 in sheet['B6':'NO6']:
        for cell2 in raw_res2:
            raw_res2_list.append(cell2.coordinate)

    for heat_map_column in sheet['J74':'AK74']:
        for cell_hm1 in heat_map_column:
            heat_map_column_list.append(cell_hm1.coordinate)
        # if sheet[heat_map_column_res].value == sheet[raw_res1].value:

    for heat_map_row in sheet['AL75':'AL102']:
        for cell_hm2 in heat_map_row:
            heat_map_row_list.append(cell_hm2.coordinate)

    for scores in sheet['B12':'NO12']:
        for cell_s in scores:
            scores_list.append(cell_s.coordinate)

    res_scores.append(raw_res1_list)
    res_scores.append(raw_res2_list)
    res_scores.append(scores_list)


# loops through raw data then set-up chart that contains no residue
redundancies + correct formatting
def make_heat_map():
    for i in range(0, (sheet.max_column - 1)):
        for j in heat_map_row_list:
            if sheet[res_scores[0][i]].value == sheet[j].value:
                for y in heat_map_column_list:
                    if sheet[res_scores[1][i]].value == sheet[y].value:
                        score_column = sheet[y].column
                        score_row = sheet[j].row
                        score_change = sheet[res_scores[2][i]].value
                        sheet[str(get_column_letter(score_column)) +
str(score_row)] = '=' + str(score_change)
                        score_cell = str(get_column_letter(score_column)) +
str(score_row)
                    else:
                        pass
            else:
                pass

def save_file():

wb.save(r'C:\Users\tyler\Dropbox\2018_Tyler_Odom\Code\PyRosetta_results\00_co
mpiled_csv_data.xlsx')
```

```
generate_lists()
make_heat_map()
save_file()
```